Ecological and experimental constraints for field trials to study potential effects of transgenic *Bt*-crops on non-target insects and spiders

CGM 2014-06

ONDERZOEKSRAPPORT

PLANT RESEARCH INTERNATIONAL

WAGENINGEN UR

Ecological and experimental constraints for field trials to study potential effects of transgenic Bt-crops on non-target insects and spiders

Kees Booij

Plant Research International, part of Wageningen UR
Business Unit Biointeractions and Plant Health
October 2014

Report 592

Plant Research International, part of Wageningen UR
Business Unit Biointeractions and Plant Health

Address	:	P.O. Box 69, 6700 AB Wageningen, The Netherlands
	:	Wageningen Campus, Droevendaalsesteeg 1, Wageningen, The Netherlands
Tel.	:	+31 318 48 06 70
Fax	:	+31 317 41 80 94
E-mail	:	info.pri@wur.nl
Internet	:	www.wageningenUR.nl/en/pri

Table of contents

Preface

When an application for cultivation of an insect resistant GM crop is filed, potential risks of the GM crop to non-target organisms (NTOs) are assessed. Applications usually contain data of laboratory experiments as well as field trials. Both types of studies are performed to investigate whether NTOs might be affected by the GM crop.
Current standards require field trials to be carried out at a number of locations and in several years. In a typical field trial the GM crop and the conventional control are planted in several plots (replicates) and the number of NTOs is determined at several time points during the growth period of the crop. Finally, a comparison is made between the number of NTOs in the GM crop and the conventional control.
As most NTOs are mobile most of their life, they can easily switch between different plots when the plot size is smaller than the distance normally covered by them. The mobility of NTOs reduces the confidence with which conclusions can be drawn from field trials. Another factor limiting the confidence of conclusions from field trials is the natural variation in the size of NTO populations during the year and between years.
To investigate to what extent the current field trials are able to detect potential effects on NTOs, COGEM commissioned a research project. This research project was carried out by dr. ir. C.J. H. Booij (Plant Research International). The resulting report provides an overview of the factors complicating the value of field trials, and draws conclusions on the ability to use field trials to investigate whether a GM crop adversely affects NTOs.

Dr. P.M. Bruinenberg
Chair of the advisory committee

Advisory committee:
Dr M. Bovers, COGEM staff
Dr. P. M. Bruinenberg, AVEBE U.A.
Dr. D.C.M. Glandorf, GMO office, National Institute of Public Health and the Environment
Dr. T.J. de Jong, Leiden University
Prof. dr. ir. J.J.A. van Loon, Laboratory of Entomology, Wageningen University
Dr. R.Y. van der Weide, ACRRES Wageningen University

Summary

Before insect resistant transgenic crops are approved for commercial growing, potential effects on non-target organisms (NTOs) are studied in laboratory studies and field trials. A certain consensus has been reached about how such experiments should be performed and about the selection of non-target organisms to be studied.
Recent meta-analyses have shown that up till now in field trials with transgenic Bt-crops in the last 15 year virtually no significant effects on non-target insects could be shown when compared with non-transgenic isogenic crops. Based on these evaluations the cultivation of current commercial transgenic Bt-maize varieties is considered to be safe for NTO's.

The question remains whether effects of transgenic insect-resistant crops are really absent or whether it is very difficult or even impossible to detect such effects by field trials as they have been and currently are performed. Arguments that challenge the non-effect conclusions include the scale of experiments, the selection of the right organisms, the way potential effects are defined and measured and the way data are processed and interpreted. In particular the spatial and temporal variability in abundance of non-target organisms is of concern and hinders the analysis and clear cut interpretation of experimental results. A further concern is that many NTO's are very mobile and movement between plots and between experimental fields and the surrounding area's masks small effects, in particular in small scale experiments. This lack of experimental robustness is leaving risk managers with uncertainty about the judgement of experimental results. Meanwhile it feeds the scientific community with more questions and incentives to improve experimental set-ups.

This report aims to explore the experimental and ecological methods employed in the field trials done last years with an emphasis on the temporal and spatial dynamics of NTO's that form an obstacle for detecting effects of transgenic crops on those NTO's. The exploration on experimental methodology in non-target field trials is done by critically analysing a number of recent case studies in insect resistant Bt-maize with respect to experimental design, scale, sampling procedures for abundance assessment and interpretation of results in relation to the ecological properties and function of the species / taxa studied. Also the implication of combining species in functional groups is treated.

From this analysis it appeared that in most case studies no a priori knowledge was available about expected density, variance in number or mobility of the NTO's groups studied. As a consequence experimental set-up or sampling plans are based on experience from earlier similar studies e.g. from pesticide effect studies. Most studies were not focused on particular species but targeted a broad spectrum of insects or spiders using general sampling devices such as pitfall traps, sticky traps or visual observations. The scale of experiments with regard to plot size, number of replicates and number of years was mostly based on thumb rules and costs involved within the guidelines provided in literature.

In many cases different sampling methods are used to get some measure reflecting the local abundance of non-target insects and spiders. Classical methods such as pitfall traps, water traps, sticky traps and visual observation of plants are commonly used and are efficient ways to get an impression of abundance for a wide variety of organisms, but all these methods have major drawbacks. Typically such assessments are made at intervals during the cropping season. It is a key issue to realize that all the methods do not necessarily reflect absolute density (numbers present/m^2) but are the result of a combination of density and activity. The latter varies not only by species but also by weather, food conditions, crop structure and even by trap position. Even though care is taken in some experiments that a reasonable number of traps is used or observations are made to cope with heterogeneity within plots, the methodology often gives only a rough and imprecise picture of the momentary abundance of organisms under field conditions. One of the reasons that numbers for many species are too low or too variable to reach statistical power may be that sampling is not intensive enough or inappropriate for those species.

To cope with low numbers and numerical variability in many species two major strategies are used. The first is to pool species in a small number of functional (often taxonomic) groups. The second is to sum-up or average the

numbers over the full sequence of sampling dates within a season. The advantage of this is that the data variance tends to decrease and a better estimate of an overall season abundance is reached. The drawbacks, however, are that any temporal effects (e.g. just after pollen shed) are obscured, that no conclusions can be drawn about effects on particular species, and that organisms are pooled while having different ecological functions, value or impact. The identification of all species captured throughout the season is to be considered waste of time when afterwards they are pooled for statistical reasons. In practice functional groups sampled in most field trials are dominated by only a few species and about 10 species might be sufficient to cover the major beneficial organisms.
It is concluded that the current weakness in data analysis is not in the statistical procedures but in the data itself. Instead of using multiple methods and getting data from a wide variety of organisms, more focus on some carefully chosen species with the best methodology could provide more valuable information with less effort.

The survey of studies also showed that many maize/corn systems are a rather simple and poor habitat compared to other agricultural crops, which is reflected in the low numbers of most species caught with different sampling techniques (hindering the statistical analysis and effect detection). Higher numbers of beneficial insects are only observed for some Carabidae (notably *Pterostichus* and *Harpalus* species), 2 or 3 orb-web spiders (*Erigone* and *Meioneta*), 1 or 2 ladybeetles (*Coccinella* and *Hippodamia*), one predatory bug (*Orius*) and sometimes some other such as the lacewing *Chrysopa carnea* and hover flies such as *Episyrphus balteatus*. Also some phytophagous species such as aphids, thrips and leafhoppers can be more abundant but their numbers are highly variable.
All these species that are frequently analysed because they are numerous tend to be mobile to very mobile and disperse and redistribute quickly to avoid unfavourable conditions and find favourable conditions. They are common and can thrive in multiple habitats and some of them are fast colonizers of fields after temporarily unfavourable conditions.

Their fast redistribution tends to level out differences at a smaller scale and results in a more even distribution (hence less variability in samples). As dispersal rates in these organisms can be in the order of tens to hundreds of meters a day or more, common plot sizes of 20x20m to 50x50m that are often used should be considered to be too small. Frequent movement of insects and spiders between, from and into experimental plots under such conditions is likely to be a common phenomenon. To avoid exchange between treatment in field trials plot size should be at least 1 ha for many NTOs.

Knowing all the experimental drawbacks connected with the current experimental approaches it seems unlikely that slight effects of Bt-crops on common and highly mobile species can be detected, even though the statistical power suggests that such conclusions are warranted. On the other hand, adverse effects of insecticide treatments in the same or similar experiments are more obvious. This suggests that if any effects of current transgenic Bt-maize exist they seem to absent or less clear than at least some frequently applied insecticides. This might be due to the fact that most of the species currently studied are not sensitive, or not (or only indirectly) exposed to the Cry toxins. It might also indicate that pesticide effects (in particular when broad spectrum) are much stronger indeed or that they are due to indirect effects such as prey depletion.

From the report it is concluded that current field trial methodology often is inappropriate to study potentially small effects of current insect resistant crops on NTO's because of the dynamic nature of most insect and spider populations in particular due to mobility aspects and sampling obstacles.
It is suggested that effects should be more focused on resident species or life stages (such as larvae) or carefully chosen species. In this context also less beneficial but more sensitive and presumably exposed organisms such as field margin bound Lepidoptera (in case of Cry 1Ab) or leaf beetles (Chysomelidae) in case of Cry3Bb) could be interesting study objects. But also in those cases it might appear difficult to exclude dispersal mediated interference between treatments or between treatments and the environment. Only larger plot sizes and intensive sampling procedures may help to create more powerful experiments. Knowing this, it might be better to look for alternative approaches including modelling to extrapolate laboratory results to impacts under natural conditions.

Samenvatting

Voordat transgene insecten-resistente gewassen worden toegelaten voor commerciële teelt worden laboratorium proeven en veldstudies gedaan om potentiële effecten op niet-doelorganismen te onderzoeken. Er bestaat een zekere consensus over hoe zulke veldproeven moeten worden opgezet en er zijn selectiecriteria voor de te bestuderen niet-doelorganismen (NTOs).

Recente meta-analyses laten zien dat tot nog toe in dergelijke veldproeven vrijwel nooit significante effecten van transgene Bt-gewassen op aantallen van niet-doelwit organismen gevonden worden ten opzichte van hun niet-transgene counterparts. Op basis van deze evaluaties wordt vooralsnog geconcludeerd dat de huidige transgene Bt-maize veilig is voor niet-doelwit organismen.

Wordt hieruit terecht geconcludeerde dat de huidige insectenresistente Bt gewassen geen effecten hebben op niet-doelwit organismen of moeten we concluderen dat deze effecten er misschien wel zijn maar met de huidige experimenten niet of moeilijk aangetoond kunnen worden?

Argumenten die aangevoerd worden tegen de non-effect conclusies zijn de schaal waarop de proeven worden gedaan, de keuze van de te toetsen soorten en de manier waarop potentiële effecten worden gedefinieerd en gemeten. Ook over de interpretatie van resultaten is nogal eens discussie. Vooral de grote variatie onder veldomstandigheden tussen plaatsen en jaren maar ook binnen proeven maakt dat verschillen niet altijd goed te toetsen zijn of moeilijk te generaliseren naar andere omstandigheden. Een ander probleem is dat veel NTOs sterk mobiel zijn en verschillen mogelijk gemaskeerd worden door een sterke uitwisseling tussen de behandelingen in proeven en tussen de experimenten en het omringende gebied. Dit zou de zeggingskracht van proeven kunnen ondermijnen, vooral wanneer de experimentele eenheden kleinschalig zijn en potentiële effecten klein.
Dit rapport heeft tot doel de experimentele opzet en ecologische methoden van recente NTO effectproeven te evalueren met de dynamiek en beweeglijkheid van de frequent onderzochten soorten. Op deze wijze wordt een inschatting gemaakt omtrent de vraag of de huidige veel gebruikte proefopzetten en methodieken in staat zijn potentiële effecten betrouwbaar aan te tonen. Daarbij werd vooral gekeken naar de experimentele opzet, de ruimtelijke schaal van proeven, de bemonsteringsmethoden om populatiedichtheden van NTOs te schatten en de manier waarop proeven werden geïnterpreteerd in relatie tot de eigenschappen van de onderzochte soorten.
Uit de evaluatie van de recente studies bleek dat er in de meeste gevallen geen voorkennis was over de verwachte variatie in dichtheden en de mobiliteit van de onderzochte NTOs. Mede daardoor zijn de proefopzet en de bemonstering niet afgestemd op de gemeten soorten, maar veelal gebaseerd op eerdere vergelijkbare onderzoeken zoals het toetsen van effecten van bestrijdingsmiddelen. De meeste studies concentreren zich niet op specifieke van tevoren gekozen soorten maar eerder op een breed spectrum aan insecten en spinnen met gebruikmaking van algemene bemonsteringstechnieken om een zo ‘breed mogelijk’ beeld te krijgen. De plot-grootte, het aantal herhalingen en het aantal proefjaren is gebaseerd op vuistregels en vaak conform de algemene richtlijnen die in de literatuur te vinden zijn.

In veel gevallen worden bodemvallen, lijmvallen en visuele waarnemingen in het gewas gebruikt om een schatting te krijgen van de populatiedichtheid van verschillende NTOs in de experimenten. Dergelijke metingen worden vaak een aantal keren per seizoen gedaan. Deze algemene vangstmethoden zijn een efficiënte manier om een beeld te krijgen van veel verschillende soorten, maar helaas kleven er nogal wat nadelen aan. Een belangrijk probleem is dat het aantal gevangen dieren niet alleen afhangt van het aantal aanwezige dieren op een bepaalde plek maar vooral ook afhangt van hun activiteit. Deze activiteit per soort verschilt en hangt af van factoren zoals het weer, de voedingstoestand van de NTOs, de gewasstructuur en zelfs de positie van de vallen. Zelfs als genoeg vallen worden gebruikt om de heterogeniteit binnen experimentele eenheden te compenseren geven dergelijke methoden over het algemeen een ruw en onnauwkeurig beeld van de werkelijke populatiedichtheid. Bij veel soorten blijken de gevangen aantallen vaak te laag om statistisch te kunnen analyseren omdat de bemonsteringsmethode te weinig intensief of niet echt geschikt is.

Om met de lage aantallen en met de variatie in aantallen te kunnen omgaan worden twee verschillende strategieën gebruikt. De eerste is om soorten met lage aantallen samen te voegen tot functionele (in praktijk vaak taxonomische) groepen. De tweede is om de aantallen van verschillende tijdstippen gedurende het hele seizoen op te tellen of te middelen om de variatie te beteugelen. De nadelen daarvan zijn echter dat tijdelijke kortdurende effecten (bijvoorbeeld vlak na de aanwezigheid van stuifmeel) verloren gaan in dynamiek van het seizoen en dat soorten worden samengevoegd die ecologisch verschillende functies kunnen hebben, fenologisch of qua grootte sterk verschillen of zelfs een tegengestelde response kunnen vertonen ten opzichte van de behandeling. Bovendien is de tijdrovende identificatie van alle soorten die vaak gedaan wordt vrij zinloos als ze daarna om statistische redenen worden samengevoegd bij de effect analyses. In de praktijk wordt binnen de diverse functionele of taxonomische groepen de hoofdmoot gevormd door slechts enkele soorten. Het zou beter zijn deze dominante soorten apart te analyseren. Uit de analyse wordt geconcludeerd dat de huidige zwakke punten in veel veldproeven niet zozeer de statistische onderbouwing of procedures zijn, maar eerder de beperkte zeggingskracht van de data zelf waaronder de lage aantallen en het samenvoegen. In plaats van het gebruik van generalistische bemonsteringsmethodes en daarmee onnauwkeurige gegevens van veel verschillende soorten, zou het verstandiger zijn zich te richten op enkele goed gekozen soorten met de beste methodologie die aangepast is aan de ecologie van die soorten.

De evaluatie van veldproeven laat ook zien dat maïs ecosystemen een relatief arm habitat vormen voor arthropoden vergeleken met andere gewassen, hetgeen gereflecteerd wordt in de lage aantallen bij de meeste soorten waardoor ook het detecteren van verschillen bemoeilijkt wordt. Onder de meest onderzochte NTOs worden hogere aantallen gevonden bij enkele loopkevers (zoals *Pterostichus* en *Harpalus* soorten), 2 of 3 hangmatspinnen (*Erigone* en *Meioneta*), 1 of 2 lieveheersbeestjes (*Coccinella* en *Hippodamia*), een roofwants (*Orius*), gaasvliegen en (*Chrysopa carnea)* en zweefvliegen (zoals *Episyrphus balteatus).* Al deze soorten leven van bladluizen, cicaden en trips die veelvuldig op maïs voorkomen en talrijk kunnen zijn maar ook sterk in aantal fluctueren.
Al deze veel onderzochte soorten kunnen weliswaar talrijk zijn maar zijn zonder uitzondering ook sterk mobiel en aangepast aan de dynamiek van agrarische teeltsystemen. Ze hebben het vermogen zich snel te herverdelen in het landschap zodra de lokale omstandigheden veranderen. Ze kunnen het goed doen in verschillende habitats en herkoloniseren agrarische percelen wanneer de omstandigheden gunstig zijn.
Hun vermogen tot herverdelen leidt ertoe dat verschillen op kleine schaal snel worden uitgevlakt of versterkt terwijl op hoger schaal niveau de populatie stabiel blijft. Voor proeven met plots ter grootte van 20x20m tot 50x50m is de uitwisseling al snel te groot om ontstane verschillen te kunnen toeschrijven aan sterfte of reproductie. Voor NTOs zoals insecten en spinnen lijkt de plotgrootte van tenminste 1 ha noodzakelijk om dit soort interferentie tussen plots te verminderen.

Gezien de hierboven geconstateerde zwakke punten in de huidige veldstudies naar effecten van GM gewassen op NTOs lijkt het onwaarschijnlijk dat in dergelijke proeven potentiële maar minder sterke effecten betrouwbaar kunnen worden aangetoond ondanks de statistisch goed onderbouwde conclusies die in sommige gevallen worden getrokken.

