From Science to Policy

Improving environmental risk assessment and management of chemicals

Marlene Ågerstrand

Division of Philosophy
Royal Institute of Technology (KTH)
Stockholm, Sweden, 2012
The road we have long been traveling is deceptively easy, a smooth superhighway on which we progress with great speed, but at its end lies disaster. The other fork of the road — the one less traveled by — offers our last, our only chance to reach a destination that assures the preservation of the earth.

Rachel Carson, Silent spring, 1962

What you risk reveals what you value.

Jeanette Winterson, Sexing the Cherry, 1989

We have too many high sounding words, and too few actions that correspond with them.

Abigail Adams, Letter to John Adams, 1774
List of publications:

**Paper I**


**Paper II**


**Paper III**


**Paper IV**


**Paper V**


The five papers are reprinted with permission of respective journal.
Abstract

A complex process like risk assessment and the subsequent risk management decision making should be regularly evaluated, in order to assess the need to improve its workings. In this thesis three related matters are addressed: evaluation of environmental risk management strategies, evaluation of environmental risk assessments, and how ecotoxicity data from the open scientific literature can be used in a systematic way in regulatory risk assessments. It has resulted in the following: a publically available database with ecotoxicity data for pharmaceuticals (Paper I); an evaluation and review of the Swedish Environmental Classification and Information System for pharmaceuticals (Papers II and III); a comparison of current reliability evaluation methods and a reliability evaluation of ecotoxicity data (Paper IV); and an improved reliability and relevance reporting and evaluation scheme (Paper V).

There are three overall conclusions from this thesis:

(1) Ecotoxicity data from the open scientific literature is not used to the extent it could be in regulatory risk assessment of chemicals. Major reasons for this are that regulators prefer standard data and that research studies in the open scientific literature can be reported in a way that affects their reliability and the user-friendliness. To enable the use of available data more efficiently actions must be taken by researchers, editors, and regulators. A more structured reliability and relevance evaluation is needed to reach the goal of transparent, robust and predictable risk assessments.

(2) A risk assessment is the result of the selected data and the selected methods used in the process. Therefore a transparent procedure, with clear justifications of choices made, is necessary to enable external review. The risk assessments conducted within the Swedish Environmental Classification and Information System for pharmaceuticals vary in their transparency and choice of method. This could come to affect the credibility of the system since risk assessments are not always consistent and guidelines are not always followed.

(3) The Swedish Environmental Classification and Information System for pharmaceuticals contribute, in its current form, to data availability and transparency but not to risk reduction. The system has contributed to the general discussion about pharmaceuticals' effect on the environment and made data publicly available. However, to be an effective risk reduction tool this is not sufficient.

Keywords: environmental risk assessment, chemicals, pharmaceuticals, transparency, predictability, accuracy, consistency, voluntary initiatives, ecotoxicity data, reliability evaluation, relevance evaluation, non-standard test data, CSR, risk management.
Populärvetenskaplig sammanfattning

När man undersöker hur giftig en kemikalie är för miljön används olika typer av ekotoxikologiska tester. Vanligtvis får tre separata tester med tre arter från tre olika nivåer i näringskedjan representera ekosystemets komplexitet. I en miljöriskbedömning jämförs den koncentration som förväntas hittas i naturen med den koncentration som inte visat sig vara skadlig i testerna. Om koncentrationen i naturen är högre än den koncentration som inte var skadlig i de ekotoxikologiska testerna anses det finnas risk för negativa konsekvenser på miljön. Detta kan leda till att åtgärder vidtas för att minska de högsta miljökoncentrationerna.


Både miljöriskbedömningar och miljöinitiativ måste utvärderas för att kunna förbättras. Dessutom måste metoder som förenklar och förbättrar den nuvarande riskbedömningsprocessen utvecklas så att riskbedömningar i framtiden kan bli mer robusta, konsekventa och transparanta.


Det finns tre övergripande slutsatser från denna avhandling:

(1) Ekotoxikologiska studier från den öppna vetenskapliga litteraturen används inte i den utsträckning de skulle kunna i miljöriskbedömning av kemikalier. Detta beror delvis på att myndigheter föredrar standardiserade studier, delvis på att forskningsstudier från den öppna vetenskapliga litteraturen inte rapporteras på ett standardiserat sätt vilket ibland påverkar kvalitén och användarvänligheten. För att ändra den rådande situationen måste forskare och redaktörer för vetenskapliga tidskrifter se till att studier som publiceras lever upp till de kvalitets- och rapporteringskrav som ställs, och myndigheter måste öppna upp för en ökad användning av tillgängliga studier.
(2) En viktig del i riskbedömningen är val av studier och metoder som ska användas i bedömningen. En transparent process är därför avgörande för att riskbedömningen ska kunna granskas av utomstående. Riskbedömningarna inom det svenska miljöklassificeringssystemet för läkemedel varierar både i val av studier och metoder samt i transparens, vilket har resulterat i inkonsekventa bedömningar som inte alltid följer systemets egna riktlinjer. Detta kan komma att påverka trovärdigheten för klassificeringssystemet.

(3) Det svenska miljöklassificeringssystemet för läkemedel tillhandahåller miljöinformation om produktarna på den svenska marknaden, men för att vara ett effektivt verktyg som bidrar till minskade miljörisken är detta inte tillräckligt.
Contents

Abstract ................................................................................................................................................. 4
Sammanfattning .................................................................................................................................... 5
Acknowledgements ............................................................................................................................... 8
Abbreviations ......................................................................................................................................... 9
1. Introduction ...................................................................................................................................... 10
   1.1 Aim ....................................................................................................................................... 11
2. Background ..................................................................................................................................... 12
   2.1 Environmental risk assessment ........................................................................................... 12
   2.2 Evaluation of risk assessments ............................................................................................ 14
   2.3 Standard and non-standard test data .................................................................................. 15
   2.5 Reporting and evaluation of test data ................................................................................ 18
   2.6 Pharmaceuticals - use, characteristics, occurrence and effects .......................................... 19
   2.7 Voluntary initiatives .............................................................................................................. 21
   2.8 The Swedish Environmental Classification and Information System for pharmaceuticals... 22
3. Preview of papers ............................................................................................................................ 24
   3.1 Paper I ................................................................................................................................. 24
   3.2 Paper II ................................................................................................................................ 25
   3.3 Paper III ............................................................................................................................... 26
   3.4 Paper IV ............................................................................................................................... 28
   3.5 Paper V ................................................................................................................................ 30
4. Discussion ....................................................................................................................................... 31
5. Future work ...................................................................................................................................... 35
References .......................................................................................................................................... 36
Acknowledgements

Many generous people have given me inspiration, knowledge, and mental and practical support during the past five years. Some I have had the privilege to work with throughout the whole process and others for a shorter period.

