Problem formulation and hypothesis testing for environmental risk assessments of genetically modified crops

Alan RAYBOULD*

Syngenta, Jealott's Hill International Research Centre, Bracknell, Berkshire RG42 6EY, UK

Environmental risk assessments can provide high confidence of minimal risk by testing theories, "risk hypotheses", that predict the likelihood of unacceptable harmful events. The creation of risk hypotheses and a plan to test them is called problem formulation. Effective problem formulation seeks to maximize the possibility of detecting effects that indicate potential risk; if such effects are not detected, minimal risk is indicated with high confidence. Two important implications are that artificial test conditions can increase confidence, whereas prescriptive data requirements can reduce confidence (increase uncertainty) if they constrain problem formulation. Poor problem formulation can increase environmental risk because it leads to the collection of superfluous data that may delay or prevent the introduction of environmentally beneficial products.

Keywords: risk assessment / problem formulation / scientific method / certainty / data requirements

INTRODUCTION

Genetically modified (GM) crops are subject to intense regulatory scrutiny. Although in theory some countries only regulate certain classes of GM crop, in practice all GM crops for uncontained use must be licensed before being placed on the market (Miller and Conko, 2004). Decisions to license a GM crop for commercial cultivation are based on risk analysis, a judgment about whether the risk from use of the GM crop is acceptable. Risk analysis is often split into two activities: risk assessment, the scientific determination of the probability of specified harmful effects; and decision-making, the evaluation of whether the risk, and the uncertainty associated with its estimation, is acceptable. Acceptability depends on the objectives of public policy, along with the ability to manage and communicate the risk (Wolt and Peterson, 2000).

The separation of risk assessment from the other aspects of risk analysis cannot be absolute because policy should determine which effects are considered in the risk assessment (Stern and Fineberg, 1996); however, most students of risk analysis consider risk assessment to be scientific, and as such it should follow the scientific method (*e.g.*, Patton, 1998; Power and Adams, 1997; Wolt and Petersen, 2000). This paper suggests some general principles for design of risk assessments in the light of current philosophy of the scientific method.

* Corresponding author: alan.raybould@syngenta.com

THE SCIENTIFIC METHOD

Until early in the Twentieth Century, science was seen as working by induction – the derivation of general truthful statements, "laws", from accumulated observations of specific instances of the actions of these laws. However, as Hume pointed out in the Eighteenth Century, induction has no basis in logic: if B follows A often enough, one may expect the next A will be followed by a B; however, "this is a fact of psychology, not logic" (Magee, 1985).

An acceptable solution to the problem of induction in science was offered by Popper (1959) who showed that falsification is logically different from proof: a single instance of B not following A falsifies the theory that "B follows A"; but any number of instances of B following A cannot prove a law that B follows A. Popper proposed that for science to have a sound basis in logic it should develop theories that make testable predictions about the world; the theories are tested by comparing observations with predictions, and if observations and predictions differ systematically, the theory is falsified and a new theory is developed that can account for the discrepancies.

Popper's philosophy revolutionizes our view of how science works. Knowledge increases by a cycle of formulation, testing, falsification and reformulation of theories such that predictions are made with increasing accuracy and precision. Science does not proceed by the revelation of true theories (laws) as observations accumulate.

THE SCIENTIFIC METHOD AND RISK ASSESSMENT

Popper's theory of the continuous development of objective knowledge has four stages:

$$P_1 \rightarrow TS \rightarrow EE \rightarrow P_2$$

P_1 is the initial problem; TS is a trial solution to the problem; EE is error elimination, in which the trail solution is evaluated by observation; and P_2 is a situation of increased knowledge. Knowledge development is continuous, as P_2 is an initial problem for a further round of hypothesis testing. An important part of this theory is that observation cannot be prior to theory, as one must have a theory in order to decide what to observe (Popper, 1959). Hence, science begins with a problem, not an attempt to solve the problem, and the sources of scientific problems are attempts to solve prior problems.

Risk assessment is simple in concept, and it conforms to Popper's ideas about how science works. The first step in a risk assessment is a decision about what needs protecting from harm; this step derives explicit and unambiguous targets for protection, called assessment endpoints, from the management goals of legislation or policy (Suter, 1990). The second step is an assessment of how the proposed action, such as the cultivation of GM crops, may cause harm; this step is often referred to as the development of risk hypotheses (*e.g.*, Patton, 1998). The next stage is the collection[1] of data to evaluate the risk hypotheses, and the evaluation of these hypotheses leads to increased knowledge of risk. These steps are in agreement with Popper's theory:

P_1 = derivation of assessment endpoints from the management goals;
TS = development of a risk hypothesis;
EE = testing of the risk hypothesis;
P_2 = increased knowledge of risk.