Aan de andere kant lijken binnen dezelfde proeven effecten van bestrijdingsmiddelen wel op te treden. Maar het betreft dan vaak effecten van breed werkende middelen of van predatoren die direct maar kortdurend worden getroffen door verdwijnen van hun prooi. Toch lijkt dit verschijnsel er in ieder geval op te duiden dat eventuele GM effecten van NTOs –mochten die reëel zijn- in ieder geval geringer zijn dan van een aantal minder selectieve bestrijdingsmiddelen. Bovendien moet bedacht worden dat veel van de in veldstudies onderzochte soorten weinig gevoelig zijn of weinig blootgesteld zijn aan de Cry toxines.
In het onderhavige rapport wordt geconcludeerd dat de huidige aanpak en gebruikte methodiek in veldstudies onvoldoende geschikt is om potentiële kleine effecten van de huidige insecten-resistente Bt gewassen op NTOs nauwkeurig te bepalen vanwege de dynamische eigenschappen van de meeste insecten en spinnen soorten en met name door de grote beweeglijkheid van veel NTOs en het ontbreken van efficiënte maar goede bemonsteringsmethoden om de populatiedynamiek goed te kunnen volgen via absolute dichtheidsmetingen.
Het wordt aanbevolen voor eventuele veldstudies het onderzoek te richten op een enkele goed gekozen potentieel gevoelige maar weinig mobiele soort of op minder mobiele levensstadia (larven). In deze context zou de aandacht kunnen verschuiven naar minder functionele maar potentieel wel gevoelige soorten die aan toxines worden blootgesteld zoals rupsen van vlinders en motten (in het geval van Cry 1 toxines) of larven van bladhaantjes (in het

geval van Cry3 toxines) die op onkruiden of in perceelranden voorkomen. Maar ook hierbij kan de response van mobiele adulte stadia interferentie tussen plots veroorzaken. Te allen tijde is het belangrijk de plotgrootte aan te passen aan de kennis die er over mobiliteit bestaat om de zeggingskracht van proeven te vergroten. Wanneer de ruimtelijke of temporele dynamiek van soorten zinvolle proeven verhindert, zijn modelmatige benaderingen vaak zinvoller of aanvullend om laboratorium resultaten naar de mogelijke gevolgen voor de natuurlijke omgeving te vertalen.

1. Introduction

To protect the natural environment and the functioning of ecosystems, the introduction of transgenic crops on the European market is regulated. Guidelines have been formulated for laboratory and field tests to warrant safety before crops are approved to be cultivated (EFSA panel 2010). Depending on technological developments and increased scientific knowledge it seems likely that these guidelines will be frequently updated and evaluated when more transgenic crops will be developed.

Currently herbicide and insect-resistant maize are dominating the transgenic crop market. For insect resistance the insertion of *Bacillus thuringiensis* (Bt) genes that produce various Cry toxins is the major technology. In the case of insect-resistant crops, tests are required by the European Food Safety Authority (EFSA) to show that there are no harmful effects on non-target organisms (NTOs) such as beneficial insects and protected species. Cry toxins are assumed to be quite specific for the target organisms at level of insect orders and non-target effects are most likely to be found at the same taxonomic level, but there are indications that direct or indirect effects on other species cannot be excluded. More and more laboratory studies and field experiments are performed to detect any unexpected effects.

Field tests are required to investigate whether effects found in the laboratory also occur under natural conditions. Because under field conditions the response of non-target organisms is extremely variable, carefully designed field experiments have to be repeated and performed at different sites and at a proper scale. Procedures for such field trials have been described that are considered to be sound and cost-effective (Perry et.al. 2009, EFSA 2009a, 2009b, 2011)

Based on these procedures many field trials have been performed in the last decade for transgenic insect-resistant maize in both the USA and in Europe (see appendix for an overview). Most studies so far have not shown consistent, statistically significant effects on non-target organisms (Sweet and Bartsch 2012, Devos and de Schrijver 2012). But these conclusions are challenged by several scientists and stakeholders, based on methodological considerations. Whether effects on non-target organisms are minor or really do not exist or whether such field trials are unable to show such effects can be debated. Of major concern are potential adverse effects that are not detected due to possible weaknesses in the methodology. Before refining the methodology, studying more and more potential effects, or extending field trials to a larger scale, one may wonder to what extent the current field trial approach can provide unequivocal answers to questions about effects on non-target organisms.

In the case of organisms or ecosystem functions to protect, it is usually assumed our understanding of population dynamics and ecosystem functioning is sufficient and that current methodology for measuring effects of transgenic crops on non-target organisms is quite robust. However, the uncertainty about our conclusions from experiments may not arise from deficiencies in our experimental procedures and statistical analysis but rather in the lack of understanding of the temporal and dynamics of many species studied and the way available knowledge is neglected to guarantee a proper set up of experiments. This makes a sound interpretation of many field studies difficult and gives rise to much uncertainty in the conclusions.

One of the basic concerns arising from current ecological insights is that processes and dynamics for ecological risk assessments can hardly be understood without considering the spatial context in which organisms live (Tscharntke *et al* 2012, Galic *et al* 2010). Experimental set-ups should take movements and dispersal capacity into account in order to do trials at the right scale. The dynamic response of species in agro-ecosystems to changing conditions forms a major constraint when studying abundance of non-target organisms in field trials. An inappropriate experimental spatial or temporal scale is likely to generate false conclusions or do not the justify extrapolation to another scale (Englund and Cooper 2003).

In this report a critical analysis is made of the results from field trials that have been published for non-target effects of Bt-maize with emphasis on the experimental set-up, plot-size, sampling methods, data analysis and interpretation of the results in the context of the population dynamic characteristics of the species studied.

Weaknesses in the methodology are identified and consequences for the interpretation if results are analysed. Based on this evaluation suggestions are given to find improved or alternative research methods to study effects of transgenic crops on non-target insects and spiders.

2. Aims and procedures of the project

The aim of the project was to evaluate the current field trials for their potential to detect effects of transgenic insect-resistant crops (*in casu* maize) on non-target insects and to suggest recommendations for more appropriate or new methodologies to improve the power of field trials to detect potential effects of GM crops on NTOs.
In order to fulfil the aims formulated for the project the following tasks were carried out:

- Generating an overview, based on literature of recent field trials including information on commonly studied non-target species and the parameters and statistical analysis used.
- Critically formulating a set of criteria that was used to test the usefulness of field trials for detecting effects, taking spatial-temporal dynamics of species into account
- Linking these criteria with methodological/statistical issues such as plot size, plot interference and sampling in field trials for non-target arthropods in order to identify limiting conditions and bottlenecks
- Using the information collected to evaluate the field trials that have recently published in scientific literature for their power and spatial bottlenecks to detect effects on NTOs.
- Formulating minimal requirements, potential improvements, limits and alternatives to study effects of insect-resistant transgenic crops on non-target arthropods.

In this context recent field trial methodology and the behavioural and population dynamic properties of studied organisms was reviewed to judge the statistical and ecological significance or obtained results. Based on the background knowledge about statistical aspects of field trials (Perry *et al* 2009) and spatial ecology of farmland organisms (Toft and Riedel 1992), a set of criteria was used to evaluate field trials including spatial (movement), temporal (dynamics) and statistical (sampling, variance) aspects. Bottlenecks such as plot-interference, low-densities, spatial scale and pseudo-replicates would be included in the analysis as well. This approach was applied to a number of recent and representative case studies.

Based on the analysis it was aimed to formulate minimal requirements to study effects on NTOs in field trials or proposing alternative methodologies.

The study was limited to transgenic insect-resistant maize and to non-target insects and spiders.

Information sources used

There is a wealth of information to be found in reports and scientific publications that are related to the potential risks and benefits of transgenic insect-resistant crops. For a proper analysis the trial methodology issue has to be placed in the broader context of risk assessment and effect studies of GM crops in general. The question of non-target effects is embedded in the policy requirement to avoid potential adverse effects of transgenic crops on human health and the environment. In this framework the definition of 'effects', 'beneficial organisms', 'populations' and 'ecosystem functions' is relevant for the set-up of meaningful field trials (for example to select non-target species and relevant effect levels). The implication of the undefined use of these terms is a matter of concern for formulating hypothesis in field trials but as such not the focus of this report. This background information, however, is implicitly taken into account when critically analysing the field trial literature.

It was also realized during the project that scientists focus on either ecosystem services (beneficial NTOs), species of conservation value (biodiversity) or just looked at 'any unexpected effects' that pop up in field trials. Depending on the focus researchers study effects on single species, specific groups or effects on whole communities. These different focuses/study subjects resulted in a variety of set-ups and a multitude of parameters. For budget and time constraints this study is not exhaustive in reviewing back ground information and all relevant information covering all the NTO studies. Therefore emphasis was placed on recent reviews, a number of much cited and recent case studies for Bt-maize and some selected alternative experimental approaches. The case studies were thoroughly screened for experimental, sampling and statistical methodology, and on how the researchers handled and discussed variability in measurements and discussed population dynamic aspects.

Because mobility aspects of NTOs are often not considered in most of the papers on field trial despite that they are very relevant, a special search was made for mobility studies for the most commonly studied NTOs in the maize ecosystem.

In this project a range of key-words has been used to find relevant information in the literature using the databases Scopus and Web of Science as well as occasional internet searches with Google. Recent reviews (Sweet and Bartsch 2012, Devos en de Schrijver 2012, Wolfenbarger *et al* 2012), and EFSA (2009, 2010), EFSA &COGEM (2013) and AMIGA publications have been used as a starting point. Cross-references and case studies were an additional entrance to find relevant information. Finally ecological mobility studies not related to GM crops were searched to find mobility / dispersal information for commonly used NTOs.

The study focused on non-target organisms associated with maize ecosystems (summarize for example by Scholte and Dicke 2005, but see also Meissle *et al* 2012) , but the same key words and insect group names could be used for other GM crops. Of course each crop can has its own set of specialized associated crop feeding species but those are often plant feeding only and mostly considered as pests rather than NTOs or beneficial species.

Keywords that were used in different combinations include:
Maize, transgenic, insect-resistant, Bt, Cry, non-target organisms, field trials, interferences, plot size, statistics, environmental risk analysis, power analysis
Dispersal, mobility, flight, migration, redistribution, mark-recapture, scale effects, spatial-temporal, release-recapture
Insects, Coleoptera, carabid beetles, rove beetles, syrphid flies, Lepidoptera, ladybirds, Coccinellidae, chrysopidae, Chrysoperla, spiders, Anthocoridae, Orius, beneficial insects.
Using citations in articles and ‘cited by’ links were often helpful to find relevant related sources in earlier, or later related papers.

3. Background of current field trial methodology.

The choice of non-target species

Environmental risk assessment for transgenic insect-resistant crops has much in common with an eco-toxicological approach as is used for pesticides. Based on this an extended set of criteria is available to select and prioritize non target organisms to be studied in risk assessment studies for transgenic crops (Todd *et al* 2008, EFSA 2010, Sanvido *et al*. 2012). Their selection criteria are - though not systematically- included in the current methodology described below.

NTO studies usually start with laboratory studies measuring direct effects of the Cry-toxins either as purified proteins or expressed in GM crops. For a number of representative species the effects of exposure to the toxins at different concentrations on life history parameters (mortality is usually the main parameter studied) is measured.
The test organisms for laboratory studies are often based on their convenience for testing in the laboratory (bio-assay friendly) and adoptability to be bred in sufficient numbers. In practice this means that the same test species are chosen as those in pesticide effect studies. These species are considered representative for agro-ecosystems in general, including major field crops such as maize.

Fitness parameters such as fecundity, development and mortality are mostly used to assess effects. Other ecological traits that contribute to survival such as temperature tolerance, walking activity or flight capacity, behaviour are less easily studied under laboratory conditions. Because of this and many other reasons the laboratory studies are not used as a tool to predict the potential occurrence of effects in the field, but serve as an indication for potential effects. But they are also essential to select species for further study and to acquire knowledge about toxicity mechanisms and exposure routes.

Laboratory studies testing Cry-toxins as expressed in Bt maize thus far have shown that the direct effect of Cry toxins are quite specific for the level of insect orders (e.g. Lepidoptera) or family (Chrysomelidae) . Even though quite a number of cross-order effects are reported from laboratory studies, the evidence in not strong for most cases and those effects found are usually (but not always) in the low toxicity range (van Frankenhuizen 2013, Lövei *et al* 2009). There is hardly any evidence for such cross orders effects from field trials thus far, even though many different species are included in those trials. Further studies and the introduction of new Bt toxins may change this. A conclusion based on laboratory studies is that direct effects are most likely to occur on species that are taxonomically related to the target organism *i.c.* Chrysomelidae (Coleoptera) and Lepidoptera for the currently marketed GM crops. Though these taxa include many species, most of them are not associated with the maize crop itself and hence not exposed to the toxins. But several species can be associated with weeds in de maize fields or with natural vegetation in the field-surrounding habitats.

Going from the lab to the field more criteria are relevant for species selection because many species are linked to the maize ecosystem (Meissle *et al* 2012) and hence directly exposed to the toxins through feeding on plant parts (leaves, pollen, seeds, roots or plant remains) or indirectly exposed due to the food web links with the target organisms. The effects through the food web may be due to carry-over of Bt-toxins or to the amount of food available (the latter is an indirect effect of the Bt toxins).

In this way many organisms in the maize-ecosystem are potentially exposed or affected by the toxin itself or by changes in the food web structure when transgenic Bt-maize is grown.

Further criteria have been formulated for selecting species for non-target field trials (Todd *et al* 2008). Apart from hazard, potential exposure and food web links, ago-ecosystem function and conservational value are important criteria. For this reason insect predators, and less often parasitoids and pollinators are included in most of the field trials. With species of conservation value most attention is given to species (such as butterflies) that visit or live close to maize fields. These studies have been scarce, however.

A final criterion that often overrules the former criteria is that the species to be studied should be assessable in the field. In practice this means that the species should be abundant as the power to detect difference is statistically mainly determined by abundance (Albajes 2013). Further they should be easy to sample by traps or to be easily observed, and identified. It should be stressed that species tested in the lab often do not occur or are hard to assess in field trials. For both field trials and laboratory studies 'surrogate' species are studied that are easy to sample and handle and considered representative for the crop. Apart from effects on specific non-target organisms most field experiments are aimed at detecting any unexpected effect on whatever organisms that can be measured. In practice many field experiments tend to study all organisms that can be sampled in order to find such effects. Later in this report the potential friction between different criteria for species and their experimental properties such as abundance and accessibility will be discussed at various points in this report.

Taking into account the criteria above the taxa listed below are most commonly chosen to study adverse effect in non-target trials for transgenic crops. Each group is shortly described and information is given that is relevant in the context of field trials.

Thysanoptera (thrips). Some species specifically feed on pollen and leaves, but are mostly considered as a pest species rather that as NTO. However, they are important as prey for predators such as pirate bugs (Orius spp.) and predatory mites.Thrips numbers can fluctuate strongly over the growing season and due to their rapid development they can reach high abundances They are readily caught on sticky traps but they are difficult to count or to identify to species level. In particular the larvae are hard to study also because their tendency to hide between plant parts. Mass flights of adults may occur at particular weather conditions and when the habitat is disturbed (at harvest for example).

Aphididae (aphids). Several species of this non-target group are phloem feeders on maize leaves . Due to potential damage mostly seen as a pest species rather that a non-target, however they are important as prey for predators for other pests Numbers or levels of infestation can relatively easy be estimated by visual observation of plant parts. Numbers typically start with a few settlers from the environment or other crops and numbers tend to explode later due rapid reproduction and development.

Cicadellidae (leafhoppers). Including phloem feeders and parenchym suckers, some species feeding preferably on maize, many others on weeds and plants in field margins. Nymphs and adults are important prey for several predators. Adult activity-density is estimated by sticky traps. Larvae can be assesses by leaf samples.

Lepidoptera (butterflies and moths). Some Bt proteins that have been introduced in GM crops target lepidopteran pests. Larvae of many lepidopteran non-target species may be assumed to be sensitive to different Cry1 proteins and species that not feed on maize plants may be exposed to maize pollen that are dispersed from the field into the field margins, or that are deposited on weeds in the crop. In most cases population effects are studied by counting caterpillars on plant parts. Adult butterflies and moths are very mobile and hard to count.

Carabidae (ground beetles). Mostly generalist predators that are assumed to be only potentially affected by food web connections to maize feeding prey, Carabids can be conveniently sampled by pitfall traps which give an indication of density and activity of those beetles. Though not closely related to Cry3 sensitive chrysomelid beetles some sensitivity may be expected. Ground beetles are ecologically well studied and populations seem to be fairly stable from year to year at a single place with the same crop, some species having a wide distribution in different arable crops. The group includes very mobile but also some quite resident species.

Staphylinidae (rove beetles). Mostly generalist predator assumed to be affected by food web connections to Cry containing prey, conveniently sampled by pitfall traps. Predatory and non-predatory species are not easily distinguishable by non- specialists. Numbers can vary strongly due to aggregation and mobility.

Chrysomelidae (leaf beetles). This group is of particular importance because several currently used Cry toxins are targeted against species belonging to this family (Corn root worm, *Diabrotica virgifera* Leconte and Colorado beetle, *Lepinotarsa decemlineata* Say), and more species from this groups are considered as pest species for other (currently not yet transgenic) crops such as flea beetles in oil seed rape. But also on weeds and wild plants growing in field margins *many* species of this family occur and maybe exposed to transgenic maize pollen. So species of this group are interesting for effect studies. Many species are strong flyers and may be caught in all kind of traps such as sticky traps, water pan traps or window traps.

Coccinellidae (ladybirds). Predators with both adults and larvae predate on fytophagous insects such as aphids and leafhoppers, but also mites that feed on maize. Easily caught on sticky traps and water pans and observed by visual inspection. Adults are good flyers but larvae are stationary and easy to count. Numbers tend to fluctuate from year to year dependent on winter conditions, and adults have to colonize fields from natural environments or field margins.

Aranea (spiders). Including orb web spiders, wolf spiders and hunting spiders, that are assumed to be affected by pollen that are attached to webs and by all kind of prey items including species feeding on maize. Spiders are sampled by pitfall traps or by visual observations.
Anthocoridae (predatory bugs). In particular *Orius* is studied as a predator of lepidopteran eggs and mite species feeding on maize, aphids and leafhoppers that may contain Cry. *Orius* also feeds on pollen. Adults are sampled by sticky traps, but the nymphs (larvae) are not easy to sample due to their hiding behaviour between plant parts.

Chrysopidae (lace wings). *Chrysopa* species are frequent predators of small insects (aphids, thrips and leafhoppers) in maize. Adults are easily caught on sticky traps and larvae observed by visual inspection. Larvae can be numerous and move only over small distances. Numbers of adults tend to be low and therefore sensitive to sampling problems.

Diptera (flies). In particular hover flies (Syrphidae) are sampled by sticky traps and water pans and their larvae are counted by visual observation. The adults are pollen feeders and pollinators and the larvae are predators of aphids. Many other fly species of different families may occur but they have a wide variety of feeding habits.
Other groups that are occasionally included in effect studies are pollen feeding Nitidulidae, Ichneumonidae (parasitoids of caterpillars), and Apidae (honeybees and bumblebees), Phytoseiidae (predatory mites that feed on phytophagous mites and thrips).