First I would like to thank my supervisor Christina Rudén, for your sharp-eyed analysis, far-seeing ideas, and for believing in me. Always. I also would like to thank my assisting supervisors Magnus Breitholtz and Misse Wester for enthusiasm and encouragement. The three of you are inspiring examples of the kind of researcher I would like to be: curious, brave, open-minded, cooperative and active in the public debate.

To my colleagues at the Division of Philosophy at the Royal Institute of Technology (KTH): thanks for all the moral and ethical challenges, comments on manuscripts, tasty lunches, social activities, and support in troubled times. Special thanks to Linda Molander, Linda Schenk, and Emma Westerholm. To the MistraPharma team: I have learnt a great deal from you. To my new colleagues at the Department for Applied Environmental Science at Stockholm University: thanks for your warm welcome.

This thesis includes work conducted in collaboration with the German Federal Environmental Agency (UBA), and especially Anette Küster. Thanks for inspiring discussions. Other important contributors who I wish to express my gratitude to are Robert Kase, Florian Kiel, Åke Wennmalm, Gisela Holm, and Bengt Mattson.

I am also privileged with fantastic friends who accompany me on adventures on mountains, cliffs, snow and ice, football fields or just everyday life. Thanks for friendship, sisterhood, advice, comfort and sometimes a kick in the butt. I hope you will enjoy the party!

I have the best family, you have always been there for me. ♥

My work has been funded by Formas (the Swedish Research Council for Environment, Agricultural Sciences and Spatial Planning) and Mistra (the Foundation for Strategic Environmental Research). Financial support has also been granted by Zonta International, Technologia, Sixten Gemzéus stifteelse, the German academic exchange service (DAAD), Kungliga skogs- och lantbruksakademin, Ångpannföreningen, Knut och Alice Wallenbergs stifteelse, and Stiftelsen Oskar och Lili Lamms minne. Their support is gratefully acknowledged.
Abbreviations

API  Active Pharmaceutical Ingredient
ECHA  European Chemicals Agency
EMA  European Medicines Agency
EU  European Union
ISO  International Organization for Standardization
L(E)C₅₀  Lethal (Effect) Concentration: the lowest identified effect concentration where fifty percent of the tested population have been found to be affected in experiments
LIF  The Swedish Association for the Pharmaceutical Industry
Log Kow  Octanol/water partition coefficient
NOEC  No Observed Effect Concentration
OECD  The Organisation for Economic Co-operation and Development
OEL  Occupational Exposure Limit
PBT  P = persistence in the environment, B = bioaccumulation in organisms, T = toxicity towards organisms
PEC  Predicted Environmental Concentration
PNEC  Predicted No Effect Concentration
QSAR  Quantitative Structure Activity Relationship
REACH  Registration, Evaluation, Authorisation and Restriction of Chemical substances (European regulation, EC 1907/2006)
SECIS  The Swedish Environmental Classification and Information System for pharmaceuticals
TGD  Technical Guidance Document
US EPA/EPA  United States Environmental Protection Agency
US FDA  United States Food and Drug Administration
1. Introduction

This thesis reports from a project concerning regulatory risk assessment of chemicals. Regulatory risk assessment encompasses a chain of activities: from data requirements and the generation of exposure and effect data, via the interpretation and evaluation process, to decisions about risk management strategies. A complex process like this should be regularly evaluated, in order to assess the need to improve its workings and thereby contributing to improving the outcome of future risk assessments. Three different subjects related to risk reducing processes are addressed in this thesis: evaluation of environmental risk management strategies, evaluation of environmental risk assessments, and how ecotoxicity data from the open scientific literature can be used in a systematic way in regulatory risk assessments.

Risk management strategies can differ a great deal in both scope and design as well as in effectiveness towards reducing risks. Within the OECD (the Organization for Economic Co-operation and Development) member countries voluntary initiatives are increasing, often presented as cheap and flexible alternatives to laws and regulations. There is a considerable amount of literature indicating that companies can profit from taking voluntary actions, but little is known about their effectiveness towards reducing environmental risks (OECD 2003). The Swedish environmental classification and information system for pharmaceuticals (here called SECIS) is an example of a voluntary environmental initiative taken by the industry. In this thesis its start-up phase is scrutinized and risk assessments are evaluated.

The result from a risk assessment is only as good as the methods and data used and therefore the risk assessment process needs to be reviewed and updated as new scientific data and methods are developed. Aspects of the risk assessment process considered in this thesis are how data is gathered and then selected, how data is reported in the open scientific literature and consequences of this, reliability and relevance evaluation of ecotoxicity data, transparency of the risk assessment process, accuracy of risk assessments in relation to the guidance document, and consistency between results from different risk assessments.

Research funded by tax payer’s money should benefit the society; to maximize the usefulness, research aiming at risk identification should be designed and reported in a way that enable new findings to be incorporated into the regulatory risk assessment process. Today ecotoxicity data from the open scientific literature is not used to the extent it could be in environmental risk assessments of pharmaceuticals. This is partly because the academia and the regulators have different procedures for the scientific work, even though the objective may be the same. This thesis presents a review of available methods for reliability evaluation for ecotoxicity test, but also a new and improved method focusing on both evaluation and reporting of test data.

All five papers in this thesis concern pharmaceuticals: The consumption of pharmaceuticals is increasing and this, in combination with pharmaceuticals’ inherent design to interact with biological systems, makes environmental effects from pharmaceuticals a growing concern. So far there are
only a handful of reports with confirmed effects on non-target organisms in the environment, but laboratory studies show that pharmaceuticals can cause adverse effects in these organisms if exposure occurs.

1.1 Aim
This is a thesis in regulatory ecotoxicology with focus on risk assessment and risk management of pharmaceuticals, but the presented ideas and methods can be applied in risk assessment of other types of chemicals and in health risk assessment. The overall aim of my work is to improve the scientific basis and increase the accuracy, transparency and predictability of risk assessments of chemicals. I have addressed these issues using interview studies and literature reviews in collaboration with other researchers and risk assessors.

The thesis consists of five papers reporting from two different research projects. The first project concerns SECIS. The aim of that project was to increase the environmental effectiveness of the classification system by conducting a thorough evaluation resulting in recommended improvements. The second project is conducted within the MistraPharma research programme. The aim of MistraPharma is to identify and reduce environmental risks caused by the use of human pharmaceuticals.

Specific research questions have been:

1) What were the driving forces, intentions and expectations underlying the development and implementation of SECIS? (Paper II)
2) Are the risk assessments and classifications made within SECIS accurate in relation to the guidance document? (Paper III)
3) How would the use of ecotoxicity data from the open scientific literature affect the outcome of risk assessments conducted within SECIS? (Paper III)
4) Are the risk assessments within SECIS consistent when different companies assess the same substance? (Paper III)
5) How can the reliability of ecotoxicity data from the open scientific literature be evaluated? (Paper IV)
6) How does choice of reliability evaluation method affect the outcome of the evaluation? (Paper IV)
7) How can ecotoxicity data from the open scientific literature be used in a systematic way in regulatory risk assessment? (Paper I and V)
2. Background

2.1 Environmental risk assessment

The overall objective of the regulatory risk assessment process is to identify and characterize environmental risks and on the basis of this make decisions with the purpose to prevent unacceptable harm to the environment (European Commission 2003). The different steps of the risk assessment and risk management process are visualized in figure 1. Risk assessment is the identification and quantification of the risk connected to a specific chemical. It entails problem formulation, exposure and effect assessment, and risk characterization. Risk management is the decision making process where results from the risk assessment are combined with political, social, economic and engineering information (Leeuwen & Vermeire 2007).