The derivation of assessment endpoints and development of the risk hypothesis are called problem formulation, and the testing of the risk hypothesis is called risk characterization (US Environmental Protection Agency, 1998). The increased knowledge of risk (P_2) may be such that a regulatory authority permits release of the crop without collection of more data. If P_2 does not indicate an acceptable level of risk with sufficient confidence, new risk hypotheses can be developed and tested, leading to a further increase in knowledge of risk. A regulatory decision to permit the release of a GM plant once a certain level of knowledge is reached need not curtail

[1] Throughout this paper, "collection" of data does not necessarily imply new experiments, because sufficient data to evaluate the risk hypothesis may already exist.

further cycles of hypothesis testing. For example, post-market monitoring could be considered to conform to this scheme, provided it tests clear hypotheses that are related to protection of assessment endpoints.

POOR PROBLEM FORMULATION IN GM RISK ASSESSMENTS

Part of the reason why risk assessment for GM crops remains controversial is that the simple tenets outlined above are not followed by research scientists studying GM crops. First, there is confusion over what should be protected (P_1 in Popper's scheme). In the UK Farm Scale Evaluations of GM herbicide tolerant (GMHT) crops, an assessment endpoint was the sustainability of populations of arable weeds in fields (Squire et al., 2003). The observed reductions in arable weed populations in some GMHT crops were considered detrimental effects, because weeds were considered to be valuable biodiversity (*e.g.*, Giles, 2003). However, weed control is an intended effect of agriculture, and under some scenarios it can increase biodiversity (Green et al., 2005) – hence confusion over whether weed populations should be protected. A similar confusion occurs over whether the reduction of predators and parasitoids associated with the intended reduction in abundance of the target insect pests of GM crops should be regarded as harm (Groot and Dicke, 2001). One of the objectives of farming is to control populations of pests in crops. The judgment of whether a particular change in species abundance following pest control constitutes harm should be independent of the method of farming, although the acceptability of that harm may vary among methods due to other criteria.

The next area of confusion is the development of a hypothesis about how the GM crops could harm the environment (TS). A common mistake is equating hazard or exposure with risk. For example, the presence of transgenes in populations of wild relatives, or residues of Cry proteins in soil, are often presented as risks; however, in the absence of a hazard associated with the transgene or the proteins, these phenomena constitute exposure only (Mendelsohn et al., 2003; Raybould and Wilkinson, 2005). Similarly, the observation that a transgenic insecticidal protein is toxic to a non-target organism only demonstrates a hazard; an evaluation of risk requires an assessment of the likelihood of exposure to amounts of protein that could cause harm (Sears et al., 2001).

Mechanisms by which GM crops are postulated to cause harm are often not unique to those organisms. Unintended effects of transformation are frequently cited as a means by which GM crops could be harmful; however, mutation breeding intentionally introduces such changes, but is not subject to regulatory control in most jurisdictions that regulate GM crops (Miller and Conko, 2004).

This suggests that there is no clear hypothesis about how unintended effects of transformation could be harmful. Finally, there is concern that there are "unknown unknowns" about GM crops – a concept that explicitly states that there is no risk hypothesis (see Gray, 2004, for the origin and use of this term in GM risk assessment).

The collection of data (EE) adds to confusion when the data test no risk hypothesis; in other words, there is no indication of how the data should be used. This can happen when small parts of a problem are studied, but there is no theory to link the parts together to predict anything of interest. For example, aspects of the pollination biology of oilseed rape (*Brassica napus* L. ssp. *oleifera*) were studied intensively in the 1990s, but no theory was available link the elements to make predictions about rates of gene flow between GM crops and non-GM crops or wild relatives (Raybould, 2004; Raybould and Cooper, 2005). Data are sometimes collected before the risk hypothesis is formulated; for example, field studies of the impact of GM insect resistant crops on non-target organisms are carried out in the absence of any demonstrated hazard of the insecticidal protein, or any other difference from conventional crops that could pose a hazard.

