Observations on the final Luxembourg Environment Agency report on the “Refinement of Risk Assessment for use in Socioeconomic Impact Assessment under REACH”

The final report was presented at the workshop. It was thought to be helpful to get observations after the workshop to elaborate the issues raised in the report. One observation was received and is given below.

Observation

First of all, we would like to thank the Luxembourg Environment Agency and the contractors for producing a thought-provoking report. The communication of ‘risk’ in terms that can be used for socio-economic analysis has been identified as an important concept that will help decision-making become more transparent, and this report is a useful contribution to the debate.

Two types of environmental ‘risk’ can be dealt with under REACH. The first concerns threshold toxicity (including endocrine disrupters), for which measures of toxic impact as described in the report could be useful. The second concerns PBT/vPvB substances, for which a threshold is assumed not to exist (at least in terms of the policy goal to reduce and eliminate emissions, based on the precautionary principle). This second group of substances is likely to be the subject of the majority of environmental restriction proposals or authorisation applications. Comparing their impacts with alternative substances with threshold effects will be problematic, and this is not really addressed by the report.

Section 3.1 Life Cycle Assessment (p. 21-)

Although this section summarises a number of models, it seems to lack a critical appraisal of their strengths and weaknesses, and the ways in which they differ from a methodological point of view. The problems might partly arise due to limitations in the supporting literature for the individual models, but some of the basic concepts are not explained particularly well in our view, and there is a lot of jargon. For example:

- A number of terms are used without an adequate explanation of what they actually mean or how they are calculated (e.g. ‘eco-cost’ (p. 22), ‘mean measured response’ (p. 23), ‘PAF’ (p. 23)).
- For the CML model, how is the impact of a kilogram of dichlorobenzene estimated, and what does the apparent impact of this substance actually mean in practice? Has it been benchmarked against a real impact, or is it solely based on the intrinsic toxicity data (and if the latter, are the data reliable)?
- What are the input data needs for each model?
- How are input data ‘translated’ to an impact index?
- How are environmental fate and behaviour taken into account?
- The jumble of different impact terms presented in Figure 3.1 is a good illustration of the problem – without further explanation they are pretty meaningless.
- The example provided in Table 3.1 is also confusing, without information on the input terms and a definition of each impact term. The ‘EDIP 1997 Ecotoxicity Water Chronic’ results imply that lead is of higher concern that chromium (VI), which contradicts the output from some of the other models. The lack of seawater data for lead in eight of the fourteen models also demonstrates the drawbacks rather well. There is no information about the reliability of the underlying toxicity data that has been used either.
- From the discussion, it appears that monochlorobenzene might be between 5 and 20 times less harmful than trichlorobenzene in seawater. This could be deduced from a direct comparison of the ecotoxicity information available for each substance (which is recognised on p. 56). Whilst this is a simplistic analysis, it could be sufficient for REACH purposes (avoiding the need for development or interpretation of LCA models).

- The models still do not seem able to give a true indication of environmental impact. In effect, they appear to provide a normalised risk characterisation ratio (albeit expressed in different terms).

The conclusions make it clear that many of these models are ‘black boxes’, and this lack of transparency does not inspire confidence in their use. Since the underlying basis for each model is unclear, it would not be a surprise to find that they all give similar results if they are based on similar methodologies.

From experience, the life cycle of many substances can be quite complicated, and we wonder how easy it is to adapt the models to this type of situation. Nonylphenol ethoxylates (NPE) provide a good example of multiple uses, and in some cases, an alternative substance might be suitable as a replacement for a specific application but not others. It is not clear whether the data available from Chemical Safety Assessments will be sufficiently robust to be used as input data for the LCA models (in the case of NPE, most of the exposure scenarios in the ESR assessment – which led to restrictions! – were based on default emission factors only).

It is clear that a substance must already be incorporated into the model for an assessment to be performed, and that addition of further substances relies on “in-depth experience”, with different data needs for different models, and is “unlikely to be a trivial exercise”. There is no indication as to which substances are present in the models, although the fact that only five of nineteen fairly well known ‘problematic’ substances are included in the GaBi 4 software is a concern. Given that a number of substances might need to be assessed as alternatives, this is likely to be the major limiting feature for REACH purposes and LCA is probably not going to be used by regulators or industry very often.

Overall, we agree with the conclusions of this section.