![Figure 1. Outline of the regulatory risk assessment process and its connection to risk management (Adjusted from National Research Council 1983).](image)

In the exposure assessment concentrations in the environment resulting from different emissions are estimated, i.e. the predicted environmental concentration (PEC). A PEC is estimated by using data derived from actual measurements (retrospective assessments) or from models based on estimations such as (estimated) sales data (prospective assessments) (European Commission 2003).

In parallel, an effect assessment is conducted. Experimental data are reviewed to establish the lowest identified no observed effect concentration (NOEC) or the lowest identified effect
concentration where fifty percent of the tested population have been found to be affected in laboratory experiments ($\text{L(E)C}_{50}$). This concentration is thereafter divided by assessment factors to derive a predicted no effect concentration (PNEC). Assessment factors are applied to compensate for variations in sensitivity within and between species, transfer from laboratory data to field conditions, and for extrapolating short-term data to chronic data. The purpose of the PNEC is thus to be used as a reference concentration below which unacceptable effects in the environment will most likely not occur (European Commission 2003).

In the risk characterization a risk quotient is calculated by comparing PEC with PNEC. Dependent on the PEC/PNEC ratio it is decided whether a substance presents a risk to organisms in the environment. A risk ratio below 1, i.e. where PEC is lower than PNEC, indicates there is no concern according to the current state of knowledge, while a ratio above 1 normally leads to further evaluations (European Commission 2003). In addition to the PEC/PNEC risk characterization a PBT assessment will be performed, by combining information about the substance’s potential to persist in the environment (P), its capacity to bioaccumulate in organisms (B), and its toxicity (T) (European Commission 2003).

The European legislation for environmental risk assessment of pharmaceuticals (European Medicines Agency 2006) follows the general structure described above. However, the process is divided into two phases with the later consisting of two tiers. In Phase I the PEC for surface water is calculated and the logKow is measured. If the PEC value is equal to or above the cut-off point 0.01 µg/L a Phase II analysis should be performed. Lipophilic substances with a logKow > 4.5 should be screened for persistence, bioaccumulation and toxicity (European Commission 2003). Pharmaceuticals that are known to affect the reproduction of vertebrates or invertebrates at concentrations below 0.01 µg/L should also enter Phase II. In Phase II Tier A physico-chemical properties, environmental fate and effect studies are reviewed and a PNEC is calculated. Standard long-term toxicity tests on algae, daphnids, and fish (OECD guidelines 201, 211 and 210) are performed, and appropriate assessment factors are applied to the identified NOEC, in order to determine the PNEC. If the ratio PEC/PNEC is above 1, an extended environmental fate and effect analysis, according to Tier B in Phase II is needed (European Medicines Agency 2006).
2.2 Evaluation of risk assessments

If the purpose of risk assessment is not only to point out substances that are proven harmful but also the substances that may cause harm, risk assessors cannot depend solely on what is scientifically proven. The risk assessment process therefore contains uncertainties. Consequently the influence of expert judgment is higher compared to pure science. Wandall (2004) concludes that it is not possible to avoid values in science or risk assessment and it is therefore better to report, discuss and assess them instead of ignoring them. Transparency is thus a key issue in the risk assessment process. Areas where the transparency tends to be lower than preferred are in the selection and evaluation of data, use of assessment factors, and disclosure of the participants and their affiliation in the groups performing risk assessments.

Rudén (2002) performed a detailed comparison of 30 different cancer risk assessments for the chemical trichloroethylene. The results showed that the risk assessors came to different conclusions regarding its potential to cause cancer, and these differences were due to both scientific and policy related issues. Data selection differed and this could not be explained by whether the data was available or not at the time of the assessment. A biased data selection may therefore have influenced some of the risk assessors’ conclusions. Interpretation of data was also identified as an area where risk assessors disagreed and this also affected the overall outcome of the risk assessments.

Schenk (2010) also identified data selection and data evaluation as factors influencing risk assessments. In a review of documents for settings of occupational exposure limits (OELs) only one fifth referred to all available key studies and the evaluation of the studies varied significantly. Assessment factors were only explicitly given for one third of the OELs. When comparing OELs with an explicit assessment factor the safety margin was 2.1 higher than for the OELs without (Schenk & Johanson 2010).

Beronius et al. (2010) found that for the endocrine-active chemical bisphenol A non-standards studies were considered less reliable and therefore assigned less weight in the majority of the reviewed risk assessments. Instead all risk assessments but one used two standard test studies as key studies. Assessment factors were used in half of the risk assessments with a range between 40 and 1000, scientific justifications for this were seldom provided.

These studies highlight the importance of transparency as well as raise the concern about selection bias. Bias is defined as an influence that tends to produce results that depart systematically from the true values. It can be caused by methodological errors, e.g. in the experimental design and use of statistical methods, experimenters’ bias, e.g. collection and interpretation of data in a selective way, or other factors such as publication bias (Wandall et al. 2007). Hazard assessment guidelines and regulatory requirements often explicitly advocate the use of expert judgment in the evaluation process but further practical guidance is however seldom provided (Hamilton et al. 2006).
2.3 Standard and non-standard test data

One way to group toxicity testing methods is regarding whether they are performed according to a standard guideline provided by an official standardization organization (standard test) or not (non-standard test).

The major advantages of using standard tests are that the results are directly comparable between substances and that the data they generate will be readily accepted across jurisdictions. Test guidelines also contribute to promote the reliability and reproducibility of the study because of the detailed test procedures and extensive reporting of data that is required. The major disadvantage of standard test methods is that they do not always represent the most relevant testing approach depending on the type of endpoint under investigation and are slow to adapt to emerging technologies and effects of interest. Therefore, results from non-standardized tests may in some cases be more sensitive and thereby contribute additional and significant information to a risk assessment.

It should however also be noted that non-standardized experiments can be just as reliable (reproducible) as tests performed under strict implementation of a test standard, and that following a standard will not automatically ensure that the test has sufficient relevance for risk assessment purposes. Criticism against using only standard tests in risk assessments has been raised in the case of bisphenol A where a large number on non-standard studies are omitted from regulatory risk assessments on behalf of two standard tests (Myers et al. 2009; vom Saal & Myers 2010; vom Saal et al. 2007). Choices of data have also created disagreements in health risk assessment of decabrominated diphenyl ether used as flame retardant (Alcock et al. 2011). Here non-standard data (Viberg et al. 2003; Viberg et al. 2003; Eriksson et al. 2002) have been criticized for flaws in the experimental design by both other researchers (Vijverberg & Berg 2004) and toxicologists from the bromine industry (Hardy 2004).