The accumulation of problems at each stage of the risk assessment may mean that knowledge of risk is not increased, even though much new information has been collected. In some cases, the accumulation of data that purport to characterize risk, but do not because they test no risk hypothesis, can increase the perception of significant risk even though risks may be negligible (Johnson et al., 2007). In terms of Popper's scheme, poor problem formulation and consequent inadequate risk characterization can result in $P_2 \leq P_1$.

CONSEQUENCES OF POOR PROBLEM FORMULATION

The immediate result of poor problem formulation is that the data required to show safety are unclear, which leads to a chain of consequences that increases environmental risk. First, companies seeking to register GM crops take a conservative approach to study requirements: studies are done just in case a regulator may require the data; and once a study is done for one product, it will be done for all subsequent products. Conservatism in study requirements leads to bloated regulatory dossiers that raise the cost of complying with regulations and act as a barrier to market for small companies and public sector institutions. Even in large companies, the cost of regulation is a significant disincentive to product development (Miller and Conko, 2004). Such barriers can delay or even prevent the introduction of environmentally beneficial products and the phasing out of harmful ones, so increasing

environmental risks (Cross, 1996). Delay will also result from the increased time needed to review larger dossiers.

The increase in the size of dossiers also gives the impression that studies of currently registered GM crops are revealing new risks that need to be evaluated. This is not the case (*e.g.*, Naranjo et al., 2005; Romeis et al., 2006), and additional benefits of GM crops have been discovered after registration (*e.g.*, James, 2003; Wu et al., 2004). Nevertheless, collection of irrelevant data can increase public unease about GM crops (Johnson et al., 2007), and so lead to inconsistent or highly conservative decisions by regulators. Again, the introduction of beneficial products may be delayed, and environmental risk increased, because poor problem formulation results in production of data in the absence of a risk hypothesis.

IMPROVING PROBLEM FORMULATION IN ENVIRONMENTAL RISK ASSESSMENT OF GM CROPS

To reduce environmental risk, the objective of problem formulation should be to identify the minimum quantity of data needed for risk assessment to demonstrate the safety of a GM crop. Hume's refutation of induction and Popper's arguments about falsification show that risk assessment cannot prove that cultivation of a GM crop is safe (free from risk). However, risk assessment can test risk hypotheses that predict the likelihood and magnitude of harmful events, and so demonstrate safety, provided safety is defined in terms of the probability of an unacceptable outcome.

What constitutes unacceptable harm cannot be derived scientifically (Stern and Fineberg, 1996), and therefore a vital part of the interaction between risk assessment and other components of risk analysis is a definition of harm that serves the needs of policy, while being amenable to scientific analysis. The objectives of policy, the management goals, are often general concepts about protection of health or the environment, and are not amenable to scientific analysis. Therefore the risk assessment must derive measurable phenomena, assessment endpoints, to represent the management goals. For example, in a risk assessment of cotton expressing an insecticidal protein, a management goal of protecting ecological functions could be represented by the population sizes of non-target organisms that are important for biological control (Romeis et al., 2006). The risk assessment would seek to predict the effects of cultivating the GM cotton on the population sizes of these organisms.

The next task is the creation of a conceptual model that links the proposed action, in this case cultivation of a GM crop, to potential harmful effects on the assessment endpoints. The model should be derived from questions of the form, "What is the probability of a specified

change in the assessment endpoints following cultivation of this GM crop?" This type of question leads to specific hypotheses that can demonstrate safety with high confidence. Assessments that begin with the question, "What will happen when this GM crop is cultivated?" provide weak hypotheses or simply lead to observations that test no hypothesis.

The final task for problem formulation is the production of a plan for testing the risk hypotheses. Testing involves collection of data on the properties of the GM crop, its transgenic proteins, or both. These properties are called test or measurement endpoints, and usually constitute estimates of hazard or exposure. Confidence in the risk assessment derives from the certainty of its conclusions. High certainty results from two aspects of good problem formulation: its ability to link particular values of the test endpoints to a high probability of no unacceptable effects to the assessment endpoints; and its power to test whether the values of the test endpoints differ significantly from values that indicate acceptable risk. The implications of these conclusions for the safety testing of GM crops are discussed below.