Section 3.2 Exposure and effects based “proxies” (p. 34-)

The concept of ‘volume of affected media’ is interesting, but we wonder how useful it actually is. As the report recognises, such a volume is still only a *notional* rather than actual indicator of harm – in fact it is directly proportional to the PNEC or SSD output, so is simply a reflection of the available ecotoxicity data. In reality, chemical discharges are dissipated by partitioning, degradation and dilution processes, and bioavailability may also vary under different circumstances, so the *actual* volume of medium affected may be very different. In our view, it would be better to keep things simple. <u>Relative</u> comparisons can still be made using just PNECs (or PEC/PNEC ratios) without adjustment (e.g. the level of impact may be n times higher for substance X than substance Y, without differentiating what that impact might be). (Even though the size of assessment factors may vary between substances with different amounts of data, this simply reflects relative uncertainty.) One aspect of impact that is not considered this way is its duration (e.g. comparing a persistent versus degradable substance, such as nonylphenol versus a long chain alcohol). This is addressed to some extent by the probabilistic approach that is discussed later in the report (since it is a factor in the exposure term) but not PNECs, SSDs or volumes of affected media.

Whilst we agree that species sensitivity distributions (SSDs) are a useful way of expressing ecotoxicity data, we may need to be wary about reading too much into them. They are based on a number of explicit assumptions, and are bound by both statistical and practical constraints. Deriving an SSD and consequently a ‘potentially affected fraction’ (PAF) based on three data points (representing just three species) could be highly misleading. That is why the REACH technical guidance specifies a minimum number of taxonomic groups – though apparently arbitrary, it at least ensures that the SSD takes account of data for a reasonable spread of organisms. The PAF is just a function of the selected data distribution, not a real measure of impact. In addition, the relative position of different species within the SSD is ignored by the PAF. For example, if a particularly sensitive group were apparent (e.g. algae for a substance with herbicidal properties), then protecting 95% of all species might still lead to problems if all the algae are in the other 5%. The problem is we do not really know how the environment would respond to impacts on a certain fraction of species. We should also recall that assessment factors are meant to account for uncertainties such as intra- and inter-laboratory variation of toxicity data and single stress (laboratory) to multiple stressor (field impact) extrapolation. Such factors are not taken into account in an SSD. SSDs also lose information on dose-response of individual species (although this is also a feature of PNECs too).

[The expression of dilution on p. 36 is not quite correct, although for the purposes of the report it does not make much difference (dilution is equal to the sum of effluent flow and receiving water flow, divided by the effluent flow, so the default river flow is 18,000 m^3/d).]

The use of equilibrium partitioning to estimate terrestrial or sediment toxicity in the absence of suitable test organism data is a screening approach used to decide on the need for further data, not really a surrogate measure of toxicity as such. Therefore it is of questionable value to use such data in any probabilistic analysis.

Section 3.3 Probabilistic risk assessment (p. 39-)

We welcome the approach described in this section, *where the available data allow such calculations to be made*. Again, the output is not a true measure of impact, but it at least allows something to be said about relative confidence in different levels of impact, which is not provided by PEC/PNEC ratios.

Section 3.5 Results (p. 44-)

The report recognises that for many of the candidate substances, insufficient toxicity data were available to make the required calculations. Whilst REACH may be expected to improve this situation, it is unlikely to provide rich (i.e. more than three taxonomic group) data sets for the vast majority of substances.

We have not reviewed the examples in detail. We note that the HC5 for freshwater sediment for HBCDD ranges from 0 to 14.5 mg/kg. With such a wide variation, is a statistical approach justified? (N.B. the volume of affected medium is 'N/A' to > 40 M m^3/d!)

Section 4 "Rapid" Method Case Studies (p. 51-)

The approach focuses on local emissions and therefore neglects regional background concentrations. Whist this is acceptable for the purposes of the report, there are occasions when the background concentration makes a significant contribution to the level of risk (as was the case for NP).

We understand that time constraints and data availability will have limited the number of comparisons that could be made in the case studies. However, it may be helpful to note that the comparison of risks/impacts of different substances needs to take account of all relevant media, as well as human health. For example, some of the study substances appear to have impacts on WWTP micro-organisms, but this is not considered in the report. Also, focus on freshwater (as a volume of medium affected) may miss potential impacts on sediments, which may be more important for many substances with high adsorption potential.