An overview of risk assessments guidance documents for several chemical groups show that standard tests are promoted. Phrases like “preferably”, “recommended”, “need to be carried out in accordance with”, “must be conducted according to”, and “can be accepted if the guideline is comparable with those guidelines mentioned” are used to indicate their importance (Table 1). However, several guidance documents also open up for use of other types of data when applicable. According to all guidance documents, studies “should” or “must” be conducted in compliance with Good Laboratory Practices (GLP), the implications of this is that data from independent researchers could not be used since very few, if any, non-commercial laboratories are GLP-certified (US FDA 1998; European Medicines Agency 2006; European Medicines Agency 2004; European Chemicals Bureau 2000; European Chemicals Bureau 2003; European Commission 2001; European Chemicals Agency 2008).
<table>
<thead>
<tr>
<th>Substance group</th>
<th>Guidance document</th>
<th>Recommended test methods</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pharmaceuticals</td>
<td>Guidance for Industry. Environmental Assessment of Human Drug and Biologics Applications. US FDA. 1998.</td>
<td>Section IV. D. Test methods and report formats are provided in the FDA Environmental Assessment Technical Handbook. Equivalent tests, such as those provided by the EPA (40 CFR 796 and 797), the Organization for Economic Cooperation and Development (OECD), or other validated, peer-reviewed methods can be used.</td>
</tr>
<tr>
<td>Human pharmaceuticals</td>
<td>Guideline on the environmental risk assessment of medicinal products for human use. EMA. 2006.</td>
<td>Section 5. Experimental studies should preferably follow the test protocols issued by the European Commission, Organization for Economic Co-operation and Development (OECD) or the International Organization for Standardization (ISO).</td>
</tr>
<tr>
<td>Veterinary pharmaceuticals</td>
<td>Guideline on the environmental impact assessment for veterinary medicinal products phase II. EMA. 2004.</td>
<td>Section 2.5 The specific test guidelines/protocols recommended in Phase II are those finalized by OECD/ISO.</td>
</tr>
<tr>
<td>New substances</td>
<td>Technical Guidance Document on Risk Assessment. Part II. European Chemicals Bureau. 2003.</td>
<td>Part II, Chapter 3, Section 3.2.1.2 The tests for new substances need to be carried out in accordance with the EU test guidelines as laid down in Annex V to Directive 67/548 or, if no EU guidelines are available or they are not applicable, following internationally recognized guidelines, preferably those of the OECD.</td>
</tr>
<tr>
<td>Existing substances</td>
<td>Technical Guidance Document on Risk Assessment. Part II. European Chemicals Bureau. 2003.</td>
<td>Part II, Chapter 3, Section 3.2.1.2 Any new tests carried out for risk assessments under Regulation 793/93 should be conducted according to the testing methods laid down in Annex V to Directive 67/548, or if no EU methods are available or they are not applicable, in accordance with internationally recognized guidelines, preferably those of the OECD (1993b). Greater weight should normally be attached to studies carried out according to current methods (e.g. EU, OECD, or US EPA).</td>
</tr>
<tr>
<td>Biocides</td>
<td>Technical Notes for guidance on data requirements for active substances and biocidal products. European Chemicals Bureau. 2000.</td>
<td>Chapter 1. Section 1.3 According to Article 8(8), as a general principle, tests must be conducted according to the methods described in Annex V of Council Directive 67/548/EEC, according to the most recent adaptation to the technical progress. These are based on those recognized and recommended by international bodies in particular OECD. In the event of a method being inappropriate or not described, other methods used should, whenever possible, be internationally recognized and must be justified.</td>
</tr>
<tr>
<td>Substance group</td>
<td>Guidance document</td>
<td>Recommended test methods</td>
</tr>
<tr>
<td>-------------------------</td>
<td>--------------------------------------------------------</td>
<td>------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
</tbody>
</table>
Tests conducted in accordance with internationally recognized guidelines (even if not specifically recommended in the Annex II or III) can be accepted if the guideline is comparable with those guidelines mentioned in Annex II or III. Tests with species mentioned in the aforementioned guidelines are in principle acceptable, although not all species are indigenous in Europe. |
According to REACH, Article 13(3), tests required for generating information on intrinsic properties of substances shall be conducted in accordance with the test methods included in a Commission Regulation or in accordance with other international test methods recognized by the Commission or the Agency as being appropriate.  
Chapter R.11 Section 1.3.3  
As the aquatic T criterion is based on a NOEC for pelagic organisms, the standardized chronic tests on fish, daphnids and algae are preferred to assess the NOEC. |

2.5 Reporting and evaluation of test data

It is not unusual that there is a mismatch between how independent researchers’ test data are reported (and conducted) and regulators’ needs for risk assessments. Criticism against non-standard data often concern experimental design and statistical analyses (Alcock et al. 2011). Evaluation of ecotoxicity (and toxicity) data usually relies to a significant extent on case-by-case assessments based on expert judgment. However, there are also attempts to make the evaluation process of ecotoxicity (and toxicity) data more structured, either as check-lists or pre-defined evaluation criteria (Klimisch et al. 1997; Hobbs et al. 2005; Durda & Preziosi 2000; Schneider et al. 2009; Küster et al. 2009; Mensink et al. 2008).

A major advantage of using a more structured way of evaluating data is increased transparency and predictability of the risk assessment process. For instance, both a check-list and pre-defined criteria will contribute to ensuring that at least a minimum and similar set of aspects are considered in each evaluation. Pre-defined evaluation criteria may also contribute to increased transparency of the evaluation process to the extent that these criteria are clearly reported to the relevant actors. Disadvantages of using pre-defined evaluation criteria and check-lists are less flexibility and a focus on the general aspects of a study. Risk assessment will always include an element of expert judgment, but it is also important to continuously seek to increase the predictability and transparency of this process.

Most of the evaluation methods currently used for risk assessment purposes focus on the reliability of the data. According to the European Commission’s Technical Guidance Document (TGD) (2003), an evaluation of the data reliability should ensure “the inherent quality of a test relating to test methodology and the way that the performance and results of the test are described”. Basically this evaluation should answer the question: Has the experiment generated and reported a true and correct result? A few methods also include criteria for evaluation of the relevance of different data sets. The assessment of the relevance of the data should describe “the extent to which a test is appropriate for a particular hazard or risk assessment” (European Commission 2003), e.g. answer questions like: Is the tested endpoint relevant for the test species? Does the test exposure scenario exist for the tested substance? How do the tested doses relate to measured or predicted environmental concentrations?
2.6 Pharmaceuticals - use, characteristics, occurrence and effects
The use of pharmaceuticals is increasing worldwide for a variety of reasons; examples include increased world population, an older population, increased wealth, new and improved treatments, unhealthy life styles, market forces, time constraints in the health care sector, and low acceptance for illness and diseases. Inevitably these substances will enter the environment; via excretion and wash off from human and veterinary use, via the manufacturing process, and by improper handling of unused products and a thorough assessment of the environmental impact of these substances must also be performed.