ARTIFICIALITY *VERSUS* REALISM

The purpose of tests of the risk hypothesis is to contribute to decision-making, not necessarily to broaden scientific knowledge in general (Hill and Sendashonga, 2003). Therefore, problem formulation should concentrate on increasing the possibility of detecting effects that indicate a potential risk. The corollary of this approach is that if such effects are not detected, the confidence that unacceptable effects are improbable is increased. Are there general principles that may guide the problem formulation to maximize the possibility of detecting potential harmful effects?

There has been a long-running argument among ecologists about the relative merits of laboratory and field experiments to detect significant effects. Problems with field studies include incorrect attribution of significance to the manipulated variable because of chance correlation with an uncontrolled factor, and the converse that significance may be missed because of insufficient replication in the face of numerous uncontrolled factors. Laboratory studies offer better control, but can, therefore, "magnify incidental or trivial factors… indeed, laboratory experiments can likely show some effect of any factor by using sufficiently extreme conditions. Laboratory studies are effective in isolating the response to a factor but the response may not be ecologically relevant" (Peters, 1991).

The problem of laboratory studies over-estimating the ecological relevance of factors can be used to advantage in risk assessment. A rule of thumb for effective problem formulation could be to create conceptual models that identify phenomena that are essential for a risk to be realized, and the existence of which can be tested in the laboratory. If the phenomenon cannot be detected in the laboratory, where "incidental or trivial factors" are magnified, then one has high confidence that it will not occur in the field, and hence the likelihood of unacceptable harm is minimal. If the phenomenon is detected in the laboratory, then its relevance in the field can be tested with laboratory experiments that more closely mimic field conditions, or with semi-field or field studies. These are the principles behind tiered testing and risk assessment (Garcia-Alonso et al., 2006; Touart and Maciorowski, 1997).

This approach is particularly effective for the safety testing of GM crops that express insecticidal proteins. A powerful method for risk assessment is the identification and evaluation of differences between a GM plant and a non-GM counterpart that is considered safe (Kuiper et al., 2002). Therefore if plant characterization data show that the only significant difference between a GM crop and its non-GM progenitor is the production of the intended insecticidal protein, the problem formulation for the environmental risk assessment suggests the risk hypothesis that the insecticidal protein will have no harmful effects on non-target organisms at exposures *via* the crop in the field. A rigorous way to test this hypothesis is to evaluate toxicity in the laboratory where confounding variation can be minimized and organisms can be exposed to higher concentrations of the proteins than they will meet under worst-case exposure in the field. If toxicity is not detected under these circumstances in a representative selection of sensitive indicator species, the probability of unacceptable effects to non-target organisms resulting from cultivation of the crop is highly likely to be negligible, and the GM crop can be considered safe.

The method of isolating and testing for essential phenomena under conditions that increase the probability of detecting them is not limited to risks from toxic effects of proteins. For example, Raybould and Cooper (2005) suggested a scheme to assess the risks of gene flow from disease-resistant crops to wild relatives. Exposure, the likelihood of gene flow, could be assessed under worst-case conditions by attempting to hybridize the crop and wild relative in the laboratory using artificial pollination and embryo rescue; if no hybrids are produced under these conditions, it is highly likely that exposure in the field will be negligible. Hazard, the potential for the descendants of hybrids of GM disease-resistant crops and wild relatives to be weedier than existing populations of the wild relative, could also be assessed in the laboratory. The first step is to test whether the wild relatives are sensitive to the pathogens associated with the disease. If the wild plants are immune to infection in the laboratory, then there is high certainty that the probability of increased weediness in the wild relatives, due to protection from a

disease that previously controlled their population size, is minimal. Simple higher tier tests are available if the wild relatives are not immune. Field studies with GM hybrids, or simulation models that attempt to predict the population dynamics of the hybrids, are unnecessary if the lower tier studies indicate minimal risk.

When used with suitable problem formulation, the artificiality of laboratory studies increases confidence that unacceptable harmful effects are unlikely. If no harmful effects are seen in such studies, increasing the realism of the study does not increase the confidence of no harmful effects in the field; in fact if unacceptable uncertainty remains even though no harmful effects have been detected, a more unrealistic test should be considered, as this would increase the probability of detecting an effect. If harmful effects are detected in the laboratory, studies with increased realism may be warranted because they have a different purpose: the highly artificial laboratory studies test for the existence of potential hazards; studies with increased realism aim to quantify risk, given that a hazard has been identified. Increasing artificiality increases the probability of detecting hazards, whereas increasing realism improves evaluation of whether a known hazard constitutes an unacceptable risk.