Choice of Effect Parameters

Field trials are aimed at providing information about potential adverse effects of transgenic crops on the abundance and diversity of non-target organisms. Without going into detail about the scientific impact of this aim it should be stated here already that for reasons of feasibility, current field trials focus on what is practically measurable under those circumstances.

Based on scientific consensus this means that almost all studies assess numbers of individuals of a defined taxon per unit of sampling effort. Any change or difference in this response parameter can be considered as a measure for a potential effect, and a significant effect when appropriate statistics have been applied. Whether a statistically significant effect is regarded as ecologically significant is a matter of debate. When a statistically significant is not consistent over time and over different locations it is often considered as ecologically insignificant.
Measuring numbers of individuals by a range of different sampling methods is an common way to estimate abundance and assumed to in indicator of population density in the timeframe and scale that it is measured. It should be noticed that real absolute densities (individuals per square meter) are rarely measured.
It is clear that treatment effects in field trials on abundance can only be established if the assessment methods are sound, performed in the same way among treatments and if measurements between plots are independent. Notice here that movement between plots or between plots and the environment is disturbing this independence.

It should be realized that changes and differences in abundance can be due to many factors including differences in reproduction, mortality and development, but also due to immigration and emigration. Factors that influence movement –such as food deprivation or other adverse conditions- also affect the 'apparent' abundance when the assessment method depends on activity (such as most traps). All these factors should be taken into account when evaluating the data that are collected in field trials.

Most field studies are focused on finding difference in local momentary abundance instead of populations trends. To study population trends the spatio-temporal context is extremely important because average levels over time and space are crucial for maintenance and constancy in ecosystem services. To follow demographic changes in populations repeated (preferably absolute) density estimates can be taken.

Diversity measures are derived from abundance measures for the species that occur in samples with one or more individuals. So every inaccuracy, bias or methodological drawback that is attached to the abundance assessment method is reflected in diversity measures.

Using above-species-level parameters

Apart from numbers of individuals as a measure for abundance for separate species other parameters are used in risk assessments.

Species in taxonomic or functional groups are frequently pooled as one response variable. This is often justified by lack of expertise or time to identify all specimens or because the numbers of separate taxa are too low to enable proper statistical analysis. Combining ecologically related species as a functional group may have its merits, but there is a risk that species are lumped that are ecologically very different, for example summing up large species with small species knowing that body size is strongly correlated to predation capacity. For example in carabid beetles biomass (and hence food uptake) per individual for big species can be 100 times that of individuals of small species. Translating numbers to biomass may diminish this problem but is never done. Even worse is that taxonomic closeness is often confused with function e.g. staphylinid beetles include detritivorous as well as carnivorous species. Moreover, density fluctuations, phenology and spatial-temporal behaviour varies much among species.

Other ecosystem response parameters such as diversity measures and multivariate approaches are outside the scope of this report. In both cases however any potential species specific are hidden in community response that is usually much more complex than effects on single species.

4. Design of experiments: statistical and ecological constraints

Under European GM regulation on the deliberate release of transgenic crops, a GM crop is only admitted to the EU market if scientific evidence demonstrates that there are no unacceptable risks for human health and the environment. Its aim is to safeguard biodiversity including non-target organisms. Before introduction of insect resistant crops on the market for cultivation can take place, field tests to study potential adverse effects on NTO's are required that are scientifically sound and provide sufficient information to warrant safety.

Even though guidelines from the EFSA are available (EFSA 2010), debate and research is going on among researchers about how adverse effects on non-target organisms should be studied in field trials. Studies aim at detecting effects in the field that are inferred from lower tier studies or from knowledge that indicates a possible effect based on ecological relations in the agro-ecosystems and the surrounding environment. In addition also studies are performed to detect any unexpected effects on species or ecological communities.

In order to carefully design useful experiments, the experimenter should have a reasonable idea about the effect size that is relevant and could be detected in a well deigned field experiment. General methodology for such trials has been proposed and discussed in several recent papers. Two recent papers provide extensive overviews in particular about the statistical aspects that are important in the set-up and data analysis of field trials for NTOs (Perry *et al* 2009, Semenov *et al* 2013).

For conducting proper experiments the following issues should be included in order to design useful experiments for non-target organisms in transgenic crop environments. A similar checklist is given by Semenov *et al.* (2013)

1. Problem definition in the context of insect-resistant transgenic crops and related to the current and near-future cropping practices and geographic regions
2. Breakdown of the problem into clear questions related to ecological risk assessment
3. Defining non-target organisms to be studied
4. Formulating testable hypothesis that give answers to the questions with a defined (un)certainty level
5. Defining system characteristics, spatial dimensions and boundaries
6. Defining treatments and controls as comparators
7. Deciding on statistical difference or equivalence testing, or both.
8. Deciding on required significance levels and statistical power
9. Precisely defining parameters to be measured in order to test the hypothesis
10. *Ex-ante* evaluation of the type and statistical characteristics of the data to be collected and choosing the appropriate statistical analyses for the type of data and relationships to be tested
11. Selecting methods and sampling plan to measure the parameters with a defined precision including sample size based on a priori knowledge about variance and spatial aggregation parameters of the species to be studied?.
12. Defining on the number of replicates in time and the arrangement in space based on the expected or *a priori* known variance in the response parameters (mostly average number per sample).
13. Designing the layout, scale and spatial setup of the experimenting such a way that data can be statistically analysed, confounding factors are minimized and interference between treatments, between plots and the surrounding environment is avoided.

Without going into detail about all the underlying methodological issues that can affect the outcome, all these issues have been used for this project to evaluate current practices in field trials and identify lack of methodological robustness in current NTO research. To what extent field experiments can fulfil the criteria for an useful experiment is discussed in later in this report For a further discussion on statistical aspects in field trials for non-target effects of GM crops the reader is referred to the reviews of Perry *et al* (2009) and Semenov (2013).

5. Analysis of Field Trials

There have been numerous studies that were aimed at testing whether any adverse effects of transgenic insect-resistant crops can be detected under field conditions. For this evaluation –that was chosen to be focused on maize- a range of papers have been considered that are mainly from the last decade. It includes experiments being carried out in the USA and Europe and only those where transgenic Lepidoptera-resistant Bt-Cry1Ab1 maize crop (LRM) or transgenic Coleoptera-resistant Bt-Cry3Bb1 maize crop (CRM) was tested against isogenic non-transformed maize either untreated against insects (ISO-) or treated with a chemical pesticide (ISO+)or with a Bt-spray formulation ISO-Bt.
In this project field trials for other transgenic crops were not included even though they may provide other useful methodologies or experimental set-ups. It is assumed however, that field trials conditions for most arable crops will be similar.

To evaluate the field trials a simple format is used including:

- The treatments that were compared in the experiment
- the design of the experiments (replicates, plot-size and field layout)
- the sampling methods used
- the species(-groups) studied
- the taxonomic resolution (whether specimens are identified to species or higher taxonomic level)
- statistical data analysis
- the -for this report relevant- conclusions that are drawn and issues discussed
- short comments that were formulated by the author of this report to list experimental shortcomings and limitations

In Appendix I a comprehensive overview of the most relevant papers about non-target studies with Bt-maize is given including experimental details, authors conclusions and evaluation comments. The list of field studies does not provide a complete overview of the field studies that were published. Only recent and in the view of the author relatively well-designed studies were compiled for this report. Several other studies had a very limited scope or were judged to be poorly designed. Studies that were not public accessible were not included either.

In Table I, data from this information is summarized in a simplified way to review the approaches and outcomes of recent field trials and to give a quantitative impression of the type of data collected.

Table 1 Overview of the non-target studies analysed for this report, summarizing quantitative experimental aspects in a simplified form.

Publication first author	Comparison[1] CRM LRM ISO- ISO+ ISO-Bt	Sites	year	Replicates per site	Plot Size (ha.)	plants visually inspected	Pitfall traps	Sticky traps	Water traps	Soil cores	Taxa sampled[3]	Numbers/ sample/ plot [4] low/ mid/ high	Analysis [6]	Single species analysed	Groups analysed	Range of effects[5]	Significant Comparisons[7]
Ahmad 2005	CRM ISO- ISO+	3	2	4	0.03		4 (3x)			8	Ca, St, O, Mi, Cl	1-4-15	ANOVA		9	0.5-1.4 (0.4)	0 (18)
Ahmad 2006	CRM ISO- ISO+	3	2	4	0.03	5 (2x)					Co, Or	0.1- 0.5-3.0	ANOVA	4		0.7-1.5 (0.7)	0 (8)
Bhatti 2005	LRM ISO- ISO+	1	3	4	0.03			3			Co, Or, Lh, Ch Chrysopa	1-20-100	RM ANOVA	7		0.8-1.2 (0.2)	1 (27)
Bruck 2006	LRM ISO- ISO+	1	2	3	0.36	10	5	5			Sp, Lh	0.5-1.5-4.0	ANOVA	1	4	0.25-1.5 (0.15)	1 (12)
Candolfi 2004	LRM ISO- ISO+ ISO-Bt	1	1	3	1.4		12 (8x)		5		Ca, Sp, St, CL, Fl, Lh	Not provided	Canoco	7	15	0.2-1.3 (0.1)	0 (15)
Daly 2005	CRM ISO-	2	2	4	0.05	10-20	2-3				Lh, Ch	0.3-1.0-3.0	RM ANOVA		5	0.5-2.0 (-)	1 (20)
De la Poza	LRM ISO- ISO+	2	3	4	0.5	10-25 (5x)	4 (5x)				Ca, St, Sp, Or, Co, O	0.3-2-5.0	ANOVA	pooled	6	0.3-1.3 (0.3)	0 (18)
Farinos 2008	LRM ISO- ISO+	1	3	3	0.5		5 (8x)				Ca, St, Sp. O	0.5-2-10	RM ANOVA	0 (37 identified)	5	0.5-1.5 (0.6)	1 (15)
Gathmann 2006	LRM ISO- ISO+	2	3	4	0.25	10 (2x)					O, (Lepid Larvae)	2-5-20	ANOVA, Equivalence	2 of 8		0.7-2.7 (0.15)	0 (6)
Habustova 2013	LRM ISO-	1	3	5	0.50	4 (					Th, Ap, Or, Co	0.2-2-25	ANOVA, Canoco	9		0.5-2.5 (-)	0 (27)
Ludy 2006	LRM ISO-	3	3	1	2	10 (5x)					Sp	1-5-15	RM ANOVA		5	0.2-1.8 (-)	0 (15)
Orr 1997	LRM ISO-	1	1	3	0.4	5 (3x)					Or, Co, Chrysopa	0.2-0.8-2.0	ANOVA	3		0.7-4.0	1 (18)
Priesnitz 2013	CRM ISO-	1	3	8	0.13		1				Ca	2-10-100	Equivalence	15 (of 70 identified)	1	0.2-2.5 (-)	2 (36)
Rauschen 2008	CRM ISO- ISO+	2	2	4	0.25	5 pl + sweep net		1	1		Lh	5-20-50	ANOVA Equivalence	2 (of 5)		0.7-1.0 (0.025)	2 (4)
Rauschen 2010	CRM ISO-	2	2	4	0.25	4-10+ sweep net					Co + Ch +O	0.1-2-10	ANOVA Equivalence	4 (of 27)		0.6-1.5 (-)	0 (8)
Stephens 2012	CRM ISO- ISO+	3	3	4	0.25		20 (20x)				Ca	1-2-4	Chi-square		1	0.5-2.0 (0.5)	0 (14)
Svoboda 2013	CRM ISO- ISO+	1	3	5	0.5		5 (7x)				Sp	0.2-1-5	Canoco	3 (of 29)		0.9-1.1 (0.5)	0 (9)

1 LRM: Lepidoptera-resistant (Cry1), CRM: Coleoptera-resistant (Cry3), ISO-: isogenic not resistant maize without insecticide treatment;), ISO+: isogenic not resistant maize without insecticide treatment; ISO-Bt isogenic not resistant maize with Bt spray treatment

2 Plot size: the size of a single experimental unit in hectares

Taxa: Ca:Carabidae, St:Staphylinidae, Co:Coccinellidae, Ch Chrysomelidae (flea beetles) Lh: leafhoppers, Sp: spiders, Or: Orius; Mi: mites, Cl: Collembola; Fl (flies, mostly syrphids); Th thrips, Ap: Aphids O:other groups

4. Numbers are expressed as numbers/plot per sample occasion in the way they are provided in statistical analysis for the publication (mostly appearing in tables or graphs). The low, median and high level indicates numbers for low, median and high range for the species regarded as abundant enough for analysis It summarizes the range of 'abundances' that researchers are considered as high enough to warrant some kind of analysis. It also exemplifies that in many cases these numbers are low to very low.

5. Effect ranges are given for the largest effects found for the taxa studied (independent if they were statistically significant, which was in fact only presented in a few cases). It is expressed as the ratio of numbers observed in transgenic / isogenic. Between brackets the largest negative insecticide effect is given as the ration insecticide treated isogenic / non treated isogenic

6. RM ANOVA : repeated measures ANOVA dealing with samples repeated over time

7. Number of comparisons made for NTO's in the study that showed statistical difference between the BT transgenic crop and its isogenic comparator. Between brackets the number of comparisons made, which is mostly the number of taxa analyses x the number of years. All significant differences found in these studies are inconsistent over years, locations or sampling methods!

From the evaluation of case studies compiled in Table 1. the following conclusions are drawn.

1. Most studies more or less follow the EFSA (2010) guidelines for testing potential effects, but these guidelines are not case specific. Unfortunately most studies do no reformulate, refine or specify the problem definition in order to have more concrete and testable hypothesis for their specific case.
2. It is common practice to set up experiments in a randomized block design with 3 or 4 replicates (rarely more), to repeat the experiments for 2 or 3 years and to use plot sizes ranging from 0.10 (ca. 30*30 m) to 0.50 ha (ca. 70*70m) without testing or discussing whether this plot size is suitable for the parameters (or species) to be tested, or representative for the cropping practice.
3. It is also common to use pitfall traps, sticky traps and visual observation without discussing whether these methods are the best ones to measure effect parameters (end points) for specific species (also because specific endpoints are mostly not defined).
4. There seems to be a tendency to choose familiar sample methods for insects and perform data analysis on species or species-groups that are trapped in sufficient numbers rather than species that were assumed to be potentially affected by the treatments.
5. Sampling plans are rarely based on *a priori* knowledge or expectations about the variation in abundance and the sampling efficiency or precision. In many cases common practices are applied, often leading to ineffective sampling yielding numbers of individuals per taxon that are too small and variable to be statistically analysed.
6. There is a wealth of response parameters measured in the trials in order to find differences or effects that can be discussed as adverse, significant or insignificant, inconsistent, or hard to explain and worth to be further studied. The fact that when many parameters are measured, to find by chance at least one to be significantly different has not been considered in any of the publications.
7. Sometimes observed statistical differences are considered as biologically insignificant because they occur only in one particular week or year, or are inconsistent over years or sites.
8. In part of the case studies, non-target organisms are identified to species level, but even in those studies there is a tendency to pool species in larger functional or taxonomic groups without discussing whether this is justified (Farinos *et al* 2008, Priesnitz 2013 *et al.*) in view of ecological relevance.
9. Compared to the large number of taxa sampled and identified only a few taxa or pooled taxa are analysed. Overall of 270 comparisons made in the reviewed case studies only 9 (3%) resulted in a significant difference between BT-transgenic crops and its isogenic comparator which is likely to occur by chance, in particular because all those detected difference are inconsistent over year, locations or sampling methods within the case studies. A rough analysis shows that the frequency of significant effects was not clearly related to effect size.
10. Often only a few species of many have sufficiently high numbers to be analysed. This means that only the most abundant, common and in many case being highly dispersive species are evaluated in field trials.
11. Comparative analysis of experiments of different studies is strongly hindered by the variation in the taxa sampled, sampling techniques, sampling plans and the way data are processed, analysed and presented.

Reviewing the field trials summarized in Appendix I and Table 1, the first conclusion is that in most field experiments a compromise is being sought between the aim to obtain statistically interpretable data, the costs involved in managing field plots and sites and the labour to obtain ecological data.
Within the general guidelines set by EFSA (2010) there is a lot of freedom for researchers and applicants to determine how the field trials are designed and about which and how ecological parameters are measured. Regarding the available resources for field trials and the necessary adaptation to local ecological conditions this approach is understandable.

The message from the many laboratory and field experiments and reviews of all those experiments (Wolfenberger *et al* 2012, Devos *et al* 2012) is that it appears very difficult to detect any significant effects of currently applied transgenic Bt-maize, at least compared to the often slightly stronger effects that are frequently found in the same experiments when insecticides are applied (dependent on the toxicity of the pesticide).

The conclusion - drawn by the authors of most studies- that any unintended effects of Bt-maize are absent or at least very small and insignificant may be defendable. However, there are certainly weaknesses in the methodology as argued above that may prevent detection of small effects just as they may prevent detection of smaller effects in pesticides. Effect size below 50% is rarely statistically significant due to variability and the fact that numbers measured for over 90% of the species are in the low to very range. The small plot size in relation to mobility in most species (see below) is likely to cause frequent plot interference.

A power analysis is rarely performed and if so ex-post. A thorough power analysis on more existing data could be useful but in those cases where such an analysis was done it appeared that the data are rarely not good enough even to show effects outside the 0.5 or 2.0 effect level range. These assessment endpoints are the ones which are often considered to be ecologically relevant (e.g. Naranjo 2005, Albajes *et al* 2013, Rauschen *et al* 2010). This indicates that effect levels should be stronger and less hindered by the occurring variability.

6. Impact of insect movement on field trial performance.

The evaluation of field trials described above shows that (significant) effects of transgenic crops on numbers of non-target insects are rarely found. Is the impact of current Bt-maize indeed negligible? Or does variability prevent to detect the signal between the noise? Or do organisms move too frequently between plots and preventing any differences to become and stay apparent?