The active pharmaceutical ingredients (APIs) of the medicine have inherent properties that make them potential environmental problems if emitted into the environment. These properties include 1) ability to interact with biological processes, 2) potency, 3) persistency, 4) drug target or cellular specific interaction, 5) ability to bioconcentrate (Santos et al. 2010; Fick et al. 2010; Gunnarsson et al. 2008; Kümmerer 2008).

To date, over 160 API have been found in the aquatic environment. The concentrations vary between ng/L and µg/L, but higher concentrations have also been observed (Cui et al. 2006; Heberer 2002; Kostich & Lazorchak 2008; Larsson et al. 2007; Li et al. 2008; Lin & Tsai 2009; Kümmerer 2008; Santos et al. 2007). Today's wastewater treatment plants are not designed to remove an increased amount of various APIs and their metabolites. The treatment technology in use was initially designed to remove organic matter, nutrients and pathogens, therefore additional treatment steps (or up-stream solutions) are needed to prevent the substances from reaching non-target organisms. Treatment with activated coal and ozonation are two of the suggested methods that has proven to remove APIs (Cuklev et al. 2012; Lundström et al. 2010a; Ikehata et al. 2006; Snyder et al. 2007; Lundström et al. 2010b)

With thousands of chemicals present in the environment, a causal link between a specific substance and an effect in the environment is difficult to establish. Still two examples exist where APIs have caused adverse effect on wild populations. The sex hormone ethinylestradiol has proven to cause feminization of male fish living in waters polluted by wastewater effluents (Purdom et al. 1994; Jobling et al. 2002; Larsson et al. 1999) and the use of the anti-inflammatory drug diclofenac is responsible for a rapid and devastating decline (up to >95%) of three vulture species in Pakistan and India (Oaks et al. 2004; Shultz et al. 2004).

Since 2006 an environmental risk assessment is required for all new marketing authorization applications for APIs (European Medicines Agency 2006). There is no publically available record over how many environmental risk assessments this has resulted in but so far 80 assessments have been performed in Germany (I. Rönnefahrt, UBA, oral presentation at the SETAC conference in Berlin May 22nd).
The open scientific literature also provides examples of environmental risk assessments and risk ranking of pharmaceutical substances. The focus, methods and data used in these studies differ, examples include: risk to aquatic host-parasite relationship and risk of developing resistant strains of parasites (Morley 2009), biological effects in fish (Corcoran et al. 2010), comparison to other groups of chemicals (Damásio et al. 2011), endocrine disrupting substances (Caliman & Gavrilescu 2009), comparison between different pharmaceutical classes (Sanderson et al. 2004), risk characterization in sewage treatment plant discharges (Christensen et al. 2009), risk assessment of the most used substances (Jones et al. 2002), case studies of individual substances (Oakes et al. 2010; Liebig et al. 2010; Küster et al. 2010), risk assessments on a national level (Besse & Garric 2008; Stuer-Lauridsen et al. 2000), marine and estuarine environment (Cooper et al. 2008), and use of quantitative structure activity relationship (QSAR) models (Sanderson et al. 2004).

Roos et al. (2012) compared prioritization methods for ecotoxicity testing and environmental monitoring of APIs. For pharmaceuticals in general it was concluded that a successful prioritization scheme should be based on risk, but for well-studied substances like ethinylestradiol, diclofenac, and fluoxetine methods based on hazard were more successful in priority-ranking these substances correctly. The fish plasma model, where a theoretically estimated fish plasma concentration is compared with a known human therapeutic plasma concentration (Huggett et al. 2003), also showed a high success rate.
2.7 Voluntary initiatives

In the field of environmental risk management voluntary initiatives has been used as an alternative to laws and regulations since the early 1990s (OECD 2003). The general idea is that such initiatives offer “win-win” situations: The regulator achieves the desired results with decreased enforcement costs and the regulated community is offered flexibility in the development and implementation of the initiatives (Daley 2007). In countries with high legal compliance voluntary initiatives are typically used to encourage companies to do more than the mandatory regulation, while it is used as a help to remedy non-compliance in countries where regulatory power is weak and enforcement is low (Blackman & Sisto 2006).

Commonly used criteria for systematic evaluations of voluntary initiatives include: environmental effectiveness; economic efficiency; equity; competitiveness; acceptability; inclusiveness and public participation (OECD 1999; OECD 2003; Sullivan 2005; Cabugueira 2001). The opinions whether voluntary approaches are more efficient than other environmental policy instruments differ. Some argue that they offer a chance to handle environmental problems in a flexible manner at low cost, and at the same time gain consensus with involved stakeholders. Others state that the environmental targets often are modest and set by the industry themselves and therefore provide few environmental improvements beyond what would have been achieved anyway. Voluntary initiatives have also proven to have potential for free-riding: i.e. when companies obtain the benefits of an initiative (e.g. avoiding stricter regulation) while not performing beyond “business as usual” (OECD 2003; Sullivan 2005; Blackman et al. 2006).

In order to increase the performance of environmental initiatives several measures have been suggested, e.g. clearly stating objectives and goals, setting quantifiable targets, development of standardized approaches for reporting, identifying precise requirements for involved partners, identifying reviewing and benchmarking mechanisms, setting appropriate sanctions for failing to meet stated requirements and involving an external reviewer (Sullivan 2005; Croci 2005; de Clercq & Bracke 2005).

Voluntary initiatives cover a variety of different arrangements and according to OECD (1999; 2003; 2003) they can be divided into four general types: (1) unilateral commitments where individual firms or coalitions of firms set up programs, (2) private agreements between polluters and pollutes, (3) public voluntary schemes where participating firms agree to standards developed by public bodies, or (4) negotiated agreements where the authorities and the industry agree on a contract. The Swedish initiative SECIS is a unilateral commitment where pharmaceutical companies on the Swedish market participate in a program set up by the pharmaceutical industry.
2.8 The Swedish Environmental Classification and Information System for pharmaceuticals (SECIS)

The initiative to develop SECIS, a national and voluntary classification system for pharmaceuticals, was taken in 2005 by The Swedish Association for the Pharmaceutical Industry (LIF) (Wennmalm & Gunnarsson 2005; Swedish Medical Products Agency 2004; Mattson 2007). The majority of the pharmaceutical companies on the Swedish market are participating (Mattson 2007). However, participation only means that the initiative is supported, far from all companies have provided environmental data and risk assessments for the pharmaceuticals they produce.