STANDARD DATA REQUIREMENTS AND CASE-BY-CASE RISK ASSESSMENTS

Each GM crop (combination of crop species and genetic modification) is regulated individually – so-called case-by-case assessment. Case-by-case assessment leads to the concept of data requirements, the idea that a prescribed standard set of data must be acquired for each GM crop to assess risk.

Standardization of data requirements for risk assessments is suitable for products where the problem formulations are similar. Chemical pesticides are an example of a class of product for which standard data requirements for environmental risk assessment make sense. As chemical pesticides are designed to be toxic to a broad spectrum of pests, there is good reason *a priori* to expect that they will have some effect on the various types of wild species that are likely to be exposed. Also, pesticides have broadly similar patterns of use, and consequently similar organisms will be exposed to each pesticide. Therefore, problem formulation can be similar for each pesticide: is the predicted environmental concentration of the pesticide sufficiently lower than that which causes unacceptable toxic effects? The toxicity of each pesticide can be assessed with a standard set of studies on sensitive indicator organisms that represent all potentially exposed species, and the exposure predicted from models based on use rates and the physical chemistry of

the active ingredients (European and Mediterranean Plant Protection Organization, 2003).

The management goals and assessment endpoints for GM crop and chemical pesticide risk assessments are similar, and the clarity of the problem formulation for chemical pesticide risk assessment is useful to illustrate the concepts of environmental risk assessment (*e.g.*, Raybould and Cooper, 2005). Nevertheless, standard data requirements analogous to those for pesticides may unnecessarily restrict problem formulation for GM crops.

Extensive testing has shown that the Cry proteins expressed by GM crops in commercial use and in development have highly specific modes of action and narrow ranges of bioactivity at the concentrations in the GM crops (Schnepf et al., 1998). Hence there is an *a priori* expectation that when expressed in plants, Cry proteins will be toxic to organisms closely related to the target pest and non-toxic to those less closely related to the target. It follows that a standard set of test species, such as is used for safety testing of pesticides, may not be the most effective means of identifying data requirements for safety assessments for GM crops expressing pesticidal proteins.

In the case of a crop expressing a protein that targets a beetle pest but has no activity against pests in other taxa, toxicity tests with carabid beetles, ladybirds and other beetles are more likely to detect effects than are tests with species that are not beetles. Failure to detect toxicity in ladybirds and carabids may indicate minimal risk of this protein to all potentially exposed species that are not beetles with far higher certainty than, say, tests on honeybees and lacewings, even though the latter studies are more "realistic" in that they involve species that are not beetles. In other words, beetles are better indicator species for this protein than are lacewings or honeybees. If the protein targets a specific lepidopteran pest, non-target Lepidoptera would be better indicators. Of course, if the single objective of the risk assessment were protection of honeybees, rather than all potentially exposed non-target organisms, then a honeybee study may indicate the degree of risk with higher certainty than studies of beetles or Lepidoptera.

It follows that for Cry proteins and other narrow-spectrum pesticides, such as vegetative insecticidal proteins (Vips) (Estruch et al., 1996), it is possible to derive highly specific and powerful risk hypotheses. For a protein that is active against beetle pests in a single Family, a risk hypothesis of the form, "At concentrations in the field, this protein is not harmful to non-target Coleoptera", may be the most effective risk hypothesis for protecting all potentially exposed non-target organisms. Therefore requirements for non-target organism toxicity data would be limited to beetles. Analogous

risk hypotheses, and similarly specific data requirements, could be formulated for proteins with different target pest spectra. At present, such strategies are unlikely to be acceptable because of insufficient confidence that screens of activity against pests indicate the range of toxicity of Cry proteins and Vips. Nevertheless, knowledge of the mode of action and spectrum of activity against pests gained during the development of insecticidal proteins offers the opportunity to tailor data requirements for specific products: if the protein has only shown activity to a specific beetle pest, then the testing of non-target beetles should be emphasized; if the proteins has only shown activity against certain Lepidoptera, testing fewer beetles and more Lepidoptera may be more effective.