At least, even increasing experimental robustness (e.g. using more replicates) and advanced statistical methods did not result in more detected differences. But realizing that plot sizes rarely exceed the scale of 0.5 hectares and knowing that most arthropods are continuously on the move, plot interference may hinder detection of small but relevant effects. For example what does it mean when over 30% of the non-flying ground beetles are exchanged between 2 adjacent fields of 4 ha within one season Thomas *et al* (2006)? It is likely that spatial processes will interfere with the effects of treatments when plots are small in relation to the mobility of the species studied.
In this chapter the problem of insect movement and scale of field trials will be explained and documented to show the intrinsic problem for field trials with highly dynamic organisms such as those measured in the studies discussed. It is not possible to provide a clear-cut / standardized overview for mobility species studied in NTO field trials. This is because mobility data are scattered and dispersal distances (as meters travelled per day) are extremely variable within species and depending on field conditions, time of the year. To get a picture of common distances and order of magnitude in dispersal, a number of studies are reviewed below, focusing on species that are relevant for NTO field trials. Information on mobility is retrieved from both agricultural literature but also from natural habitats. Temporal and spatial heterogeneity in the abundance of insects at different scales is a basic factor that has to be taken into account when doing field experiments with naturally occurring populations. Due to limited resources (money and time) ecological experiments tend to be performed at scales that are often smaller than the scale of systems and processes of interest (Scale and Cooper 2003).

In the course of a typical summer month an estimated three billion insects fly through a 1 km^2 'window' of sky in England (Chapman *et al* 2003). Realizing these mass movements of insects should make ecologists aware of the problems when measuring densities at a specific field at a particular moment. Although the broad impact of insect movement on the local distribution, abundance and species interactions is (Turchin 1998), it is not easy to quantify the impact in field trials because movement is hard to measure and can be very variable under different conditions. Many insect species in agro-ecosystems cope with the unstable and transient conditions by strong dispersal capacities and responding to adverse food or breeding conditions. Cross-habitat movement in agricultural landscapes has a big impact on the phenomena measured on a smaller scale and is the major factor in the maintenance in the abundance and diversity of species (Tscharntke 2012). Also among pest species that occur in various crops short range dispersal rates ranging from 10-300 meters per day are common. For a review see Mazzi and Dorn (2012).
However for many species real data on movement is hard to obtain due to methodological problems and complex dispersal behaviour of most organisms. Therefore information remains scattered and to limited to understand the impact of movement on the dynamics of most species. New methods such as genetic approaches, mark-release-recapture experiments and new tracking techniques confirm that many insects and spiders are able to disperse at high rates and frequently move between fields and between fields and surrounding areas. This is in particular true for species adapted to the dynamic nature of agricultural fields, such as predators and pollinators that depend on local temporal food sources, but also for many phytophagous species, including insect pests.

Recent genetic studies, for example show that there is little genetic structuring in several agriculture-bound species indicating that gene flow and hence exchange is strong and that they move over large distances (see references in Raymont *et al* 2013). This has been shown for example in aphids, planthoppers and in syrphid flies that can easily bridge distances of hundreds of meters or even kilometers in a short time.

At a smaller scale it has been shown for ladybirds (*Hippodamia* spp.) and pirate bugs (*Orius* spp.) that, when crop attractiveness or prey abundance changes, they can easily move over distances of 20-30 m/day (Prasifka 1999). In landscape studies it was shown that cropping practices, prey abundance and colonization conditions equally impact the build-up of populations of spiders, ladybird beetles and spiders in fields. In those cases predator abundance is strongly influenced by immigration from nearby uncultivated fields at a scale of 1-2 kilometres (Prasifka 2004). Mass movement between fields with clover and maize (in response to crop status and prey) was observed in Egypt for Orius bugs, ladybirds and lacewings and spiders with populations doubling in a couple of weeks due to frequent immigration (Shoeb *et al* 2008).

For ground beetles (Carabidae) many movement studies have been performed at different scales showing that movement is driven by hunger levels and adverse conditions and that the larger species can easily walk 20 meters or more per day. Most of the smaller species are winged and massive movements occur the beginning of the cropping season and after harvest. The fast colonisation of arable fields is well documented, as well as the recovery of populations after insecticide sprays in one or two weeks. The population dynamics of most species is determined by processes above field scale (Holland *et al* 2005).
Mass movement of many other insects also takes place in early spring not only by many natural enemies but also by pest species that colonize the fields from their overwintering habitats and hosts. The settlement in the fields depends on the environment of the cropped area and distances among the patches (Prasifka *et al*. 1999).

Also during the season redistribution by movement is likely to be common. Decrease or increase in numbers in specific patches does not necessarily mean that the 'population' as a whole at landscape scale is changing. For many crop bound species, the field level does not necessary reflect the overall population level as many species occur in various habitats at different densities. The spill-over from relatively good to less favourable patches is an issue in current ecological theory about source-sink dynamics which could lead to a better understanding about whatever effects of agricultural practices really means for population maintenance (Sisterson 2007). For several spider species the densities at the field level are determined by the landscape complexity at a scale starting from 10 ha up to 2000 ha (Schmidt *et al*. 2008)
It is clear, however, that frequent random movement may easily level out differences among plots at a lower scale. By contrast it may also exaggerate a difference between plots when insects easily move from the non-preferred to the preferred plot, without causing mortality or population decrease. When an increase or decline in numbers is locally observed this can be interpreted as a n effect of plot treatments but also as a transient shift in spatial distributions. So in particular at smaller scale differences can just arise or fade away depending on the mobility of the species studied.

To what extent all the movement phenomena described above disturbs the reliability of current field trials for the most commonly studied species is hard to say. Semenov *et al* (2013) provide a table presenting the required minimal plot sizes to avoid plot interference for species in different dispersal categories based on a small set of references, noting that substantial buffer zones should be used to avoid spill-over. They suggest that for species with fairly high or high dispersal rates plot sizes from 0.5 ha – 2.5 ha are most suitable to avoid interference. Only for slow moving species (e.g. springtails and mites) and immature stages of many insects, smaller plot sizes are acceptable. Their conclusions are based on a limited number of papers that studied dispersal in different context and it's unclear how they derive plot size from those studies. A real test or modelling approach for plot interference for field trials with measured data on dispersal capacities of different species in the field, however, is not available so far.

In the current practise of using plots of 50x50 m or 70x70 m in field trials to study effects of insect-resistant crops or pesticides on non-target insects and spiders, it seems likely that these effects are more easy to detect for species that do not move more than a meter per day than species that easily move 10 meters per day. Daily exchange of individuals between plots may tend to wipe out differences in mortality between plots that may take weeks to become apparent. Unfortunately the species that are numerous or abundant enough to do statistical analysis, such as the common ladybirds, syrphid flies, ground beetles, *Orius* bugs and many spiders, are typically highly mobile (moving tens or even hundreds of meters per day), blurring local temporal effects but also causing variability among sites and years as they respond to heterogeneity at a larger scale.

Another experimental problem of mobility occurs for the interpretation of data from pitfall traps and sticky traps that are extensively used in NTO studies. In such traps the number of insects trapped directly depends on mobility. The more they move the higher their abundance seems to be. Though this problem has been acknowledged already many years ago (Thiele 1977), ecologists have settled down with the idea that there is hardly an alternative and they accept the bias and uncertainty in this trapping method, not only for ground beetles and rove beetles but also for spiders (Topping and Sunderland 1992). Most real density estimates require very laborious methods such as taking soil-core samples, using suction-traps with high efficiency and taking whole-plant samples. When such methods are used and repeated in time they can give a much better picture of the real demographics than pitfall traps. Nevertheless pitfall traps and sticky traps remain mainstream methods for sampling a wide variety of insects, and the term 'activity/density' is fully accepted. As long as the characteristics of experimental plots are not likely to influences activity, this parameter may be acceptable as an indicator for difference in 'abundance'. However, in case the transgenic crop affects microclimate or prey abundance, this likely affects walking behaviour in ground beetles and hence the chances to be caught (Thomas *et al.* 2006). Also differences in prey- or host densities are known to influence movement of predators and parasitoids and hence cause a false picture of 'abundance'. Hungry animals tend to move more and increase the apparent abundance under unfavourable food conditions. When prey is scarce predators may suffer from lack of food, but the when their searching activity increases, the measured 'apparent abundance' in traps may increase. In this way crops may look more favourable than they really are.
All these mobility factors may be considered as factors of less significance when treatments cause strong and long-lasting effect on organisms. However, when one is interested in smaller and more subtle effects, they all become relevant. In order to make better decisions on plot size in NTO field experiments mobility information should be taken into account to avoid extensive exchange between experimental plots and between plots and the natural environment.

7. Ecological properties of insect groups and assessment constraints

This chapter will shortly review experimental constraints for a number of non-target insect groups and exemplify issues mentioned in earlier chapters. It may help to bring more focus in field studies for non-target insects taking into account spatial aspects of their ecology as these are rarely incorporated in set-up of field trials.

Lepidoptera

Butterflies and moths are of special concern as the Bt-Cry1 toxins targeted at maize stem borers are known to affect many other lepidopteran species in laboratory conditions (Adinda *et al* 214). So when non-target species would be exposed and ingest those toxins in field condition they are likely to be affected. The concern for risk managers obviously is not for species living on maize as these are considered as primary or secondary pest species, but also non-target species living on weeds or on other plant species in the neighbourhood of the fields could be exposed. Not many species are directly connected to the maize plants, but no complete inventory could be found for species that may feed on weeds in the maize fields or plant species that frequently grow in the field margins of maize fields.

Studies have shown that pollen –that are taken up by caterpillars when feeding on leaves may spread in a small zone around maize fields but only a few studies are available on species that live on non-maize plant species growing in or near the fields (Gathmann 2006).

Butterflies and moths can be strong fliers (bridging distances of more than 100 m/day). Although spatial studies are available from nature conservancy, little information was found on dispersal for species in agro-ecosystems with relevancy for field trials. Well studied species such as potato tuber moth and diamond back moth that widespread in agriculture, are strongly dispersive and can move hundreds of meters per day (Cameron *et al* 2009, Mo *et al*. 2003).
Also mark-release-recapture studies indicate dispersal rates of adults moths from 20-100 meter per day (Mazzi & Dorn 2012), but other species like micro-lepidoptera might be more stationary although little information is available. Most adult butterfly species of conservation value are monitored by visual counts following transects.
The larvae stages (the caterpillars) are stationary anyway and move far less than 1 m/day). Caterpillars are mostly sample by counting their presence on whole plants, though with younger stages this may be laborious. While the maize ecosystem is expected to be relatively poor in non-target butterfly and moth species, the focus could be furthers shifted to weed-feeding caterpillars in the field margins which are exposed to wind-born pollen from the field. In parallel with survival studies the uptake of toxins should be monitored.

Coleoptera

Non-target species are potentially vulnerable for Cry1 proteins (in maize targeted against the corn root borer), as these toxins were show to affect several species to some extend in a wide range of beetle families under laboratory conditions. As they tend to be common and easily trapped in agro-ecosystems, beetles are a favourite to be included in NTO effect studies.

Coleoptera constitute a very diverse group of beetle species including many plant feeders (all parts, including pollen), carnivorous predators and scavengers and species having a mixed diet. Also within families species have different feeding patterns. Therefore pooling species from samples into functional groups based on taxonomy can be tricky because families such as rove beetles (Staphylinidae), ground beetles (Carabidae) and even ladybirds (Coccinellidae) include species having very different diets and different functions in the ecosystem. Also the variation in body mass (and hence predation capacity) per species does not allow species to be summed as an estimate for the functional capacity

Within the above mentioned groups and other beetle families that inhabit arable fields, most species have a good dispersal power (in the range of 10-500 meters/day for many flying species), hence quickly responding as colonisers of place with favourable conditions. For these flying beetle species local numbers tend to fluctuate strongly due to rapid immigration and emigration. So the suitability of species for effect studies should be considered case by case. At least non flying beetle species and larvae of some beetle groups may be suitable for such studies.

Carabidae

The bias towards ground beetles (Carabidae) for non-target effect studies is since many species occur in arable cropping systems, they are easily caught by pitfall trapping, they are relatively easy to identify to species level, many species play a role as predator for pest species and their biology is relatively well known. However, high levels of spatial aggregation and heterogeneity of the activity densities is found in field experiments Priesnitz (2013), presumably due to different abiotic such as crop coverage, moisture and soil type (Booij & Noorlander 1992, Thomas *et al* 2006). Distribution patterns in the field with high and low density levels of carabids were evaluated in several other studies (Holopainen *et al*. 1995, Holland *et al*. 2000, Holland *et al* 2005). Abundance of different species varies from year to year and from site to site due to weather factors, soil type and cropping practices (Thomas 2002, Booij & Noorlander 1992). In particular when less than 4 traps are used for abundance assessment and traps are only active at intervals (which is the case in many studies), for many species the number caught is too low to be analysed. Pooling species for data analysis is common practices (and sometimes reduces variance) but ecologically questionable as species vary strongly in size and have different functions (feeding patterns) in the maize ecosystem. Species differ strongly with respect to their function in the ecosystem and predation capacity. Common species are either big and able to spread tens of meter a day by walking (Ekbom 2000) or small and dispersing easily by flight. So one may wonder how suitable ground beetles are to study potential adverse effects of transgenic crops. At least when useful information is needed experiments should be done at the above hectare level (preferably 2-4 hectares) than using plots of less than 0.5 ha (Holland et 2005, Thomas *et al*. 2006). The drawbacks of pitfall trapping to assess abundances is an additional problem as they always reflect a combination of activity and density instead of density alone (Topping and Sunderland 1992, Thomas *et al* 2006).

Staphylinidae

As rove beetles can be common in arable fields including species that have different routes for potential Cry uptake, this group is of interest for adverse effect studies. The maize ecosystem, however, seems to be rather unfavourable compared to other crops (Topp *et al* 2013), and the use of pitfall traps for rove beetles encounters the same experimental drawbacks as observed for ground beetles that are often jointly caught. Species can have very different ecological functions (predatory of detritivorous), vary in size and are very mobile (many species fly > 100 m /day under good weather conditions. Moreover, they are harder to identify to species level than ground beetles. Their abundance is at least as variable as that of ground beetles and so far no single species was suitable in recent studies to be useful for statistical analysis because numbers were too low.

Coccinellidae

Ladybirds are regarded as an important non-target group to be studied as most ladybird species are predators of aphid pests that are able to quickly reduce aphid outbreaks. In particular the voracious larvae are able to consume many aphids per day. The adults are extremely mobile and can fly many kilometres per day under favourable weather conditions (Jeffries 2013). This means that abundance measurements in adult ladybirds at small scale fields can only be viewed as a short term response to good conditions, *in casu* the presence of abundant prey. In mark-release-recapture experiments it was shown that over 80% of the hundreds of released ladybirds in a field of 0.36 ha had left the plot within 24 hours (Van der Werf *et al* 2000). Ladybirds can be trapped by sticky traps or water traps but those measurements reflect activity rather than density. Also sweep net samples are used to estimate more accurate abundances in the crop.

The larvae, however are quite stationary (not moving more than a few plants away from where they hatch) and when established their demographics are likely to be a function of prey density and quality; hence that could well indicate indirect effects or transgenic crops.

Other coleopteran families

Other beetle families such as Nitidulidae can be interesting non-target species to study possible effects when they feed on pollen or maize plant debris. However little is known so far about the feeding habits and ecology of many species, and most pollen feeders are likely to be good fliers due to the fact that pollen sources occur patchy and very temporarily.

Diptera

This insect order includes many fly families that are plant feeding, pollen feeding, detritus feeding, predators and parasitoids, and they can live on all kind of substrates.

Attention has been paid to fly species of which adults or larvae are pollen feeding or predators of plant feeding species such as aphids and leafhoppers that occur on maize. Adults of many species are strong fliers that move between patches where they can find food (pollen) for themselves or prey for their larvae (aphids). Hover flies (Syrphidae) are most frequently included in NTO studies. Species such as *Syrphus balteatus* which is dominant in all agricultural crops all over Europe is an extremely good flier and dispersal is fully overruling local dynamics (Raymond *et al* 2013). Assessing abundance of adult hover flies is difficult because their presence and activity is mainly a function of good weather conditions and nearby sources where they forage or breed (other habitats, field margins). Therefore visual observations, water traps and sticky traps give only a rough indication of abundance in the plot? under assessment. As they respond mainly to actual high prey densities links to transgenic crops effects will be hard to prove if they easily move to neighbouring fields when there is more prey. Assessing the survival of more stationary larval densities by visual inspection of would be better but is laborious.

Pollinating fly species were found to transport pollen (and hence fly) over distances of at least 100 m in one day (Rader *et al*. 2011) As such specialized pollen feeding species may be interesting objects for non-target direct effect studies, but assessing only abundance probably is not the way to detect any as most pollen feeding species are likely to be mobile and feeding on a wide array of plant species outside the fields as well. The added value of field trials compared to lab studies therefore is likely to be limited.

Chrysopidae / Chrysoperla

Lacewings are active fliers that search for different prey types in both natural and agricultural habitats where prey is abundant (e.g. aphids, white flies and leafhoppers).
It was shown by Sivakoff *et al*. (2012) that the adults easily move over distances of several hundreds of meters in a short time when the habitat becomes unfavourable. The larvae can disperse by moving from plant to plant in a crop canopy, but are not likely to move over distances farther than 10 or 20 meters. Therefor lacewing larvae could be good study objects in NTOs trials, numbers should be followed during the season to detect effects. Unfortunately good sampling by visual inspection of plants is quite laborious.

Hymenoptera

For bees and bumblebees ranges of activities easily cover hundreds of meters up to 2 km from their nesting places outside the field. Though flowering maize can be visited when little other pollen or nectar is available, bees prefer other crops and uptake of maize pollen is likely to be low in most situations and probably to infrequent to be of any significance for abundance effects.

Of course parasitoid Hymenoptera come into the picture that oviposit on and develop in the larvae of the target species (the European corn borer Ostrinia and other maize inhabiting species). Potential effects on population levels can be expected due to lower abundance or host quality in maize fields.

Marino and Landis (in Ekbom 2000) for example state that the semi-natural habitats and the landscape complexity is a primary factor in determining the abundance of parasitoids in cropped fields. Parasitoid populations just use the cropped area temporarily to exploit abundant hosts. If these would be unavailable they will suffer from lost searching time. So the overall dynamics at a larger scale is likely to be affected by negative crops factors mainly when the unfavourable crops dominate in the landscape. Of course, when specific insect resistant-crops are grown at a large scale any parasitoids that is strictly bound to the target insect is likely to be reduced in parallel with its preferred host.

The dispersal capacity of parasitoids is probably linked to the size of species, but no quantitative field information was found in literature. Parasitoids can be caught by sticky and water traps, but most studies focus more on parasitation rates (percentage of hosts parasitized) than on density estimates.