Whilst the calculations based on acute data sets help to illustrate the points, the focus of risk is generally long-term impact, which is protective of acute impacts.

The examples based on QSAR data are of questionable value. The use of QSARs needs to be justified on a case-by-case basis, and the mixed picture that is presented is not particularly surprising. In the future, it might be possible to make extrapolations to increase the size of the data set using tools like the US EPA's Web-ICE software, and we understand that an Ecetoc project is looking into this further. However, this is likely to rely on the mode of action of the particular substance, which may be different between alternative substances. It also adds further uncertainty in some respects.

We noted some inconsistencies in the data tables (and could provide some additional references), but these are not vital to the overall report. However, the comparison of NP data with alcohol ethoxylates is misleading – the identified risk was based on NP as a persistent degradation product of ethoxylate use. It would therefore be more appropriate to consider data for the alcohol (assuming this does not co-degrade with the ethoxylate chain). This example also raises an interesting additional aspect. One of the main concerns about NP is its potential to elicit oestrogenic effects, which alcohols do not possess. This is not considered by the SSD approach as such. Even if the two substances were of equal toxicity, presumably the endocrine-related nature of the NP effects would need to be recognised in the analysis.

Similarly, the main focus of concern for short-chain chlorinated paraffins is its PBT properties. Long chain chlorinated paraffins may exhibit toxicity, but does not have a PBT profile of concern. How can these two substances be compared?

Section 5.3 Applying ecosystem services within SEA of restrictions under REACH (p. 111-)

Figures 5.1 and 5.2 (causal chains) are not helpful in our view. They provide a qualitative indication of possible impacts in different media, but these remain notional and are unquantifiable. The same sort of diagram could be drawn for many other substances. We therefore agree with the conclusions that they provide no additional analytical benefit.

[The remainder of the comments on this section were provided by an advisor.]

The major problem is that the report does not pick up the link between the final welfare end points and ecosystem analysis. At the moment there is a disconnect between the valuation discussion in Section 5.3 and ecosystem analysis in Sections 5.1 and 5.3. A discussion of this type has to distinguish between final (i.e. welfare relevant) and intermediate ecosystem services, which though potentially relevant to final services, serve no immediate welfare function. The earlier analysis in the paper seems to provide a discussion of proxy measures and a description of intermediate services.

Ecosystem service analysis is not just about causal chains, but about the final welfare outcomes resulting from the effect of allowing a substance into the environment. This rather elementary point means that one cannot spot the value that the analytical approach has in illuminating the concept of ecosystem resilience. The attitude one might have to the risk associated with this type of ecosystem outcome is also not explored - the default of risk neutrality probably does not apply to the sort of risks associated with hazardous chemicals. These risks map into uncertainty and are likely to be both pervasive in many cases (particularly to certain members of the population) and potentially non-marginal in effect. This means that some or all of the population are likely to be risk averse to some hazardous chemical-associated risks and uncertainty.

The essential point in the discussion of ecosystem valuation is whether the aggregate of the welfare end points associated with the affected final ecosystem service that result from an authorisation or restriction captures the whole effect of the proposal. Or is it the case that chemical-by-chemical decisions mean that we are missing something from the valuation analysis? Most of these effects are likely to be unknowns rather than risks. What is the attitude of the paying public likely to be to allowing something to happen which is unknown, but likely to be adverse and potentially strongly so?

In summary, the authors of the report have not understood the ecosystem valuation approach. The attitude of the paying public towards the build up of hazardous chemicals should be explored further.

Section 7 Recommendations (p. 117)

We believe that probabilistic risk assessment approaches are a promising area for further guidance to be developed, along with suitable freely available tools. The likelihood of an impact is an important concept, although it is still not possible to ascribe the impact to any particular ecosystem service as such (other than in a broad sense). In addition, as pointed out at the workshop, lack of data will still be a limiting factor for some time to come.

Exposure-based proxies are not so useful in our view, given the comments above. We fear that they could be mis-interpreted as some sort of real measure by economists unfamiliar with their derivation, and it is simpler to make comparisons between PNECs or risk characterisation ratios for the purpose of making simple statements about relative risk.

We also have doubts about whether LCA models can be developed sufficiently to make them useful for REACH purposes, at least in the short term. We would need to be convinced about the additional benefit they would bring, along with better transparency about what the model outputs mean and how they are derived.

S Dungey
Chemicals Assessment Unit
Environment Agency

24 March 2011