Within SECIS, the pharmaceutical companies provide environmental data and classify their products according to pre-defined criteria and a guidance document for environmental risk assessment of pharmaceuticals (Swedish Association of the Pharmaceutical Industry 2007). The guidance document is developed for the purposes of SECIS, but it is based on the environmental risk assessment guidance documents published by the European Medicines Agency’s (EMA) (European Medicines Agency 2002; European Medicines Agency 2006) and the European Commission (European Commission 2003).

SECIS is based on an assessment of risk, i.e. combining information about the substances’ hazardous properties with information about the estimated environmental concentrations. The PEC/PNEC ratio is used to classify each product according to the criteria into a risk category and corresponding risk phrase. The system has four risk classification categories; insignificant, low, moderate, and high risk. A classification according to the SECIS guidance document furthermore includes an assessment of the pharmaceutical substances’ potential to persist in the environment and to bioaccumulate in organisms (Swedish Association of the Pharmaceutical Industry 2007).

An external consultant reviews the risk assessments and the classification proposal before the information is made public on LIF’s webpage (www.fass.se). Information is presented per product, which implies that several risk assessments and classifications can exist for the same pharmaceutical substance. Neither SECIS nor the EMA guideline suggests or implement any risk management measures to reduce environmental impact. The purpose is only to provide information and according to the EMA guideline the result cannot affect the outcome of an market application (Swedish Association of the Pharmaceutical Industry 2007; European Medicines Agency 2006).

The SECIS risk assessment process differs from the EMA guideline in three general aspects: 1) Within SECIS actual sales figures for the total sales volume of the pharmaceutical substance is used when calculating PEC. EMA, on the other hand, states that PEC should be derived from data on maximum daily dose and percentage of market penetration for a specific product and company. 2) SECIS allows for use of short-term data and non-standard data. This allows for more pharmaceuticals to be assessed in SECIS since long-term data are missing for a majority of the substances. 3) The EMA guideline only applies to new registrations while SECIS applies to all
products on the Swedish market (Swedish Association of the Pharmaceutical Industry 2007; European Medicines Agency 2006).
3. Preview of papers

3.1 Paper I

Paper I presents the development of WikiPharma, an easily accessible, comprehensive and up-to-date database of effects caused by pharmaceuticals on non-target organisms. The paper describes the construction and structure of the database. It also provides guidance for users and presents quantitative characteristics of the database.

The database, available at www.wikipharma.org, currently contains effect data for 151 substances from 212 scientific papers and it is updated yearly. The database only contains published data and the peer-review system is used as a quality control. WikiPharma provides extracted data and bibliographic references but it does not give direct access to the papers from where the data are gathered.

Over 60% of the effect data originates from tests conducted according to standardized methods. The most frequently occurring standard tests are those issued by OECD, the U.S. EPA, and ISO. Tests performed using standard kits like Microtox, Thamnotoxkit and Spirotox are also common. There are over hundred different species represented in the database with *Daphnia magna*, *Vibrio fischeri* and *Pseudokirchneriella subcapitata* being the most frequently used. More than half of the compiled data represent short-term exposure and acute responses. The two pharmaceutical groups showing lowest median effect concentrations are sex hormones and antibiotics.

This concentration of ecotoxicity data for pharmaceuticals compiled in one single place can be of help when assessing potential risks of pharmaceutical ingredients in the environment, as well as for identifying data gaps. Since the publication of the database in 2009, we know from email correspondence and webpage statistics that it has been used by researchers, risk assessors and regulators in both Europe and North America.
3.2 Paper II
This paper reports from an empirical investigation of the motivations, intentions and expectations underlying the development and implementation of SECIS. A questionnaire, containing 19 questions, was sent out to the 29 persons involved in the start-up process of the classification system. The questions were either composed as open-ended questions, or as statements asking the respondents to indicate agreement on a five point scale.

The response rate for the survey was 72% (n=21). The answers were coded and the data analyzed using descriptive statistics. The answers from the respondents were analyzed according to affiliation and the respondents were therefore divided into three different groups: persons affiliated with the pharmaceutical industry, persons affiliated with governmental agencies and persons with other affiliations.

The results from this inquiry indicates that the decision to implement a classification and information system for pharmaceuticals was the result of a combination of several driving forces, mainly political pressure and a possibility to increase the industries' goodwill, while at the same time keeping the process under the industries' control. The expected possible effects of the system, other than increased goodwill, are according to this survey assumed to be low. The system offers little guidance for end-users in the substitution of one pharmaceutical for another. One possible reason for this could be that LIF needs to observe the interests of all its members and should not affect competition. The affiliation of the involved actors correlates to how these actors view and value the system, but this has not hampered the collaborative process to develop and implement it.

There are three general conclusions that can be drawn from the results of this study. First, the likelihood that the system will have an impact on prescriptions and use of pharmaceuticals are questioned since the system provides only limited guidance to end-users that want to explore the possibility to substitute one pharmaceutical for another. Second, if the owner of the system, in this case LIF, cannot provide information that will enable comparisons between different pharmaceutical substances due to their role as a representative for all their member companies, then the efficiency of the system might benefit from being hosted by a more independent body. Third, there seems to be a common perception among the respondents that developing an environmental classification system has a value in itself.

In summary, SECIS is an ambitious initiative that is unique in its kind. The system has great potential to contribute to scientifically well-motivated decisions on risk management. However, the results from this investigation suggest that there is room for further improvements of the system.
3.3 Paper III

In this paper the accuracy and the consistency of the environmental risk assessments conducted within SECIS were evaluated. Three main aspects were considered:

(1) Were the risk assessments and classifications made within SECIS accurate in relation to the SECIS guidance document: How were data selected for the risk assessments made so far? What types of data were included (e.g. standard vs. non-standard test data), what were the main sources of the data, and how complete were the data-sets?

(2) If the data-sets did not include all available and relevant data; would a supplemented data-set alter the PEC/PNEC ratio and the risk classification?

(3) How consistent were the risk assessments when different companies had assessed the same pharmaceutical substance?

For these analyses we selected all the pharmaceutical substances classified by at least one company and for which ecotoxicity data could be found in the open scientific literature. This resulted in a set of 36 substances and 48 assessments that were scrutinized (the number of assessments is larger than the number of substances since several companies are producing products with the same pharmaceutical substance).

The results showed:

(1) Based on the review of the 48 assessments above, alternative PEC/PNEC ratios could be determined for six substances. Five of these were results of altering the assessment factor in assessments where long-term data had been used for extrapolation to short-term exposure or vice versa. In the sixth assessment ecotoxicity data from blue-green algae was not used in an assessment of an antibiotic substance, even though it is specifically acknowledged in the EMA guideline and by the reviewer to be an important species when it comes to ecotoxicity testing of antibiotics. The alternative PEC/PNECs for these six substances would, if implemented, lead to four new risk classifications. In three cases the substance would have been classified into a higher category and in one case a lower.

The scrutiny also brought about five assessments which, in SECIS, had resulted in risk classifications even though the data provided was not sufficient for classification according to the guidance document.