Anthocoridae / Orius

Predatory bugs of the Anthocorid family including pirate bugs (different *Orius* species) that are important predators of pest insects such as aphids, thrips, mites, and leafhoppers. Apart from small insects and larvae they also predate on eggs of several insect species and feed on pollen when available. As these food sources all occur on maize, this could makes them basically ideal non-target insects to be studied for adverse effects of transgenic crops.
As *Orius* tend to be relatively common in the maize ecosystem and uptake of maize feeding prey and maize pollen is substantial. Their abundance has also been assessed in some field studies, mainly by sticky traps or by visual observations.
Even though no effects could be shown in these field studies, the statistical power mostly is weak due to variance and sampling problems. Moreover, Orius species are known to be good dispersers that lives in many habitats. Hard figures on dispersal distances are not available, but observed colonization processes suggest than they can move hundreds of meters per day.

Aranea (spiders)

The potential interactions of transgenic Bt crops with spiders have been extensively reviewed by Peterson *et al* (2011). Their work perfectly shows how the complex interactions potentially can be. Through a short time-window spiders are exposed during pollen shed (which adhere to webs) and all over the season by predation on a wide range of prey type including aphids and leafhoppers feeding on maize.
They also mention the huge variation in natural densities (among agricultural fields 100 fold or more), the variation in species composition and the strong dominance of some species among the great richness in other species.
Due to the variability powerful statistics can rarely be applied to the individual species while this actually necessary with regard to the variation in ecological characteristics between species. The implications of movement by ballooning for experimental plot interference is not mentioned in that study, this is a very important characteristic in the dominant orb web spiders that inhabit agricultural fields. In good weather conditions many spiders bridge distances of 10 to 100 meters a day, but a smaller fraction the population travel much farther.
In contrast, Thorbeck and Topping (2005) modelled spider distribution in heterogeneous landscapes including a near realistic movement process based on frequent short distance displacement (10 m/day) and long distance ballooning under optimal conditions (200 m/hour). Their simulations showed that spatial landscape diversity (number of habitat types available for the spiders) is crucial for the persistence of spiders, but that spatial heterogeneity (spatial arrangement of patches) only had little impact on spider abundance, while patches free of pesticide and enough prey are crucial for abundance. The simulations indicated that the combination of high dispersal abilities and high reproductive rate enables these spiders to exploit the transient resources of the different habitats in the agricultural landscape.

The significance of frequent short and long distance ballooning is spiders is also well described by Wyemann (1993, 2000) Changes in air movement are an immediate trigger for take-off behaviour, and food deprivation is confirmed as a short-term moderator affecting ballooning frequency. Ballooning motivation of spiders inhabiting arable farmland is frequent all over the seasons and covers tens to hundreds of meters. Ballooning by spiders found on arable farmland is suggested to have evolved primarily as a risk-spreading strategy to maximise survival in unpredictable habitats. Frequent redistribution and the web-building behaviour tend to make spiders over-dispersed, which has to be accounted for in statistical analysis.

Due to the frequent and complex dispersal behaviour, redistribution patterns and broad diet of spiders it is unlikely that possible small indirect effects of transgenic crops can be showed in a small scaled setting.
A specific problem for orb web spiders in maize is that during the late main cropping season (June-September), many species are in the juvenile stage which prevent species identifications to species level. Most spiders are active

at and just above the soil surface level. As they move daily, they are easily trapped by pitfall trapping. Just a few common species usually constitute over 90% of all spiders caught in traps.

Potential effects on arthropods in other crops and ecosystems.

The foregoing review of arthropod groups is put in the context of maize systems in accordance with the focus of this report. But this information is also useful for other arable crops and even other annual crops such as vegetables, because many species living in arable land tend to be generalists and are widespread. Yet, every crop has its own characteristic fauna and consequently additional and other non-target species that may be exposed and affected in another way. The basic properties of the arthropod groups described remain relevant for most annual crop ecosystems. Perennial crops such as vineyards, orchards, and tree nurseries, however, harbour a quite different fauna and other factors may dominate the population dynamics in such systems. The all-year round presence of crops promotes stability and a more resident fauna. Under such conditions some species may be more exposed to substances that are present in the crop or the food webs.

Several studies have already pointed to the vulnerability of arthropods to Bt-toxins in other systems such as forests and aquatic systems (Carstens *et al* 2012) where arthropod communities tend to be different and more divers in species and exposure routes can be more complex.

8. Suggestions for better field trials and alternative approaches

Based on scientific evidence from lower tier studies or ecological inference, there are sometimes reasons to assume that potential effects of current or future insect transgenic crops may occur on non-target insects, spiders or other organisms.

In such cases, field studies to investigate such effects can be useful, but only if a number of conditions is fulfilled. It is necessary that such studies should have a clear focus, for example by studying only one or a few species that are selected by strict criteria (Todd 2008) and for which the ecological context is sufficiently clear. It is also needed that clearly defined effects can be reliably assessed in order to get results than can equivocally be interpreted. When any doubts exist about confounding factors that can mask effects (by strong movements or high temporal dynamics for example) it might be better to look for alternative research approaches.

Assuming that species to be studied have low to moderate mobility (see below) the following issues should be considered for experiments to be useful:

- A precise effect definition should be formulated, a relevant effect size and a species-specific sampling plan for each species should be available. Such a plan includes the selection of the best sampling method to estimate (changes in) absolute densities and the intensity to collect sufficient numbers and cope with (known or unknown) variance
- It is often helpful to consider how imaginary results will be analysed and interpreted before starting the experiments. For example how the interpretation of results would be with different outcomes.
- Plot size should be defined in relation to mobility information to avoid excessive interchange between plots. Modelling exercises may help to estimate sufficiently large plot-sizes based on effect-size to be detected, mobility, the duration of experiments and expected variance. In general it seems necessary for the more mobile species (adult stages of flying species) to use much bigger plots than are currently used.
- Variability (due to sampling variance and natural small scaled heterogeneity, difference between site and years) in most species requires that 5 or even better 10 replicates for each treatment should be used (personal communication, van de Voet), preferably at different sites and in 3 or more years (making it more easy to generalize results).

  Ex-post analysis has confirmed that 5 replicates are still not sufficient to reach acceptable statistical power for the common species studied (Prasifka *et al* 2008) and species with lower densities (in the order of < 10 per sample occasion) (Comas *et al.* 2012). So any a priori knowledge about variance can be helpful to improve experimental power. Albajes *et al.* (2013) showed that abundance is the most influential factor determining the capacity to detect differences. But also the residual variance increases with abundance (Rauschen 2010)
- Intensive and appropriate (species-specific) sampling procedures and methods make abundance estimates more reliable and tend to reduce variance and to increase statistical power.
- When one is interested in direct effects rather than indirect effects, the quantity and quality of food sources (such as prey) should be monitored as well and the uptake of Cry-toxins by the non-target should be assessed as well to provide a direct link of the toxin and any effect on numbers (see e.g. Stephens 2012).
- Including a positive control (e.g. a pesticide that has proven toxicity in the field for the species to be studied) to give an indication of the effect size that can be detected in the field experiment.
- Direct estimation of fitness parameters such as larval mortality and fecundity in the field trials can give additional or better insight on what is really happening in the field.
- Analysing differential change in numbers (population growth or decline) maybe more informative that just comparing numbers between plots. It should also be assured that the natural population development is not influenced by other cropping factors than that of the transgenic effect and that change can be ascribed to reproduction or mortality rather than emigration or immigration.

- Using release-recapture procedures might be useful for species that can be marked and easily re-trapped where differences in recapture rates may indicate 'adverse effects' of unfavourable conditions resulting in mortality or emigration from the plot (Frederiksen 2013)

To avoid any experimental problems with the movement of organisms the only real option is to choose less mobile species (or life stages) or to create physical barriers between plots (which is rarely possible). For some non-flying ground and rove species one could use barriers (e.g. closed fences) between plots. Several of those species disperse by walking rather than by flying and as they also eat all kinds of prey the uptake of Cry-toxins can be indicative for the toxin spread in the maize food web. Detailed studies for one or two of the most critical and less mobile species (such lady bird larvae, or exposed caterpillars) may give more leads to the potential impact of transgenic crops than studying the whole community.

It is obvious that larval stages of for example ladybirds, lacewings, syrphid flies and predatory bugs are less mobile NTOs for which the spread of toxins in the food chain is likely. Closely monitoring change in larval populations can be a potential way to detect transgenic effects when sampling or low numbers are not hindering accurate measures. Also soil mesofauna species such as spring tails and soil mites can be considered as more or less stationary. However, such effects may be more easily studied in more controlled mesocosm studies in the laboratory.
As the uptake of Cry toxins is likely to occur in species that either feed on plant tissue or which consume or are exposed to pollen, such species are most interesting for non-target studies in particular the larval stages. These. may also include crop feeding caterpillars, beetle larvae as long as they are considered as harmless and non-target Beneficial quiet stationary organisms such as Orius and predatory mites that feed on pollen and small insects are closely connected to the maize plants are potentially good study objects.

Also stationary larvae of non-target species living on weeds inside the crop or on wild plants in the field margins may be good study objects in particular when it concerns lepidopteran or chrysomelid species. This field has hardly been explored so far. Complexity in weed and field margin ecosystems, however, may be an extra methodological obstacle for experimental studies.

Measuring Cry-content in NTOs during any experiment is helpful for a better interpretation of results in particular when it concerns cause and effects. For example a reduction of numbers in plant feeding species can be due to toxin effects or due to increased predation levels. When high toxin uptake is measured the first interpretation becomes more likely.

For risk managers and other stakeholders it may be helpful to translate temporal small scale effects that are found in field trials into the real world large scale context of the agricultural environment.

To put any experimental results into a context of a larger space and time scale than a single field, is necessary to judge ecological impact for example using a modelling approach. More important processes and (agronomic) factors determining local abundances may be at hand that may limit the ecological significance of detected field effects. Many species are able to recover from disturbances (Macfadyen *et al* 2014) when these are transient (e.g. transgenic maize in a rotation with other crops).

Can fields trials be used to studypotential adverse effects of NTOs? Yes, when it concerns species with low mobility, when effects are very clear and not very sensitive to plot interference, that could be possible. In particular when lower tier studies indicate significant effects on mortality parameters it can be useful to evaluate such effects in field trial. In that case a full focus on the particular species and a sound experimental set-up can provide additional information. But it might be less useful to trace spurious effect that are not triggered by lower tier studies? The very costly field trials may not be the most straightforward way.

Looking for any unknown non-indicated effects might be more useful in post market release situations. Any trends detected of course can lead to new hypothesis, insights and experiments. But is looking for the unknown and non-science based hypothetical questions something that can be used for regulation? The number of potential effects of transgenic and non-transgenic crops is infinite. It should be realized that also in post market environmental

monitoring similar methodological problems may be present, but at least the realistic embedding at farm scale and covering trends over more years may provide harder evidence for the presence or absence of effects when densities of species are estimated with sufficient accuracy (see e.g. Albajes *et al.* 2012)
Impact studies that focus on ecosystem functions rather than effects on single species were not discussed in the report, but some ecological tools used in those studies could be used to improve research on NTOs. For example by measuring predation rates on artificial prey (egg-masses for example) or parasitation rates (percentages of infested hosts) may give a better indication of the size population of beneficial species in the crop than by merely counting trapped numbers. For an example see Orr and Landis (1997).

Suggestions for contemporary field trials for some NTO groups.
As long as the consequences of plot interference due to mobility and abundance assessments errors due to sampling remain unclear, it is hard to formulate new advices for concrete experiments that give more valuable results. Yet some preliminary recommendations are given below for ongoing research on commonly studied NTOs in field trials. For all groups it is suggested to take mobility aspects into consideration and not to rely only on the guideline for field trials in literature. Below some suggestions are given for specific NTO groups that should be considered as **additional** to advices given elsewhere in literature (e.g. EFSA 2010).

Lepidoptera. Species of this group are potentially sensitive for the Lepidoptera targeted Cry toxins. Considering the high mobility, abundance assessments for adult butterflies or adult moths in field trials are considered to be meaningless when looking for GM-crop effects. Due to low mobility and better sampling methods trials with caterpillars might be more informative. The survival of caterpillars of non-target Lepidoptera living on weeds within the crop or in field boundaries could be estimated by comparing numbers before and some weeks after pollen shed as the main source of exposure (see Gathmann 2006). As caterpillars are not likely to move further than 1 or 2 meters during development, plots can be small (10*10 m) but ranges of pollen shed should not overlap between plots. Buffer zones that avoid wind driven pollen contamination should be taken into account.

Chrysomelidae. Species of this group are potentially sensitive for Chrysomelidae targeted Cry-toxins. Most species of interest for non-target studies in and around maize ecosystems are likely to be found on weeds and other host plants in the field margins. For the species studied mobility is likely to be high or at least moderate. By counting adults it can be assumed that mortality cannot be distinguished from emigration in plots smaller than 0.5 ha. With larger plots direct counts of adults on plants could be an option, but sticky-traps or water traps may just indicate what flies 'over the crop' and are not very useful. Following larvae of particular weed-associated species before and after pollen shed may give indication for mortality effects when studied in plots of at least 20*2O m with sufficient isolation to avoid pollen interference between plots.

Carabidae. The carnivorous common species that disperse mainly by walking such as Pterostichus melanarius or Pterostichus melanarius may be suitable for study but experimental plots should be large (> 1 ha) and sufficient replicates (>4) are needed to cope with density variations which are easily caused by microclimate and soil differences. When food and crop cover conditions among plots are not different, pitfall traps (at least 5/plot) can fairly reflect absolute densities when applied over the full season. Any density differences among plots should preferably be correlated by measured uptake of toxin to show causal effects.

Staphylinidae. It is doubtful whether any species is resident and numerous enough to be studied in field trials. For the moment it is assumed that this group is less suitable to study in field trials.

Coccinellidae. Some ladybird species are common and important predators for aphids. The adults themselves are not suitable for NTO field trials due to their high and frequent mobility. Egg laying is likely to be independent of crop variety unless aphid prey is at equal levels. When this can be assured it might be useful to measure mortality of the lady bird larvae in NTO field trial as these are resident and countable on a plant by plant base. When intensive sampling in a time series is applied in plots of > 0.25 ha and prey populations are monitored as well this could be a useful approach to study effects on predators of aphids that feed on GM crops. Measuring Cry-uptake by ladybird larvae during the trials well support the interpretation of any potential differences to be found.

Lyniphiidae Some common species (such as Erigone atra) occur all over Europe. The adults spiders tend to peak in June, and the juveniles later in the season are hard to identify to species level. They have a broad diet of small insects of which only a few are likely to contain Cry toxins. As they can easily disperse over moderate and large distances, their suitability for field studies is doubtful, but large plots (1 ha or preferably more) may overcome this problem. Sampling by pitfall trapping is probably the only feasible method, but adverse conditions to increased activity may lead to higher trap catches without density changes. Therefore the results of this group will always be difficult to be interpreted.

Orius As a common predator of aphids and leafhoppers in maize ecosystems, this species is considered to be a good target for effect studies. However, one should realize that this species responds very quickly to food conditions by redistribution over large distances. That is why counting or trapping adults make little sense. Following demographics of larvae in association to prey populations could make more sense but is very difficult to perform due to the larvae's small size and hidden life style.

9. Discussion and conclusions

Field studies to address potential adverse effects of transgenic insect-resistant crops on non-target insects are required in the EU before these crops are released for commercial production. The central question for this report was whether such field trials, as they are currently done, are useful to study potential effects on NTOs and if not, under which conditions they could be useful if the methodology would be improved.
The underlying concern was that mobility of many insect species makes it hard to do field experiments that are powerful and robust enough to detect ecological and statistical significant differences in the abundance of NTOs transgenic crops compared to the abundance in non-transgenic crops. There are several reasons to assume that this concern is justified because many of the studied insects and spiders are quite mobile and scale of experiments tend to be smaller than probably needed.

The apparent lack of effects, however, might be also due to highly variable data and inappropriate experimental designs or poor sampling. An additional problem is that experimental designs and results that show to be robust in one year may be fully unreliable next year as was show in study of Naranjo in Bt-cotton (2005). Changes in populations densities (often due to weather) or movement patterns can be unpredictable and hence – even when mobility is limited- it might be best to include experimental safety buffers by maximizing plot size, number of replicates and sampling intensity.

Many initiatives are taken to improve the methodology in the experimental design and better selection criteria for non-target species (e.g. AMIGA project and the EFSA 2010) However, one may wonder if by collecting more data and by applying better statistics, problems are solved as long as the ecological properties of the species are not better taken into account. This is particularly true for species of which numbers fluctuate strongly within the season and between years and those that respond quickly to disturbing factors by small and large-scale migrations from field to field. These phenomena are characteristic for many species inhabiting agro-ecosystems.

The current dominant practise of measuring 'abundance' of multiple species by a variety of methods with the aim to detect some species that respond while taking the spatial-temporal dynamics for granted, is likely to result in more and more variable data that are hard to interpret, but not leading to robust conclusions.

By critical analysis of current practices and methodologies in recent field trials it becomes clear that the quality of field trials that are carried out often is low due to insufficient focus in research questions, and ineffective allocation of money and time to gather the useful data and set up rigorous field trails that can answer all the questions. Instead of collecting precise data for specific questions, experiments are set up to collect many data to increase the likelihood that some significant differences or in-equivalences can be shown.

For many reasons it should be better to make specific choices about which changes in abundance would be relevant for a particular species, based on ecological reasoning and sufficient background knowledge including their population dynamic behaviour in time and space, movement aspects and precise methodology to assess abundance (see also Rauschen 2010 and Peterson 2013).

From recent field studies it appears that a limited group of species of ground beetles, rove beetles, ladybird beetles, pirate bugs, hover flies, spiders and a few others are numerous enough in samples to be statistically analysed. Unfortunately poor statistical power is often found in these studies for the majority of the species, prevents to detect effects if they would be present. But even worse is to realize that most of the species studied are ecologically very dynamic and mobile which is likely to blur potential differences between plots. Moreover, for many of those species, no direct effects of Cry toxins have been shown in lower tier studies. So one may wonder what should be expected. For all the other less numerous species in the experiments measured abundance tends to be too low to be statistically analysed (often due to poor sampling?). So, if research is looking for unexpected effects among these less common species by exploring poor data sets this does not look very promising unless there is more focus on less mobile species and sampling is done much more intensively.

Despite all these limitations in current field studies, one may argue that similar experiments are done in pesticide studies and effects of pesticides on non-target insects are frequently shown, also with insecticide treatments within the transgenic crop trials (a good example is given by Naranjo 2005 for cotton). Therefore, many authors tend to conclude that the effects of transgenic crops are minor compared to those of some broad-spectrum pesticides. Another way to formulate this conclusion is that current field trials can only indicate adverse (temporal within season and within field) effect when the effect is strong enough to overrule blurring factors, just as in experiments of some broad spectrum insecticides. Actually the problem of mobility in pesticide effect fields trials and consequent plot interference was recently reviewed by Macfadyen *et al* (2014). They conclude that plot size should generally be increased to make assessments more robust (plot size over 1 hectare for highly mobile species) They also advise to ensure that the scope and aim of the study is clear and the plot size is adapted to the species studied (based on mobility information).