For some pesticidal proteins, no non-target organism toxicity data may be required to demonstrate safety with high certainty. The fungi *Fusarium graminearum* and *F. culmorum* cause a disease of cereals called Fusarium head blight (FHB). The fungi produce deoxynivalenol (DON), a trichothecene mycotoxin that enhances FHB severity and is a heath hazard to humans and livestock. The protein FsTRI101 transfers acetyl groups to the C3 hydroxyl group of trichothecenes, which includes DON. A gene from *F. sporotrichoides* (*FsTRI101*) that codes for FsTRI101 was transformed into wheat and shown to provide protection against FHB (Okubara et al., 2002).

Under the Federal Insecticide, Rodenticide and Fungicide Act, any substance that plants produce for protection against pests are pesticides "if humans intend to use these substances for preventing, repelling or mitigating any pest" (US Environmental Protection Agency, 2001). FsTRI101 could be regarded as a pesticide under this definition, and hence, in the USA at least, be regulated in the same manner as insecticidal proteins. As for insecticidal proteins, the assessment endpoints for a risk assessment of plants producing FsTRI101 would be populations of non-target organisms, but it does not follow that ecotoxicology studies would be required to demonstrate safety. A more effective approach may be to test the hypotheses that the activity of FsTRI101 is limited to the C3-OH trichothecenes, and that the C3-OH trichothecenes are found only in *Fusarium* and closely related genera. These hypotheses would be tested by biochemical studies, not non-target organism toxicity tests. If these hypotheses were falsified, then standard non-target organism testing may be triggered.

These examples show that requirements for ecotoxicological studies on a standard set of non-target species for all pesticidal proteins could unnecessarily restrict the problem formulation and decrease the certainty of the risk assessment. Just as observations cannot be prior to theory, data requirements should not be prior to problem formulation.

CONCLUSIONS

Poor problem formulation in GM risk assessments can lead to the collection of unnecessary data, which may prevent or delay the introduction of environmentally beneficial products, so increasing environmental risk. By creating risk hypotheses that can be tested rigorously, good problem formulations seek to minimize data requirements, while enabling risk assessments capable of indicating with high confidence that GM crops pose minimal environmental risk. Confidence is increased by identification of phenomena necessary for harm to occur, and rigorous testing for their existence. Failure to observe such phenomena under conditions designed to maximize their detection can be regarded as an indication of safety. Two important implications are that artificial test conditions can increase confidence, whereas prescribed data requirements can reduce confidence (increase uncertainty) if they constrain problem formulation.

REFERENCES

Cross FB (1996) Paradoxical perils of the precautionary principle. *Wash. L. Law Rev.* **53**: 851–925

Estruch JJ, Warren GW, Mullins MA, Nye GJ, Craig JA, Koziel MG (1996) Vip3A, a novel *Bacillus thuringiensis* vegetative insecticidal protein with a wide spectrum of activities against lepidopteran insects. *Proc. Natl. Acad. Sci.* **93**: 5389–5394

European and Mediterranean Plant Protection Organization (2003) Environmental risk assessment scheme for plant protection products. *EPPO Bull.* **33**: 133–114

Garcia-Alonso M, Jacobs E, Raybould A, Nickson TE, Sowig P, Willekens H, van der Kouwe P, Layton R, Amijee F, Fuentes AM, Tencalla F (2006) A tiered system for assessing the risk of genetically modified plants to non-target organisms. *Environ. Biosafety Res.* **5**: 57–65

Giles J (2003) Biosafety trials darken outlook for transgenic crops in Europe. *Nature* **425**: 751

Gray AJ (2004) Ecology and government policies: the GM crop debate. *J. Appl. Ecol.* **41**: 1–10

Green RE, Cornell SJ, Scharlemann JPW, Balmford A (2005) Farming and the fate of wild nature. *Science* **307**: 550–555

Groot AT, Dicke M (2001) Insect-resistant transgenic plants in a multi-trophic context. *Plant J.* **31**: 387–406

Hill RA, Sendashonga C (2003) General principles for risk assessment of living modified organisms: Lessons from chemical risk assessment. *Environ. Biosafety Res.* **2**: 81–88

James C (2003) Global review of commercialized transgenic crops. *Curr. Sci.* **84**: 303–309

Johnson KL, Raybould AF, Hudson MD, Poppy GM (2007) How does scientific risk assessment of GM crops fit within the wider risk analysis? *Trends Plant Sci.* **12**: 1–5

Kuiper HA, Kleter GA, Noteborn HP, Kok EJ (2002) Substantial equivalence – an appropriate paradigm for the safety assessment of genetically modified foods? *Toxicology* **181-182**: 427–431