(2) Additional data which altered the PEC/PNEC ratio were available for 18 of the 48 risk assessments. These data were reported in 14 studies. Nine of the 14 studies reported results from standard tests. Half of the studies were regarded as long-term studies and the other half as short-term.
The use of additional data resulted in a new classification category for 13 of the assessments. Ten of the 13 assessments resulted in a higher classification category, for one substance the PEC/PNEC ratio increased with three orders of magnitude and the risk classification changed two steps.

(3) Seven of the 36 pharmaceutical substances selected for evaluation in this study were assessed and classified by more than one risk assessor. In two of the seven cases, different producers classified the same substance into different classification categories. This is a rather high percentage of the substances that were eligible for this analysis (2/7 or 29%), but only a small part of the substances that have been classified by at least two companies (2/18 or 11%) and only a tiny part of the total number of classified substances in SECIS in April 2009 (2/179 or 1%). To what extent our selection of pharmaceutical substances to evaluate is representative for the rest of the classified substances is not known.

Finally some recommendations to improve the classification system’s accuracy, consistency, transparency, data reporting and data selection are proposed. The recommendations include clarifying concepts and instructions in the guidance document, introduction of a standardized way of reporting data to the website, and promotion of use of non-standard test data when considered the most relevant.


3.4 Paper IV

The environmental risk assessment process within SECIS differs from the EMA method by allowing for use of non-standard ecotoxicity data. However, even though the use is acknowledged by SECIS, information on how to evaluate this kind of data is not included in the SECIS guideline. The overall aim of this study was to investigate if non-standard ecotoxicity data could be used systematically in environmental risk assessments of pharmaceutical substances. This has been done by 1) an evaluation of the usefulness of existing proposed criteria for reliability evaluation of test data, and 2) an investigation whether recently published non-standard ecotoxicity studies from the open scientific literature fulfill proposed reliability criteria.

Four reliability evaluation methods were found when searching the scientific literature: Klimisch et al. (1997), Hobbs et al. (2005), Durda and Preziosi (2000) and Schneider et al. (2009). Nine studies published in the peer-reviewed literature containing non-standard test data were used as examples and evaluated according to the four methods, this resulted in 36 evaluations.

The evaluation methods differ in their scope, user friendliness, and how criteria are weighted and summarized. Klimisch et al. (1997) has unclear criteria and not enough guidance for the evaluator which affects the user friendliness. Durda and Preziosi (2000) have a broad scope and is user friendly but using the method resulted in the evaluation result “not reliable” for all nine studies that were evaluated. Hobbs et al. (2005) also has a broad scope and is user friendly. More studies were accepted when using this method, but none of the studies were categorized as having “high quality”. Schneider et al. (2009) also has a broad scope and is user friendly. This method differentiated the selected studies more than the other methods, i.e. some studies were considered “not reliable” and others were considered “reliable without restrictions”. Using the four methods lead to an agreed evaluation result for only two studies. The evaluation results differed by one step for five studies and by two steps for two studies.

Reliability evaluation methods with more structured criteria, as an alternative to case by case evaluations, can help risk assessors from both regulatory agencies and pharmaceutical companies to use non-standard test data. However, the choice of evaluation method can affect the outcome of the reliability evaluation. Using the method described by Durda and Preziosi (2000) might result in a situation were no data from the open scientific literature can be used in risk assessment, whereas the method described by Hobbs et al. (2005) opens up for use of more data.

Overall the evaluation of non-standard test data resulted in a low number of studies with acceptable reliability. Only 15 of the 37 evaluations (42%) resulted in “acceptable quality”, “reliable with restrictions”, or “reliable without restrictions”. Many of the aspects considered important in the different evaluation methods are left out in the reporting by the authors of the selected studies. Examples of aspects that are missing are information about the controls, results from statistical evaluations, dose-response correlations, tested concentrations and clear descriptions of the test environment. Possible reasons for this could be lack of knowledge about what kind of data that are
needed in risk assessment and that the reports from the scientific studies were not written with risk assessment in mind. Using a checklist (e.g. criteria from an evaluation method) as guidance could improve the reporting and regulatory impact of the studies by guaranteeing that all relevant aspects are reported. Introducing checklists in the peer-review process would be an alternative and additional way to achieve the same goal.
3.5 Paper V

In this paper a new reliability and relevance evaluation method was presented. Since ecotoxicity studies of high reliability is needed in risk assessments it was important to combine the researcher’s and the regulator’s perspective in the development. The method is therefore the result of a collaboration between MistraPharma researchers and the pharmaceutical unit at the German Federal Environment Agency. Ecotoxicologists from Brixham Environmental Laboratory in United Kingdom and AstraZeneca in Sweden also contributed with their expertise.

The 62 reliability criteria are based on the four evaluation methods presented in paper IV, together with the OECD reporting requirements for chronic testing on algae, daphnids and fish. The criteria are divided into ten different categories: purpose and endpoint; protocol; test compound; dosing system; test organism; controls; test environment; statistical design; biological effect and other considerations. Twelve relevance criteria were also added since this aspect of evaluations has so far been performed without or with very few criteria. A more structured reliability and relevance evaluation is needed to reach the goal of transparent, robust and predictable risk assessments.

In order to clarify how data with a varying degree of reliability and relevance can contribute to the process of decision making, a model for how to summarize the evaluations was developed. Both standard and non-standard data could be evaluated using this method. Intended users of the method are both risk assessors and researchers performing ecotoxicological experiments, offering to bridge the gap between the regulators’ and the scientist’s needs and procedures of work. Furthermore, the criteria can be used for education purposes and in the peer-review process to increase the reliability of scientific papers.
4. Discussion

Evaluation of the current situation is a prerequisite for future improvements. This thesis evaluates steps in the risk assessment process concerning use of ecotoxicity data, and also a risk management initiative from the pharmaceutical industry in Sweden. The included aspects focused on; user-friendliness, transparency, accuracy, consistency, data gathering, data selection, data reporting, and data evaluation. All five papers in this thesis concern environmental risk assessment of pharmaceuticals but the presented ideas and methods can be applied to other types of chemicals and also to health risk assessment.

There are three overall conclusions from this thesis:

(1) **A closer collaboration between researchers and regulators is needed to better utilize the generated knowledge and data from the academia.**

Ecotoxicity data from the open scientific literature is not used to the extent it could be in regulatory risk assessment. This is due to regulators’ preference for standard data and since research studies in the open scientific literature is reported in a way that can affect the reliability and the user-friendliness. Researchers from academia, regulators from authorities, and representatives from industry may never agree on a common goal for their work but it might not be necessary if data sharing can be secured with an agreement regarding how to report data.

Some argue that academic research is not suitable for risk assessment since it is curiosity-driven. This is true for some studies but not for all. As long as the data are of sufficient reliability, and relevant for the particular risk assessment, the driving force behind the generation of the data is of little importance for the adequacy of the study. There are no contradictions between producing curiosity-driven innovative research and producing research of high reliability. In fact, the latter is a cornerstone in the scientific work since it is closely connected to reproducibility. It is therefore alarming that the overall reliability of the evaluated data in Paper IV is low. Whether this is due to the design and performance of the study or the reporting of it is not clear. A combination of the two is also possible.