The long term and large scale effects of multiple pesticides and transgenic effects (including future applications) on non-target insects in current agro-ecosystem and agro-landscape, however, is still not fully clear. At least many other factors in agriculture are relevant for maintaining functional agro-biodiversity and species of conservation value. The understanding about cropping practices including the effects transgenic crops on non-target species may better be based on farming system research, agro-landscape and modelling studies than on new field trials for current insect-resistant Bt crops that are weak in detecting potential but probably limited effects (see also Sweet 2011, Ervin 2011, Tscharntke 2012 and many others). The problem of translating small scale effects to large scale impact on populations at higher spatial scales can only be handled by modelling approaches as was e.g. done for pesticides already in the nineties (Sherrat *et al* 1993). For many species the spatial-temporal distribution of crop management factors at farm and landscape scale is extremely important to maintain populations of beneficial or nature value (Holland *et al* 2005).

For the species that are very likely to be sensitive for the currently used and new Bt-Cry toxins (such as Lepidoptera and Chrysomelidae) it may be advisable to do specific studies on species that live on plants that are associated with maize fields, such as on weeds in low tillage systems or plants that grow in field boundaries. For an elegant experimental approach and balanced discussion, see Gathmann (2006). The experimental constraints discussed and improvement suggestion given in this report may be used to design more appropriate experimental set-ups for new transgenic crop settings taking into account the mobility aspects.

Finally one may wonder why for impact studies of farming practices on NTOs (including the use of transgenic crops) no absolute quantitative targets are being formulated that can be considered as sustainable. The only thing we know that our target species (pests) and our non-target species are mutually dependent and all populations move in the patchy landscape to find food and avoid adverse conditions. So maybe the question should be to what extent transgenic crops provide enough NTOs for the surrounding fields and follow-up crops rather than only focusing on whether some species are less frequently caught on Bt-maize fields

The finding that many current field trials have a limited value because of methodological and feasibility problems does not mean that field studies will never be useful. Improvements are possible and combination of knowledge from laboratory studies, field studies and ecological modelling may provide new insights as was argued in the related field of ecotoxicology (Artigas *et al* 2012, Beketov 2012). Meanwhile the discussion should be intensified about when any 'effects' should be regarded as adverse and harmful in order to make any research results helpful for decision making (Sanvido *et al.* 2012) .

From this report it can be concluded that field trials to study effects of insect-resistant transgenic crops are likely to be useful for species that are not very mobile, that plot size should always be adapted to the verified mobility patterns of species to be studied and that other methods such as modelling could be helpful to predict potential effects at a larger scale as long as model parameters have a sufficient scientific base.

Glossary

Abundance – mostly used as the number of individuals measured or observed by a certain method at a certain place at a certain moment (or accumulated over a period of time) often used as an indicator of population-size.

Agroecosystem – An ecosystem that is primarily aimed at producing one or more crops often consisting of different fields within a farming system that is managed by a farmer or company. Also semi-natural patches within the farming system that interact with crops are considered as part of the agro-ecosystem.

Ballooning – The behaviour of spiders to produce free silk threats by which they can be transported by wind over considerable distances.

Beneficial organisms – In the context of agriculture beneficial organisms are all species that promote plant growth or prevent any crop losses due to pests and diseases. Typical examples are growth promoting soil bacteria, predators and parasitoids that eat pest insects, pollinators such as bees. In general they contribute to the productivity of crops or quality of the products. As such they have a positive function in the agro-ecosystem. Many species are considered as beneficial even though their functionality is often not quantified and can be variable.

Block – A block within an experiment consists of two or more plots having different treatments and defined in contrast to other blocks within the experiment usually because blocks are spatially separate and homogeneity within blocks is assumed to be lower than between blocks.

Carabid beetles - Ground beetles that are soil surface dwelling beetles of the family Carabidae including many species that feed as a predator on other insects and are considered to be beneficial.

Dispersal capacity– the potential of organisms to spread by walking or flying within and between fields.

Effects -In a broad experimental sense, an effect is any change in a measured parameter (for example the numbers of spiders caught in traps) that can with some certainty be related to a factor that is experimentally changed. It is not always clear such an effect is exactly caused by the experimental factor. Effects can have different sizes (small or strong), can have different scales (plot, field, farm and so on), and can be short term (e.g. weeks) or long term (e.g. years).

NTO -non-target organism, meaning any organism apart from the species that was the original target for which the toxin producing gen construct was incorporated in the transgenic crop. NTOs can include non-target pests, beneficial organisms and species of conservation value.

Orb-web spiders – small spiders that mostly belong to the family Lyniphidae and are often abundant in agro-ecosystems. They build tiny webs in which many small insects are trapped and eaten by the spiders. Therefore these spiders are considered to be beneficial and often studied.

Plot – the smallest uniform unit in an experiment where measurements are done, and characterized by a treatment that is applied for that plot. Plots may be situated within an experimental block. (see **block**)

Population – All individuals of the same species. Populations often consist of many sub populations that are linked by migration and movement (dispersal). Sub-populations are often defined for a certain defined space such as a field, but strong connectedness to other subpopulation limits the value of such definitions

Population density – The number of individuals of one species per surface or volume (e.g. per m^2 or m^3) at a particular moment. Local densities naturally change in time due to reproduction, development, mortality, immigration and emigration. Densities accordingly vary in time and space to environmental (abiotic and biotic) factors and are not necessarily synchronized between locations and subpopulations.

Power analysis -Statistical data analysis to determine the chance that an effect of some factor can be detected based on the variability found in the data of the variable to be studied and the experimental set-up.

Resilience – The ability of an ecological system or species community to recover from disturbances, also used as the biological buffering capacity to prevent development of insect outbreaks.

Significance – Often used as ‘statistical significant effect’ meaning that it is unlikely that an observed effect is caused by chance. As a more common term, ‘significant’ also indicates that is meaningful i.e. that the effect found has some importance and has relevant impact in the context. So a small but statistical difference of a factor may be irrelevant in the context of other factors in an ecological, economic of safety context. In this report significant is only used in the statistical meaning, using relevance for context dependent impact.

References

Ahmad, A., G. E. Wilde, *et al.* (2005).
'Detectability of coleopteran-specific Cry3Bb1 protein in soil and its effect on non-target surface and below-ground arthropods.' Environmental Entomology 34(2): 385-394.

Ahmad, A., G. E. Wilde, *et al.* (2006).
'Effect of corn hybrids expressing the coleopteran-specific Cry3Bb1 protein for corn rootworm control on aboveground insect predators.' Journal of Economic Entomology 99(4): 1085-1095.

Albajes, R., G. P. Farinos, *et al.* (2012).
'Post-market environmental monitoring of Bt maize in Spain: Non-target effects of varieties derived from the event MON810 on predatory fauna.' Spanish Journal of Agricultural Research 10(4): 977-985.

Albajes, R. *et al.* (2013)
Representative taxa in field trials for environmental risk assessment of genetically modified maize. Bulletin of Entomological Research 103:724-733

Anonymous (2011).
'AMIGA : Assessing and monitoring the impacts of genetically modified plants on agro-ecosystems.' Seventh Framework Programme proposal pp 98.

Arpaia, S. (2010).
'Genetically Modified Plants and 'Non-Target' Organisms:
Analysing the Functioning of the Agro-ecosystem.' Collection of Biosafety Reviews 5: 13-80.

Artigas J., G.Arts, *et al.* (2012)
Towards a renewed research agenda in ecotoxicology. Environmental Pollution 160:201-206.

Bai, Y. Y., R. H. Yan, *et al.* (2012).
'Field response of aboveground non-target arthropod community to transgenic Bt-Cry1Ab rice plant residues in postharvest seasons.' Transgenic Research 21(5): 1023-1032.

Bakker N.V.J. en C.J.M. Musters (2006)
Non-target arthropods im conventional potato and maize crops.
Istitute of Enviromnmental Sciences. Leiden University, CLM Report 170, 47 pp.

Balog, A., J. Kiss, *et al.* (2010).
'Rove beetle (Coleoptera: Staphylinidae) communities in transgenic Bt (MON810) and near isogenic maize.' Crop Protection 29(6): 567-571.

Balog, A., A. Szénási, *et al.* (2011).
'Analysis of soil dwelling rove beetles (coleoptera: Staphylinidae) in cultivated maize fields containing the bt toxins, cry34/35ab1 and cry1f×cry34/35ab1.' Biocontrol Science and Technology 21(3): 293-297.

Banks, J. E. and B. Ekbom (1999).
'Modelling herbivore movement and colonization: Pest management potential of intercropping and trap cropping.' Agricultural and Forest Entomology 1(3): 165-170.

Beketov, M. A. and M. Liess (2012).
'Ecotoxicology and macroecology - Time for integration.' Environmental Pollution 162: 247-254.

Bhatti, M. A., J. Duan, *et al.* (2005).
'Field evaluation of the impact of corn rootworm (Coleoptera: Chrysomelidae)-protected Bt corn on ground-dwelling invertebrates.' Environmental Entomology 34(5): 1325-1335.

Bhatti, M. A., J. Duan, *et al.* (2005).
'Field evaluation of the impact of corn rootworm (Coleoptera: Chrysomelidae)-protected Bf corn on foliage-dwelling arthropods.' Environmental Entomology 34(5): 1336-1345.

Bruck D.J. , de Lopez M., *et al* (2006)
'Effects of transgenic bacillus thuringiensis corn and permethrin on nontarget arthropods.' Journal of Agricultural and Urban Entomology 23: 111-124

Bryan W., Clark, T. A. P., and J.R. Coats (2005).
'Environmental Fate and Effects of Bacillus thuringiensis (Bt) Proteins from Transgenic Crops: a Review.' J. Agric. Food. Chem 53: 4643-4653.

Candolfi, M. P., K. Brown, *et al.* (2004).
'A faunistic approach to assess potential side-effects of genetically modified Bt-corn on non-target arthropods under field conditions.' Biocontrol Science and Technology 14(2): 129-170.
Carrasco, L. R., T. D. Harwood, *et al.* (2010).
'Dispersal kernels of the invasive alien western corn rootworm and the effectiveness of buffer zones in eradication programmes in Europe.' Annals of Applied Biology 156(1): 63-77.
Carrière, Y., C. Ellers-Kirk, *et al.* (2009).
'Landscape effects of transgenic cotton on non-target ants and beetles.' Basic and Applied Ecology 10(7): 597-606.
Carstens, K., J. Anderson, *et al.* (2012).
'Genetically modified crops and aquatic ecosystems: considerations for environmental risk assessment and non-target organism testing.' Transgenic Research 21(4): 813-842.
Chapman, J. W., D. R. Reynolds, *et al.* (2003).
'Vertical-looking radar: A new tool for monitoring high-altitude insect migration.' BioScience 53(5): 503-511.
Comas, J., B. Lumbierres, *et al.* (2013).
'Ex-Ante Determination of the Capacity of Field Tests to Detect Effects of Genetically Modified Corn on Nontarget Arthropods.' Journal of Economic Entomology 106(4): 1659-1668.
D, S. J. a. B. (2012).
'Studies to evaluate research on biological issues on GM plants of relevance to Swiss envrionments.'
Daly, T. and G. D. Buntin (2005).
'Effect of Bacillus thuringiensis transgenic corn for lepidopteran control on nontarget arthropods.' Environmental Entomology 34(5): 1292-1301.
Dávalos, A. and B. Blossey (2011).
'Matrix habitat and plant damage influence colonization of purple loosestrife patches by specialist leaf-beetles.' Environmental Entomology 40(5): 1074-1080.
Devos, Y., S. de Schijver *et al.* (2012). Bt-maize event MON 88017 expressing CryBb1 does not cause harm to non-target organisms. Transgenic Research 21:1991-1214
Ekbom B., M. Irwin and Y. Robert eds. (2000) Interchanges of insects between agricultural and surrounding landscapes. Kluwer Academica Puiblishers.. 239 pp.
Englund G. and S.D. Cooper 2005
Scale effects and extrapolation in ecological experiments. Advances in Ecological Research 33: 161-213
De la Poza, M., X. Pons, *et al.* (2005).
'Impact of farm-scale Bt maize on abundance of predatory arthropods in Spain.' Crop Protection 24(7): 677-684.
De Roos, A. M., E. McCauley, *et al.* (1991).
'Mobility versus density-limited predator-prey dynamics on different spatial scales.' Proceedings of the Royal Society B: Biological Sciences 246(1316): 117-122.
Devos, Y., A. de Schrijver, *et al.* (2012).
'Bt-maize event MON 88017 expressing Cry3Bb1 does not cause harm to non-target organisms.' Transgenic Research 21(6): 1191-1214.
Dicke, C. D. S. M. (2009).
'Designing experimental protocols to investigate the impact of GM crops on non-target organisms.
Duan, J. J., M. Marvier, *et al.* (2008).
'A meta-analysis of effects of Bt crops on honey bees (Hymenoptera: Apidae).' Plos One 3(1).
EFSA Panel on genetically Modified organisms (GMO) (2010).
Scientific Opinion on the assessment of potential impacts of genetically modified plants on non-target organisms. EFSA Journal 8 (11):1877 [72 pp.] available online: www.efsa.europa.eu/efsajournal.htm
EFSA (2010).
'Guidance on the environmental risk assessment of genetically modified plants.' EFSA Journal 8: 111 pp.
EFSA & COGEM International Workshop
'Non-Target Organisms and Genetically Modified crops': Assessing the effects of Bt-proteins' 30 pp available online at www.cogem.net

Englund, G. and S. D. Cooper (2003).
Scale effects and extrapolation in ecological experiments. Advances in Ecological Research. 33: 161-213.

Ervin, D. E., L. L. Glenna, *et al.* (2011).
'The theory and practice of genetically engineered crops and agricultural sustainability.' Sustainability 3(6): 847-874.

Farinós, G. P., M. de la Poza, *et al.* (2008).
'Diversity and seasonal phenology of aboveground arthropods in conventional and transgenic maize crops in Central Spain.' Biological Control 44(3): 362-371.

Follett, P. A., W. W. Cantelo, *et al.* (1996).
'Local Dispersal of Overwintered Colorado Potato Beetle (Chrysomelidae: Coleoptera) Determined by Mark and Recapture.' Environmental Entomology 25(6): 1304-1311.

Folloni, S., D.-M. Kagkli, *et al.* (2012).
'Detection of airborne genetically modified maize pollen by real-time PCR.' Molecular Ecology Resources 12(5): 810-821.

Frederiksen, M., J. D. Lebreton, *et al.* (2013).
'Identifying links between vital rates and environment: A toolbox for the applied ecologist.' Journal of Applied Ecology.

Galic, N., U. Hommen, *et al.* (2010).
'Potential application of population models in the european ecological risk assessment of chemicals II: Review of models and their potential to address environmental protection aims.' Integrated Environmental Assessment and Management 6(3): 338-360.

Galic, N., A. Schmolke, *et al.* (2012).
'The role of ecological models in linking ecological risk assessment to ecosystem services in agroecosystems.' Science of the Total Environment 415: 93-100.

Gathmann, A., L. Wirooks, *et al.* (2006).
'Impact of Bt maize pollen (MON810) on lepidopteran larvae living on accompanying weeds.' Molecular Ecology 15(9): 2677-2685.

Habustova O., Dolezal, P. *et al.* (2012)
Impact of Cry1Ab toxin expression on th enon-target insects dwelling on maize plants. Journal of Applied Entomology 2012: 1-5

Harwood, J. D., W. G. Wallin, *et al.* (2005).
'Uptake of Bt endotoxins by nontarget herbivores and higher order arthropod predators: molecular evidence from a transgenic corn agroecosystem.' Molecular Ecology 14(9): 2815-2823.

Holland, J. M., C. F. G. Thomas, *et al.* (2005).
'Farm-scale spatiotemporal dynamics of predatory beetles in arable crops.' Journal of Applied Ecology 42(6): 1140-1152.

Holst, N., A. Lang, *et al.* (2013).
'Increased mortality is predicted of Inachis io larvae caused by Bt-maize pollen in European farmland.' Ecological Modelling 250: 126-133.

Jeffries, D. L., J. Chapman, *et al.* (2013).
'Characteristics and drivers of high-altitude ladybird flight: Insights from vertical-looking entomological radar.' Plos One 8(12).

Jensen, P. D., G. P. Dively, *et al.* (2010).
'Exposure and nontarget effects of transgenic bt corn debris in streams.' Environmental Entomology 39(2): 707-714.

Julie A. Peterson, J. G. L., and James D. Harwood (2011).
'Interactions of transgenic Bacillus thuringiensis insecticidal crops with spiders (Araneae).' Journal of Arachnology 39(1): 1-21.

Leslie, T. W., D. J. Biddinger, *et al.* (2009).
'Carabidae population dynamics and temporal partitioning: Response to coupled neonicotinoid-transgenic technologies in maize.' Environmental Entomology 38(3): 935-943.

Lovei, G. L., D. A. Andow, *et al*. (2009).
'Transgenic Insecticidal Crops and Natural Enemies: A Detailed Review of Laboratory Studies.' Environmental Entomology 38(2): 293-306.

Lovei, G. L. and S. Arpaia (2005).
'The impact of transgenic plants on natural enemies: a critical review of laboratory studies.' Entomologia Experimentalis Et Applicata 114(1): 1-14.

Ludy, C. and A. Lang (2006).
'A 3-year field-scale monitoring of foliage-dwelling spiders (Araneae) in transgenic Bt maize fields and adjacent field margins.' Biological Control 38(3): 314-324.

M., S. E. J. a. D. (2005).
'Effects of insect-resistant transgenic crops on non-target arthropods: first step in pre-market risk assessment studies.' COGEM, Bilthoven.

Macfadyen. S. *et al*. (2014).
'Using semifield studies to examine the effects of pesticides on mibile terrestrial invertebrates'. Annual Review Entomology 59:383-404

Maier, C., K. Bachinger, *et al*. (2013).
'European Stone Fruit Yellows: A Mark, Release and Recapture Experiment Tracking the Dispersal of its Vector Cacopsylla pruni (Hemiptera: Psyllidae) in a Model Apricot Orchard and Epidemiological Studies in Lower Austria.' Journal of Phytopathology 161(10): 713-722.