Magee B (1985) Popper. Fontana Press, London

Mendelsohn M, Kough J, Vaituzis Z, Matthews K (2003) Are *Bt* crops safe? *Nature Biotech.* **21**: 1003–1009

Miller HI, Conko G (2004) The Frankenfood Myth. Praeger, Westport, CT

Naranjo SE, Head G, Dively GP (2005) Field studies assessing arthropod nontarget effects in *Bt* transgenic crops: introduction. *Environ. Entomol.* **34**: 1178-1180

Okubara PA, Blechl AE, McCormick SP, Alexander NJ, Dill-Macky R, Hohn TM (2002) Engineering deoxynivalenol metabolism in wheat through the expression of a fungal trichothecene acetyltransferase gene. *Theor. Appl. Genet.* **106**: 74–83

Patton DE (1998) Environmental risk assessment: tasks and obligations. *Hum. Ecol. Risk Assess.* **4**: 657–670

Peters RH (1991) A Critique for Ecology. Cambridge University Press, Cambridge

Popper KR (1959) The Logic of Scientific Discovery. Hutchinson, London

Power M, Adams SM (1997) Perspectives of the scientific community on the status of ecological risk assessment. *Environ. Manag.* **21**: 803–830

Raybould AF (2004) A decade of gene flow research: improved risk assessments or missed opportunities? *Aspect Appl. Biol.* **74**: 27–33

Raybould AF, Cooper JI (2005) Tiered tests to assess the environmental risk of fitness changes in hybrids between transgenic crops and wild relatives: the example of virus resistant *Brassica napus*. *Environ. Biosafety Res.* **4**: 125–140

Raybould AF, Wilkinson MJ (2005) Assessing the environmental risks of gene flow from genetically modified crops to wild relatives. In Poppy GM, Wilkinson MJ, eds, Gene Flow from GM Plants. Blackwell Publishing, Oxford, pp 169–185

Romeis J, Meissle M, Bigler F (2006) Transgenic crops expressing *Bacillus thuringiensis* toxins and biological control. *Nature Biotech.* **24**: 63–71

Schnepf E, Crickmore N, Van Rie J, Lereclus D, Baum J, Feitelson J, Ziegler DR, Dean DH (1998) *Bacillus thuringiensis* and its pesticidal crystal proteins. *Microbiol. Mol. Biol. Rev.* **62**: 775–806

Sears MK, Hellmich RL, Stanley-Horn DE, Oberhauser, KS, Pleasants JM, Mattila HR, Siegfried BD, Dively GP (2001) Impact of *Bt* corn pollen on monarch butterfly populations: a risk assessment. *Proc. Natl. Acad. Sci. USA* **98**: 11937–11942

Squire GR, Brooks DR, Bohan DA, Champion GT, Daniels RE, Haughton AJ, Hawes C, Heard MS, Hill MO, May MJ, Osborne JL, Perry JN, Roy DB, Woiwood IP, Firbank LG (2003) On the rationale and interpretation of the Farm Scale Evaluations of genetically modified herbicide-tolerant crops. *Phil. Trans. R. Soc. Lond* B **358**: 1779–1799

Stern PC, Fineberg HV, eds (1996) Understanding Risk: Informing Decisions in a Democratic Society. National Academy Press, Washington, DC

Suter GW (1990) Endpoints for regional ecological risk assessments. *Environ. Manage.* **14**: 9–23

Touart LW, Maciorowski AF (1997) Information needs for pesticide registration in the United States. *Ecol. Appl.* **7**: 1086–1093

US Environmental Protection Agency (1998) Guidelines for ecological risk assessment. *Fed. Regist.* **63**: 26846–26924

US Environmental Protection Agency (2001) Regulations under the Federal Insecticide, Fungicide and Rodenticide Act for Plant-Incorporated Protectants (formerly Plant-Pesticides) *Fed. Regist.* **66**: 37772–37817

Wolt JD, Peterson RKD (2000) Agricultural biotechnology and societal decision-making: the role of risk analysis. *AbBioForum* **3**: 39–46

Wu F, Miller DJ, Casman EA (2004) The economic impact of *Bt* corn resulting from mycotoxins reduction. *J. Toxicol. Toxin Rev.* **23**: 397–424

To access this journal online:
www.edpsciences.org/ebr