Another argument against use of academic research in risk assessment is that investigated endpoints are not always connected to adverse effects and therefore not suitable. However, in the WikiPharma database the majority of the test endpoints are the same as for standard tests. This also applies for choice of test species and exposure time (Paper I). For research where non-traditional endpoints have been used there are methods developed for how to determine causality. The adverse outcome pathway is a conceptual framework that portrays the linkage between a molecular event and an effect at population level (Ankley et al. 2010; Kramer et al. 2011), and U.S. EPA has developed a standardized decision tool for causal analysis that helps assessors identify likely causes of impairments (Suter et al. 2010).
To close the gap between academic research and regulatory risk assessment several things need to be done. The reliability of ecotoxicity studies has to improve to meet the regulatory needs. This could be achieved by researchers using reporting criteria as a guide when designing, performing and reporting studies. Editors of scientific journals can also contribute towards the aim by introducing evaluation guidelines for reviewers and reporting requirements for researchers.

More knowledge does not necessarily make risk assessment easier since data can point in different directions. The relationship between the regulator and the chemical industry is already established through the regulatory process, and this provides the regulator with standard data for which evaluation experience and criteria exists. Inclusion of data from academic research can therefore be a burden for the regulator who works under time pressure; contacts have to be established, databases have to be searched, and innovative research methods have to be evaluated. To promote inclusion collaborations between researchers and regulators could be established through workshops, method developments, and risk assessment guidelines that instruct industry to perform a thorough data gathering from the academic research.

(2) **A risk assessment is the result of the selected data and selected methods used in the process and therefore a transparent procedure is necessary to enable external review.**

An important starting point for the risk assessment is to gather all available data instead of settling for data that only fulfill the basic criteria of three separate tests from three different trophic levels. In paper III it is concluded that the pharmaceutical companies have had different strategies regarding data gathering, some have included publically available data while others have settled for their own data. Data from the open scientific literature is generally not used to the extent it could be in the risk assessments within SECIS.

Using reliable and relevant test data in risk assessments is considered fundamental. Still, it is seldom mentioned in risk assessment documents how ecotoxicity data were evaluated for reliability and relevance. Paper III confirms that none of the evaluated risk assessments specified how the data evaluation was conducted and Paper IV concludes that the choice of evaluation methods can have an effect on the outcome of a reliability evaluation.

Neither the EMA legislation (European Medicines Agency 2006) nor the SECIS guidance document (Swedish Association of the Pharmaceutical Industry 2007) clarifies how data evaluation should be conducted. In contrast, both the TGD (European Commission 2003) and the REACH guidance document (European Chemicals Agency 2008) address how to evaluate the reliability and relevance of test data by providing a framework with some guidance to risk assessors. There is also several evaluation methods presented in the open scientific literature. However, with the exception of Klimisch et al. (1997), these have so far had low use.
(3) In its current form, SECIS contributes to data availability and transparency but not to risk reduction.

SECIS has contributed to the general discussion about pharmaceuticals’ effect on the environment both on a national and international level. It has also made data, which previously were considered company secrets, publicly available. However, to be an effective risk reduction tool this is not sufficient. Examples of evaluation variables for voluntary initiatives include; purpose, third-party review, transparency, sanctions, participation rate, and level of compliance (Paper II).

SECIS has the purpose to provide available environmental information about all pharmaceutical substances on the Swedish market by 2010. The information provided does not differ much from the requirements in the EMA legislation but it applies to all pharmaceuticals instead of only new market applications. The host of an environmental initiative can affect and limit the user-friendliness, and consequently also the environmental effectiveness, of the initiative by defining a purpose and a structure that limit the user’s possibility to act on the information provided. In SECIS the information is presented in a way that does not encourage further actions such as substituting one pharmaceutical for another (Paper II).

The results from Paper III indicate that a third-party reviewer is not a guarantee for transparent and accurate risk assessments. Failing to follow the guidelines without justifying the exceptions or not explaining why a decision was made can affect the credibility of the initiative. Furthermore, there are no sanctions, economic or non-economic, for non-compliance within SECIS. Sanctions are considered important for the effectiveness of voluntary initiatives (Cabugueira 2001). When guideline recommendations give room for interpretation it can result in inconsistent risk assessments with varied content (Paper II and III). By using a standardized reporting format the transparency and comparability of the risk assessments could increase.

The participation rate in SECIS is relatively high; this could be seen as a signal of success according to Lange (2009) and Southworth (2009). However, the variation of the content of the risk assessment suggests that the level of compliance differs (Paper III). Competitive pressure, the environmental commitment in the upper management of the company, regulatory pressure and the size of the company are examples of factors that could affect the level of compliance (Wu 2009).
Recommendations for improvement of SECIS

- Facilitate substitution; the system should be adjusted so that the environmental risk assessments for pharmaceutical substances with the same pharmacological effect can be more easily compared.

- Investigate the underlying factors to the current inconsistency in the risk classifications performed by different companies and make required adjustments in routines and the guidance document.

- Define central concepts such as “short-term” and “long–term” data in the SECIS guidance document.

- Clarify which aspects of the SECIS guidance document that is mandatory or recommended.

- Promote the use of non-standard test data when these are considered the most relevant. It should be clearly stated that the aim is to always use a relevant and sufficiently sensitive test method.

- Clarify how non-standard test data can be evaluated when it comes to relevance and reliability.

- Create a reporting format that helps companies report all relevant information to the website in a harmonized way.

- Require that the open scientific literature as a rule is searched for relevant data and, when available, these data should be taken into consideration in the risk assessment.

- Require that every action that is not supported by the SECIS guidance document should be transparently justified.
5. Future work
Together with regulators and risk assessors from several European governmental authorities and consulting firms I will continue the development of the reliability and relevance evaluation scheme. In the new version plausibility and weight of evidence is added so that whole data sets can be evaluated for use in environmental risk assessments for chemical substances. A ring test with risk assessors from academia, industry, consultant firms and governmental authorities is planned since the experiences and demands from users are important in the design and function of the evaluation method.

Future work also involves the development of a reliability and relevance evaluation scheme for ecotoxicity data for nanoparticles. Experiences from paper IV and V will be used as the basis for this scheme and we will collaborate with expertise working with ecotoxicology of nanoparticles to create a tool that is user friendly and helpful for risk assessors and researchers. The aim of the work is to enable use of data from the open scientific literature in regulatory risk assessments.

We have also continued to use the WikiPharma database for a comparison of ecotoxicity effect values. The aim of this study is to investigate how the current regulatory risk assessment method can be improved.
References


Commission Regulation (EC). No 1488/94 on Risk Assessment for Existing Substances, Part II.