Marvier, M., C. McCreedy, *et al*. (2007).
'A meta-analysis of effects of Bt cotton and maize on nontarget invertebrates.' Science 316(5830): 1475-1477.

Mazzi, D. and S. Dorn (2012).
'Movement of insect pests in agricultural landscapes.' Annals of Applied Biology 160(2): 97-113.

Meissle, M. and A. Lang (2005).
'Comparing methods to evaluate the effects of Bt maize and insecticide on spider assemblages.' Agriculture, Ecosystems and Environment 107(4): 359-370.

Meissle M, Álvarez-Alfageme F, Malone LA & Romeis J (2012).
'Establishing a database of bio-ecological information on non-target arthropod species to support the environmental risk assessment of genetically modified crops in the EU.' Supporting Publications 2012:EN-334, EFSA, [170 pp], http://www.efsa.europa.eu/en/supporting/doc/334e.pdf

Mesa, L. A., B. G. Howlett, *et al*. (2013).
'Changes in the relative abundance and movement of insect pollinators during the flowering cycle of Brassica rapa crops: Implications for gene flow.' Journal of Insect Science 13.

Mo J., Baker,G, and R. Roush (2003)
Local dispersal of the diamond-back moth (Plutella xylostella L.)) (Leipdoptera:Plutellida). Environmental Entomology (32): 71-79

Musters, B. N. V. J. a. C. J. M. (2006).
'Non-target arthropods in conventional potato and maize crops.'

Naranjo, S. E. (2005).
'Long-term assessment of the effects of transgenic Bt cotton on the abundance of nontarget arthropod natural enemies.' Environmental Entomology 34(5): 1193-1210.

O'Rourke, M. E. and L. E. Jones (2011).
'Analysis of landscape-scale insect pest dynamics and pesticide use: An empirical and modeling study.' Ecological Applications 21(8): 3199-3210.

Orr, D. B. and D. A. Landis (1997).
'Oviposition of European Corn Borer (Lepidoptera: Pyralidae) and Impact of Natural Enemy Populations in Transgenic Versus Isogenic Corn.' Journal of Economic Entomology 90(4): 905-909.

Perry, J.N., C.J. ter Braak et. al. (2009).
(Statistical aspects of environmental risk assessment of GM plants for effects on non-target organisms. Environmental Biosafety Research 8 :65-78

Perry, J. N., Y. Devos, *et al.* (2010).
'A mathematical model of exposure of nontarget Lepidoptera to Bt-maize pollen expressing Cry1Ab within Europe.' Proceedings of the Royal Society B: Biological Sciences 277(1686): 1417-1425

Perry, J. N., Y. Devos, *et al.* (2012).
'Estimating the effects of Cry1F Bt-maize pollen on non-target Lepidoptera using a mathematical model of exposure.' Journal of Applied Ecology 49(1): 29-37.

Perry, J. N., S. Arpaia, *et al.* (2013).
'No evidence requiring change in the risk assessment of Inachis io larvae.' Ecological Modelling 268: 103-122.

Peter A. Hambäck, M. V., Teja Tscharntke, Carsten Thies and Göran Englund (2007).
'Top-down and bottom-up effects on the spatiotemporal dynamics of cereal aphids: testing scaling theory for local density.' Oikos 116: 1995-2006.

Peterson, J. A., J. J. Obrycki, *et al.* (2009).
'Quantification of Bt-endotoxin exposure pathways in carabid food webs across multiple transgenic events.' Biocontrol Science and Technology 19(5-6): 613-625.

Piechnik, D. A. (2013).
'Trophic levels colonize sequentially but effects of habitat size and quality are transient.' Acta Oecologica 47(0): 85-94.

Prasifka, J. R., K. M. Heinz, *et al.* (2005).
'Relationships of landscape, prey and agronomic variables to the abundance of generalist predators in cotton (Gossypium hirsutum) fields.' Landscape Ecology 19(7): 709-717.

Prasifka, J. R., R. L. Hellmich, *et al.* (2005).
'Assessing the Effects of Pest Management on Nontarget Arthropods: The Influence of Plot Size and Isolation.' Environmental Entomology 34(5): 1181-1192.

Prasifka, J. R., R. L. Hellmich, *et al.* (2008).
'Selection of nontarget arthropod taxa for field research on transgenic insecticidal crops: Using empirical data and statistical power.' Environmental Entomology 37(1): 1-10.

Prasifka, J. R., P. C. Krauter, *et al.* (1999).
'Predator conservation in cotton: Using grain sorghum as a source for insect predators.' Biological Control 16(2): 223-229.

Priesnitz, K. U., U. Benker, *et al.* (2013).
'Assessment of the potential impact of a Bt maize hybrid expressing Cry3Bb1 on ground beetles (Carabidae).' Journal of Plant Diseases and Protection 120(3): 131-140.

Project, A. (2011).
'Assessing and Monitoring the Impacts of Geneticallly Modified Plants on Agro-ecosystems. Collaborative Project in the 7th Framework Programme: GM crops in the EU.

Rader, R., W. Edwards, *et al.* (2011).
'Pollen transport differs among bees and flies in a human-modified landscape.' Diversity and Distributions 17(3): 519-529.

Rauschen, S., J. Eckert, *et al.* (2008).
'An evaluation of methods for assessing the impacts of Bt-maize MON810 cultivation and pyrethroid insecticide use on Auchenorrhyncha (planthoppers and leafhoppers).' Agricultural and Forest Entomology 10(4): 331-339.

Rauschen, S., F. Schaarschmidt, *et al.* (2010).
'Occurrence and field densities of Coleoptera in the maize herb layer: Implications for Environmental Risk Assessment of genetically modified Bt-maize.' Transgenic Research 19(5): 727-744.

Raymond L, M. Plantegenest *et al* (2013)
Migrationa and dispersal may drive to high genetic variation and significant genetic mixing: the case of two agriculturally important, continental hoverflies (Episyrphus balteatus and Sphaerophoria scripta) Molecular Ecology (22) 5329-5339).

Romeis, J., D. Bartsch, *et al.* (2008).
'Assessment of risk of insect-resistant transgenic crops to nontarget arthropods.' Nature Biotechnology 26(2): 203-208.

Rose, R. and G. P. Dively (2007).
'Effects of insecticide-treated and lepidopteran-active Bt transgenic sweet corn on the abundance and diversity of arthropods.' Environmental Entomology 36(5): 1254-1268.
Sanvido, O., C. Hippe, *et al.* (2012).
'Evaluating occupational health risks in the approval process of plant protection products.' Agrarforschung Schweiz(9): 422-427.
Sanvido, O., J. Romeis, *et al.* (2012).
'Evaluating environmental risks of genetically modified crops: ecological harm criteria for regulatory decision-making.' Environmental Science & Policy 15(1): 82-91.
Saska, P., W. van der Werf, *et al.* (2013).
'Temperature effects on pitfall catches of epigeal arthropods: A model and method for bias correction.' Journal of Applied Ecology 50(1): 181-189.
Schmidt M.H. , C. Thies, *et al.* (2008)
'contrasting responses of arable spiders to the landscape at different spatial scales.' Journal of Biogeography 34:157-166
Scholte and Dicke (2005)
Effects of insect-resistant transgenic crops on non-target arthropods: first step in pre-market risk assessment studies *(COGEM REPORT 2005-06)*
Schuppener, M., J. Mühlhause, *et al.* (2012).
'Environmental risk assessment for the small tortoiseshell Aglais urticae and a stacked Bt-maize with combined resistances against Lepidoptera and Chrysomelidae in central European agrarian landscapes.' Molecular Ecology 21(18): 4646-4662.
Semenov, A. V., J. D. van Elsas, *et al.* (2013).
'The use of statistical tools in field testing of putative effects of genetically modified plants on nontarget organisms.' Ecology and Evolution 3(8): 2739-2750.
Shelton, A. M., S. E. Naranjo, *et al.* (2009).
'Setting the record straight: A rebuttal to an erroneous analysis on transgenic insecticidal crops and natural enemies.' Transgenic Research 18(3): 317-322.
Sherrat T.N. and P.C. Jepson (1993)
A metapopulation approach to modelling the long-term impact of pesticides on invertebrates. Journal of Applied Ecology 30: 696-705
Shoeb, M. A., F. A. Attalla, *et al.* (2008).
'On the migration of common predatory species from egyptian clover to adjacent maize fields in Egypt.' Egyptian Journal of Biological Pest Control 18(2): 385-389.
Sisterson, M. S., Y. Carrière, *et al.* (2007).
'Nontarget effects of transgenic insecticidal crops: Implications of source-sink population dynamics.' Environmental Entomology 36(1): 121-127.
Sivakoff, F. S., J. A. Rosenheim, *et al.* (2012).
'Relative dispersal ability of a key agricultural pest and its predators in an annual agroecosystem.' Biological Control 63(3): 296-303.
Stephens, E. J., J. E. Losey, *et al.* (2012).
'The impact of Cry3Bb Bt-maize on two guilds of beneficial beetles.' Agriculture Ecosystems & Environment 156: 72-81.
Stevens, V. M., C. Turlure, *et al.* (2010).
'A meta-analysis of dispersal in butterflies.' Biological Reviews 85(3): 625-642.
Sujii, E. R., P. H. B. Togni, *et al.* (2013).
'Field Evaluation of Bt Cotton Crop Impact on Nontarget Pests: Cotton Aphid and Boll Weevil.' Neotropical Entomology 42(1): 102-111.
Svoboda Z., Habustova, *et al.* 2012
Epigeic spiders are not affected by genetically modified maize MON 88017. Journal of Applied Entomology 137: 56-67

Sweet, J. and Bartsch 2012
Synthesis and overview studies to evaluate exiting research and knowlegde on biological issues on GM plants of relevance to Swiss enviroments. Available at: http://www.vdf.ethz.ch/service/3498/3499_synthesis-and-overview-studies_oa.pdf
Swezey, S. L., D. J. Nieto, *et al.* (2013).
'Dispersion, Distribution, and Movement of Lygus spp. (Hemiptera: Miridae) in trap-cropped organic strawberries.' Environmental Entomology 42(4): 770-778.
Thiele H.U. (1977)
Carabid Beetles in their Environments. Springer Verlag, Berlin
Thomas C.F.G., D.A. Kendall *et al.* (2006)
Carabid movement and vegetation density: implications for interpretation of pitfall trap data from split-field trials. Agriculture, Ecosystems and Environment 113:51-61.
Thorbek, P. and T. Bilde (2004).
'Reduced numbers of generalist arthropod predators after crop management.' Journal of Applied Ecology 41(3): 526-538.
Thorbek, P. and C. J. Topping (2005).
'The influence of landscape diversity and heterogeneity on spatial dynamics of agrobiont linyphiid spiders: An individual-based model.' BioControl 50(1): 1-33.
Toft S. and W. Riedel (eds.) 1992
'Arthropod natural enemies in arable land I: desnity, spatial heterogeneity and dispersal.' Acta Jutlandica LXX:2: 1-214
Todd, J. H., P. Ramankutty, *et al.* (2008).
'A screening method for prioritizing non-target invertebrates for improved biosafety testing of transgenic crops.' Environmental Biosafety Research 7(1): 35-56.
Topp, W. and W. Trittelvitz (1980).
'DISTRIBUTION AND DISPERSAL OF EPIGEIC ARTHROPODS IN AGROECOSYSTEMS .2. STAPHYLINOIDEA.' Anzeiger Fur Schadlingskunde Pflanzenschutz Umweltschutz 53(3): 33-36.
Topping, C.J. and Sunderland K.D. 1992
Limitations to the use of pitfall traps in ecological studies exemplified by a study of spiders in a fields of winter wheat. Journal of Applied Ecology 29:485-491
Tscharntke, T., J. M. Tylianakis, *et al.* (2012).
'Landscape moderation of biodiversity patterns and processes - eight hypotheses.' Biological Reviews 87(3): 661-685.
Turchin P., (1998)
Quantitative analysis of movement: measuring and modeling population redistribution in plants and animals. Sinauer Associatiot, Sunderland, MA.
Van Den Brink, P. J. and C. J. F. Ter Braak (1999).
'Principal response curves: Analysis of time-dependent multivariate responses of biological community to stress.' Environmental Toxicology and Chemistry 18(2): 138-148.
van der Voet, H., J. N. Perry, *et al.* (2011).
'A statistical assessment of differences and equivalences between genetically modified and reference plant varieties.' Bmc Biotechnology 11.
Van Der Werf, W., E. W. Evans, *et al.* (2000).
'Measuring and modelling the dispersal of Coccinella septempunctata (Coleoptera: Coccinellidae) in alfalfa fields.' European Journal of Entomology 97(4): 487-493.
Weyman, G. S., K. D. Sunderland, *et al.* (2002).
'A review of the evolution and mechanisms of ball ooning by spiders inhabiting arable farmland.' Ethology Ecology and Evolution 14(4): 307-326.
Wolfenbarger, L. L., S. E. Naranjo, *et al.* (2008).
'Bt Crop Effects on Functional Guilds of Non-Target Arthropods: A Meta-Analysis.' Plos One 3(5).

Appendix I Non target field studies

Ahmad et a; 2005 (soil surface and soil)
Comparison: CRM (Cry3Bb1) vs NTM+ (isogenic+ insecticde seed treament) vs NTM- (untreated)
Design:, 3 sites with 4 replicates each, plot size ca 16 * 16 meter (200o2) and 7*9 meter (2003), randomized complete block
Sampling: pitfall raps on 3 occasions (3 weeks during silking), 4 traps/plot, soil microarthropods by cores and tullgren, 8 cores per plot
Taxonomy: to family level (beetles, ants) and spiders(total). Carabida, Elateridae, Staphylinidae and Cry protein measurements in soil.
General conculsions: No significant differences at the family level of above ground soil dwelling arthropods; no significant differences among collembola and soil mites (even after more years). There is an effect in some groups of seed treatements (Elateridae). Very low amounts of Cry were detected in the soil supporting the no effect measured.
Comment: the non-effects of seed treatment on a number of soil surface species (carabids and spiders) may indicate interference (redistribution) between the relatively smalll plots.

Ahmad *et al* 2006 (above ground, pollen effects of Cry3Bb1)

Comparison: CRM (Cry3Bb1) vs NTM+ (isogenic + insecticide seed treatment, neonicotinoide) vs NTM- (untreated)
Design : 3 sites with 4 replicates each, plot size ca 16 * 16 meter (200o2) and 7*9 meter (2003), randomized complete block
Sampling: visual counts on 5 plants per plot (2 times in 2002, 4 times 2003) Coccinellidae and Antocoridae (Orius).; predation experiments on O.nunilalis eggs
Taxonomy: 3 Coccinellidae and Orius to species level both adults and immatures.
General conclusion: In 2002 no differences between all(!) treatments on Orius insidiosis adult and and nymphs and Hippodamia, Coliomegilla and Scymnus Coccinellidae) adults, hardly any effect detected in 2003 on all variables including ladybird larvae. No effects found between treatment for predation capacity on eggs.
Comments: the non-effects of seed treatment on all above ground groups may indicate interference (redistribution) of highly mobile species between the relatively small plots. This may also explain no predation differences.

Bhatti *et al* 2005

Comparison: transgenic LRM, non-transgenic NTM (ISO-), non-transgenic + permethrin NTM+ (ISO+)
Design: 1 site, 4 replicates, 3 years, plot size 18*18m (0.04 ha), split plot with seed treatment and foliar spray.
Sampling: 3 sticky traps / plot, 2 weeks period and repeated
Taxonomy: Chrysomelidae, Coccinellidae, Orius, aphid, leafhoppers, Chrysoperla
Statistics.: repeated measure ANOVA
General conclusions: many significant differences between year (strong year effects), strong foliar spray effects on lady birds and Orius but not consistent over years. No consistent effects on NTOs.
Comments: very small scale compared to the mostly flying insects. Adverse effects on ladybirds were found after foliar spray (probably due to prey depletion). Effects were highly variable and maximum 0.2.

Bruck *et al*. 2006

Comparison: transgenic LRM, - ,
Design: 1 site, 3 replicates, 3 treatments, plot size = 61*61 m (0.36 ha) with small 3 alley-ways between plots, 2 years

Sampling: pitfall traps (5 / plot), sticky traps (5/plot), weekly visual plant examination (10 plants), intervals from june to october
Taxonomy: Araneae, Diabrotica and leafhoppers. (only a few had a CV <75)
General conclusions: data analysis focused on differences just after insecticide applications for a best comparison and only taxa with a CV of less than 75. The use of repeated measures analysis on all data was not done because periods between applications are prone to populations recoveries.
As a strategy everything was focused to detect differences by reducing variability and avoid recovery equalizing effects. Most groups very low mean abundance per trap (<2), only spiders abundant an significant effect of insecticide, also Elateridae are affected by the insecicides, Macrocentrus cingulum (Hym. Parasitoid of O. nubilalis) was higher in non-transgenic compared to transgenic or insecticde. Nitidulidae are reduced in BT-corn probably because of their preference for damaged ears and kernels. They discuss and mention the repopulation process of many common arthropods that hinders to detect sustained effects.
Comments: despite plot size and intensive sampling numbers generally low, too much variation and only spiders gave instant response to insecticides (not transgenic).
The authors state that even insecticide effects are not easily show due to repopulation. (see also Prasifka *et al* 1999 who mention movement of about 25 m/day)). When transgenic effects are smaller they are even more difficult to show.

Candolfi *et al* 2004 (multivariate, faunistic analysis)

(according to Plant protection guidelines for fields studies of within season effects of pesticides)
Comparison: Cry1Ab BT corn vs. untransformed no insecticides vs untransformed+cyhalothrin vs untransformed+ Bt spray
Design: 1 site, 3 replicates, 4 treatments, plot-size 1.18 – 1.68 ha (= ca 135*135 meter !) 1 year
Sampling: 12 pitfall traps / plot, 8 intervals mainly july-september; 5 watertraps / plot, 10 beating-funnel per traps, 7 times in the season.
Taxonomy: soil dwelling: Linyphiidae spiders to species level, Carabidae to species level, Staphylinidae, Collembola, Formcidae and Phalangium opilio (most common harvestman). Plant dwelling: spiders (family level), Cantharidae, Elateridae and diptera, aphids (Rhopalosiphum) and cicadellids (zyginidea), Hymenoptera and Chrysopa perla, Thysanoptera.
Statistics: CANOCO principal response curves,
General conclusions: for pitfalls no response or very transient, beating samples only insecticide gave clear and significant and consistent response in plant dwelling(negative impact on some groups responsible for community response, in particular Orius had short term (4 weeks) lower numbers and also some other species only transient (recovery by emergence or re-colonization?), in flying animals (water traps) no overall community response, but the authors conclude that this can be fully due to rapid recolonization of the highly mobile taxa. Some species, however responded on some days, and the Syrphidae, Lonchopteridae and Mycetophilidae were responsible for the community response noticed, possibly a response to pollen shed is a factor for the pollen feeding flies such as Episyrphus balteatus.
Comment: Conclusions indicate that only insecticides gave short term response and both variability and mobility (recovery) are obscuring any potential effects of Bt on flying insects. The community response is mostly due to some dominant species, weakening this approach

Daly & Buntin 2005

Comparison 2 LR (CryAb1Ab) Bt-corn strains vs 1 isogenic and 1 conventional strain
Design: randomized complete block, 4 replicates, 2 sites, 2 years , plots 21*25 m (0.05 ha)
Sampling: 10 /20 plants per plot weekly visual, 2 or 3 pitfall traps / plot, weekly emptied from seeds to harvest
Statistics: ANOVA with repeated measures, with years, sites, treaments as effects.
Taxonomy: bugs, flea beetles, sap beetjles (Nitidulidae) and leafhoppers visual more or less to species, corn thrips (F. williamsi) (plant samples), and visual predators (bugs and ladybirds to species, spiders).

General conclusions: Out of 30 comparisons no consistent (year, site) effects on non-target insects except slight effects on Nabis bugs (very low numbers, no robust analysis)) and 2 species probably attracted to slightly Lepidoptera infected kernels (corn ear samples).
Comments: plot size above minimum. Sampling quite intensive. Numbers however mostly low < 2 (3)/sample/date, which makes analysis weak.

De la Poza 2005

Comparison LR, insecticide (imidachloprid seen dressing) and non-treated
Design: 2 sites, 3 years, 3-4 replicates, plots 0.4-0.7 ha (= 65*65 – 75*75 m)
Sampling: visual 10-25 plants/plot 5 times per season; pitfall traps 3-5 / plot, intervals (all activity/density or momentary abundance)
Taxonomy: all to family level substantial numbers: Araneae, Anthocoridae (Orius), Nabidae, Coccinellidae , Carabidae, Staphylindae, Chrysopidae
General conclusions: variation between sites > between years > between treatments for most groups. Difference between per group in the order of 2-5 times (in the non zero values), indicating mainly the random and sampling variance.
Comments: Though difference in aphids was recorded, no response of predators could be detected, interpretation considers separate dates as replicates which is not valid. No species level conclusions.
The dominance of other factors than Bt (sites, years) is obvious or at least detectable when compared to tretament effects.

Farinos *et al* 2008 (abundance in time and community structure)

Comparison: transnegic LR, non transgenic and transgenic + imidachlopríd (see treatment)
Design: 1 site, 3 treatments, 3 replicates, plot size: 0.50 ha.
Sampling: 5 pitfall traps / plot, interval samping 3 days/2 weeks june to sept.
Taxonomy: main groups: spiders, ground beetles, rove beetles (and centipedes, click beetles)
Statistics: data analysis buy ANOVA with repeated measures
General conclusions, for carabids, staphylinids and spiders strong year to year variation but only occasional erratic significant difference between treatments all over the season. (spiders and ground beetles 10-20 indiv/trap/date, so quite high), high dominance of 1 or 2 species in all groups that fully determine total numbers. Great temporal variation in number between and within years but patterns between plot were very similar
Comments: Straightforward approach. No serious comments apart from the equalizing effect of mobility which is not discussed. Dominance of some species that are useful for analysis

Gathmann *et al* 2006

Comparison: LR (Cry 1Ab) vs near-isogenic vs near isogenic + pyrethroid
Design: randomized block, 2 fields (5 + 3 reps), 3 treatments, plot size: 56*45, with each plot 20*1 m weed strip included (Sinapsis alba and Chenopodium album).
Sampling: pollen deoposition in weed strips (by adhesive tape), lepidopteran larvae on beating trays, 3 years. 10 subsamples per plot. Intensive!
Taxonomy: lepidoptera larvae to species level. Only Plutella and Pieris rapae sufficient numbers to be analysed.
Statistics: non parametric confidence intervals of ratio. With equivalence testing, ratio interval does not include 1.0
General conclusions: more pollen deposition in Bt plots!, insecticide treatment always lower than Bt-mais or isogenic without insecticide, the last 2 are not different (before and after pollen shed). The lack of any effect on the caterpillars on BT plot strips despite heavy pollen shed may be due to early larval development in relation to pollen shed.
Comments: apparently plot size and isolation of weedy strips sufficient to wipe out differences from the insecticide

plot. Elegant and useful experiment. Balanced discussion linking thought about the population effects on landscape scale. Also including remarks on the problem of variability, power analysis and high experimental costs to show significant effects only on the common species.

Habustova 2012

Comparison: LRM (Cry1Ab) vs NTM- isogenic
Design: regular design with 5 replicates for each treatment, plot size 71*71, 2 meter between plots .
Sampling: 4 plants per sampling date (in plastic to the lab) / plot
Taxonomy: aphids and thrips most abundant R. padi and M. dirhodum, F occidentalis, Orius, Chrysopa, Coccinellidae reliably assessed
General conclusions: Years are significant influence on numbers in all groups as well as sampling date. No consistent differences between treatments The homogeneous settlement of aphids is spatially dependent and trigger teh predator responses.
Comment: Orius (mostly nymphs), Chrysopa (eggs), and Cocc (larvae?) just followed their prey populations indicating that the adults respond by movement and heterogeneous egg laying. So what is actually measured?

Ludy and Lang (2006) spiders, germany

Comparison: LR (Cry 1Ab) Bt mais (field and margins) vs non Bt maize (field and margins
Design: 3 sites, 2 treatments, no replicates, 2 hectares per plot, nettle margins 6-7 meters wide), 3 years
Sampling: 10 suction samples per plot (field or margin), 5 dates
Taxonomy: 3 spider guilds (space web spiders, orb web spiders and hunting spiders) based on species identification
Statistics: ANOVA with repeated measures.
General Conclusions: both habitats 'colonized' mainly by orb web spiders (Theridiidae an Linyphiidae), abundance and number higher in nettle margins than in maize fields.
Bt treatment no effect on numbers, community structure, and species number. Rather stable numbers on guild level.
Comments: focused study, large scale but low number of sites. Quite intensive absolute sampling but still low numbers (only 5 species quite abundant)! Pooling of species in guilds defendable but still a bit tricky. Actually colonization is measured rather than standing population (in particular the higher numbers, non Bt responsive orb web spiders … this weakens the difference detection power, but at least they mention mobility in the discussion. Some inconsistent effect is attributed to a possible pleitropic effect in the maize structure.

Orr and Landis (1997)

Comparison: Cry1Ab vs isogenic nontransformed
Design: 3*2 rrandomized block design, 3 reps, 2 treatments, plot size 65*63
Sampling: 3 dates, visual countings on 5 plants / plot
Taxonomy: Orius insidiosis, Coleomegilla maculata (Cocc), Chrysopa larvae, + egg-mass predation test Ostrinia.
General conclusions: numbers are very low (0.5 – 2.5 / plant /day = 2-10 / plot) for Orius and even less Coccinellidae > Chrysopa, no significant differences.
Comments: numbers very low, prey (apart from Ostrinia) not measured.

Priesnitz et (2013)

Comparison: CR (Cry3Bb1) v.s. near-isogenic and 2 conventional maize
Design: randomized complete block , with 8 replicates, 4 treatments (hybrids), plotsize 0.13 ha (40*31m), 3 years
Sampling: pitfall traps continuous over the growing season , one (1) trap per field. Also living ground beetles were sampled for Cry-uptake tests.

Statistics: ANOVA and pairwise comparisons. And equivalence tests.
Taxonomy: Only ground beetles, all to species level
General conclusions: The differences among block is much higher that between maize varieties, species tend to be spatially clustered disturbing the analysis. The 11 most common species (4 species being really dominant) were sufficiently numerous to be analysed, most being equivalent and other inconsistently (but not significantly) different in the years. Ratio differences outside range 0.5 to 2.0 were not found (11 species * 3 years). Some higher in transgenic some higher in non-transgenic . Conclusions remain when species are pooled to the family level Situation reflect 'natural' variation in maize fields. In 200 live carabid beetles caught, in 82 carabids CryBb1 could be detected. (40%).
Comments: the study is extensive in replicates, but poorly sampled as no more than 100 to 300 beetles (including up to about totally 60 species) were trapped per plot per year in only one trap (causing to much residual variation). Many species analysed not more 10 ind. per plot on average, so variation to big to detect any differences. Though the authors recognize the variability problem they conclude on negligible environmental effects, without power analysis.

Rauschen *et al* 2008 an evaluation ...

Comparison: LR (Cry 1Ab) vs near-isogenic vs near isogenic + insecticide pyrethroid (Cyfluthrin)
Design: randomized block, 2 fields (5 + 3 reps), 3 treatments, plot size: 56*45; 2 years
Sampling: visual inspection of 4 plants per plot (adults and nymphs), 4 sweepnet transects / plot, sticky traps alles 1 maal in augutsus
Taxonomy: 5 plant/leaf hopper identified to species level from sweep/sticky and watertrap. Zyginidia scutellaris, Empoasca pteridis, Pasmmotettix alienus, Laodelphax striatella. Only Z. scutellaris abundant enough for analysis.
Statistics: equivalence tests
General conclusions: insecticide treatment much lower numbers caught that in transgenic and non transformed isogenic that have equals numbers
Comments: Apparently no significant recolonisation o fthese fytophagous leafhoppers after insect treated (mid july) plot when many in larval stage. No detectable Bt-maize effect but severe insecticide effect on larval stages (lost generation).

Rauschen *et al* 2010 (only CryBb1 treated here while also Cry1Ab data available)

Comparison: CR (Cry3 Bb1) vs near-isogenic and conventinal (not insecticide treated)
Design: systematically randomized, 8 replicates), 2 treaments, plot size 0.13 ha (50*50 m)
Sampling: 10 whole cobs/plot, sweep netting on 2-3 days per year, visual assessments 4 plant/plot,
Taxonomy: focused on Coleoptera. All identified to genus or species level.
Statistics: equivalence tests were used and variance-abundance correlation was tested.
General conclusions: though many replicates and samples , estimated abundances are very low and only for 4 species an equivalence test could be performed inclusing 2 chrysomelidae (Oulema and Phyllotetra, and one ladybeetle Propylea 14-punctata). For these species no over-dispersion could be detected.
Comments: Coleoptera densities very low, 2 harmful species in chrysomelidae are testable, + one ladybeetle being very mobile, common and judged not-senstive in all other studies.
Very low numbers were caught of most species. Only 2 species indicated effects (by chance?)

Svoboda et 2013 ***

Comparison: herbicide and coleoptera ristant CRM, insecticide (chlorpyrifos) treated, and 3 references (whole trial treated herbicide around emergence)

Design: randomized block. 1 site, 3 years, plot size 63*81 meter
Sampling: pitfall traps 5 per plot, 7 intervals of one week spread over the season, for spiders only (focused)!
Taxonomy: almost all to species level. Most abundant Oedothorax apicatus, Pardosa agrestis, Pachygnatha degeeri
General conclusions: Total 79 species (42-44 specie / year). 7 species makes up more than 90% of the total spider numbers. Total activity/abundance between years and intervals (sampling dates) most significant and varied 5-10 fold where between treatments in the same period only max 2 fold and insignificant. Other factors of field management had major influences on abundance. Total year abundance in Insecticide consistently lower (but not stat sign different) from other treatment. No further consistent differences found between treatments. Border rows had significantly higher numbers (coming from outside). The multivariate analysis of the total community gave similar significance as the numbers. They discuss the strong influence of the surrounding habitats despite the considerable plot-size.
Comments: ...

Stephens *et al* 2012 (USA) (carabids)

Comparison: BR (Cry3Bb1) vs iso insecticde (tefluthrin at seeding) vs untreated ISO
Design: all plots received gaucho. 4 Bt plots, 7 control isogenic and 7 Iso+insecticide). Plot size 50*50 meter. 3 years.
Sampling: 8 traps per plot, opened one day a week for about 20 weeks.
Taxonomy : 3 dominant (tot > 70%) carabid species were counted Agonum muelleri, Poecilius lucublanus and Pterostichus melanarius
Conclusions: no treatment effects whatsoever in all year, 2001 much higher than 2002 and 2003.
Comments: 1 day week trap open nog enough, catches very low (1 beetle per trap!), including spread over the season give enough opportunities to overcome early insecticide treatments. Strange that carabids are summed up! Because of density?

Peterson *et al* 2011 (meta-analysis spiders of BT corn, cotton, rice and egg-plant)

Comparison: LRM (Cry 1Ab) vs NTM- vs NTM+
Design: not relevant; it concerns a metanalysis over number of studies
Sampling : their analysis shows that many trapping and assessment systems are applied which make studies difficult to compare.
General Conclusions: In their meta-analysis they show that in almost all field studies (mostly maize but also other crops) no significant effects have been shows when Bt is compared to insecticide free isogenic crops, but in many cases Bt, or unsprayed non-Bt doesn better than insecticide treated non modified. But the main and serious message that they provide is that most studies are performed at the family level while many species respond differently and more taxonomic resolution to species level is required to get a clear picture They also not the huge variation in natural densities (among agricultural fields 100 fold or more), the variation in species composition and the strong dominance of some species among the great richness in other species. They also state the importance of ballooning and phenological (within year) dynamics, and the extreme variation in sampling techniques and efforts which is a caveat in comparing and interpretation of results.
Comments: Despite all the drawback in methodology they surprisingly ply for more research to get a broad picture instead of improving the studies. The implications of ballooning behaviour for interplot effects is not mentioned
.

Appendix 2 Selection Criteria for non-target invertebrates according to Todd *et al*. 2008 (tables copied from the publication)

Table 3. List of selection criteria, and the parameters used to define each of these criteria, used by the Priority Ranking of Non-Target Invertebrates (PRONTI) method. "Stressor" refers to the new GMO, while the term "receptor" is used to encompass both target and non-target invertebrates in the receiving ecosystem.

Selection criteria	Defining parameters
1) Could the stressor pose a hazard to the receptor? (H)[1]	H1. Identification of the stressor/s
	H2. Potential direct effects of stressor on receptor
	H3. Possible indirect effects of stressor on receptor
2) Could the receptor be exposed to this stressor? (E)	E1. Receptor found in receiving area
	E2. Receptor's diet
	E3. Receptor's use of plant
3) Could there be an impact on the ecosystem if this receptor is affected? (I)	I1. Receptor's biomass (S)
	I2. Food web links from receptor to other organisms in the ecosystem (S)
	I3. Receptor's special ecological function (S)
	I4. Receptor's resilience (*i.e.*, ability to avoid the hazard or reduce its exposure level) (R)
4) Do people value this receptor? (V)	V1. Value of the receptor to indigenous human cultures
	V2. Conservation value of the receptor
	V3. Value of the receptor to society
	V4. Economic value of the receptor
	V5. Links from the receptor to higher levels in the food web (including human diets)
5) Can researchers perform tests with this receptor? (T)	T1. Accessibility of the receptor
	T2. Generation time of the receptor
	T3. Rearing protocols available for the receptor
	T4. Bioassay protocols available for the receptor

[1] Representative symbols used in the text.

Table 4. Receptor attributes used to measure species' population resilience. The attributes were used to inform four resilience factors which may be used by the species to mitigate the effects of the stressor: resistance, behavior, migration and recovery. Surrogate attributes, listed in the third column, are used by the Priority Ranking of Non-Target Invertebrates (PRONTI) method when the receptor attributes listed in the second column are unknown.

Resilience factor	Receptor attributes likely to result in this resilience factor	Surrogate attributes used to estimate receptor resilience
Resistance	Published reports of the receptor showing resistance to the stressor	Published reports of taxonomic family members showing resistance to the stressor
	Published reports of the receptor showing resistance to related stressors	Published reports of the receptor showing resistance to different, unrelated stressors
	Receptor is known to be exposed (*e.g.*, known to eat the stressor, or to prey on species that eat the stressor)	Information that indicates the receptor may be exposed (*e.g.*, may eat the stressor, or prey on species that may eat the stressor)
	The stressor is known to have an effect on the receptor (*e.g.*, the receptor is known to be susceptible to the change in protein expression in the plant giving an opportunity for resistance to arise)	Information that indicates the receptor may be affected by the stressor (*e.g.*, is susceptible to a related hazard, or taxonomic family members are known to be susceptible to the hazard)
	Density of the receptor population in the ecosystem is known to be large, so that it is more likely that resistant individuals may arise	Receptor is not known to be rare or threatened in the ecosystem of interest, but actual population size is unknown
	Reproduction rate is high	Reproduction rate of other species in the same genus is high
	Number of generations per year is high	Number of yearly generations of other species in the same genus is high
Behavior	Feeding stage coincides with presence of the stressor allowing individuals to display avoidance behaviors	Feeding stage may coincide with the presence of the stressor
	Known to be a generalist feeder with a flexible diet that may allow it to avoid eating the new stressor	May be a generalist feeder
	Known to have mechanisms to detect food quality and can avoid eating parts of the stressor or prey that are unpalatable	Information indicates the receptor may be able to detect the change in the stressor or identify prey that have eaten the stressor
	Life stages that are exposed to the stressor are highly mobile, allowing them to find alternative foods	Information on mobility and dispersal distances indicates the exposed life stages may be able to disperse to find other food sources
	Receptor's diapause known to reduce exposure to the stressor	Receptor's diapause may reduce exposure to the stressor

Table 4. Continued.

Migration	Receptor known to be highly mobile and able to disperse into refuge areas or other ecosystems	Receptor has several modes of dispersal and the dispersal distances may be quite large
	Receptor's population density in the receiving ecosystem is high, increasing the likelihood that some individuals may migrate out of the receiving environment	Receptor is not known to be rare or threatened, but actual population density is unknown
	Receptor's population density in other ecosystems is high, suggesting the receptor will survive in areas to which it disperses	Receptor has been found in several other ecosystems, or has a widespread distribution
Recovery	Receptor's reproduction rate and number of generations per year will allow it to quickly re-populate an area	Estimates of receptor's reproduction rate and number of generations per year suggest it might be able to re-populate the area
	Populations in other ecosystems are large, providing a source for re-introductions	Receptor is not known to be rare or threatened or has been found in several other ecosystems or has a widespread distribution
	Receptor known to be highly mobile and able to move into vacated areas	Receptor has several modes of dispersal and the dispersal distances may be quite large