Abstract of a Danish EPA working paper in relation to establishment of POP criteria under The Stockholm Convention presented to the participants of Criteria Evaluation Group II Meeting, Vienna, June 14th to18th, 1999

Use of QSARs for Selection of POPs

Abstract.

POPs(Persistent Organic Pollutants) or PBTs (Persistent Bioaccumulative Toxics) have been addressed in regulatory contexts at various national and regional scales (e.g. Marine Conventions). Recently POPs have also been addressed at a global scale (UNEP) to stop or significantly reduce continued environmental release. Different criteria for identification of POPs/PBTs are currently being discussed in relation to the level of concern and geographic scale of environmental transport. Available experimental data on chemicals are generally scarce relative to the number of chemicals and endpoints of concern such as persistency, mobility, bioaccumulative and toxic potential. Experimental data on candidate POPs / PBTs are also relatively scarce even though a lot of experimental data are available on a limited number of well-known POPs / PBTs of global concern.

Persistence in the mobile environmental media air and water is used in two ways. It is used as a criterion for the potential of long-range environmental transport in the atmosphere, by rivers or by currents in the oceans. It is also used as a criterion for the persistency in remote areas after long-range environmental transport has taken place.

Predictive QSAR models may, like experimental test data, be employed for initial selection of POP candidates. Such QSAR based approaches may supplement approaches for identification of candidate POPs based on experimental data and for setting priorities for a thorough evaluation and / or testing.

QSAR approaches for initial selection of POPs were conducted on 166.075 chemicals (including approximately half of all chemicals registered on the EU market) using nine different sets of possible selection criteria of POP candidates, which were comparable with the draft UNEP POP criteria referring to experimental data. The selection criteria employed differed in that six of them employed different possible selection criteria relating to bioaccumulation (i.e. use of log Kow- or BCF- models). Three further sets of selection criteria focussed on selection of POPs with a significant potential for long-range atmospheric transport (i.e. criteria for volatility and atmospheric stability were also employed).

The outcome of these selection procedures was presented, so that the number of chemicals at each separate selection step can be identified in relation to bioaccumulation and to persistence in water, soil, sediment and air. The exercise contributed to the discussions about which criteria should be established for POPs of global concern in the framework of UNEP, e.g. what would be the appropriate POP criteria triggers for the bioconcentration factor in aquatic organisms and for the half-life in water, soil and sediment.

The QSAR based exercise indicated that number of potentially globally concerning POPs / PBTs of widespread use is relatively small regardless of which of the QSAR approaches for environmental persistence and bioaccumulation potential that were employed. Other QSAR methods regarding baseline (minimum) toxicity for fish toxicity and mammalian toxicity (mutagenicity, cancer and

reproductive toxicity) were employed on QSAR selected candidate POPs. These showed that all identified candidate POPs may be quite toxic to fish and that a fairly high percentage may have a mutagenic / carcinogenic / reproductive toxicity potential.

Finally it was discussed that regardless of whether data are generated by employing QSARs or obtained from standard or simulation types of experimental tests, it is necessary to extrapolate and / or interpret such data in relation to the POPs criteria.

Introduction

Policy issues and concerns

The issue of POPs (Persistent Organic Pollutants) is defined and described in a variety of recent policy documents, scientific papers and working documents due to the increasing attention this type of organic chemicals has received also on an international regulatory scale (AMAP, 1996, Webster et al, 1998; Waina F. & D. Mackay, 1996). The concern caused by POPs has been recognised explicitly for POPs of global concern in relation to the development of international conventions and protocols such as the recently adopted UN ECE in relation to air pollution with POPs (i.e. the Protocol concerning Long Range Transboundary Atmospheric Transport of POPs, ECE LRTAP POPs) and in the context of the recently started international activities under UNEP in relation to establishment of an international legally binding instrument for global action against the continued release of POPs. The same type of concern has however for several years also been expressed in relation to international conventions, declarations and actions for protection of especially the marine environment. These refer typically to control of releases, losses and discharges of persistent, bioaccumulative and / or harmful or toxic chemicals, PBTs (Persistent Bioaccumulative Toxics, c.f. e.g. national programmes on chemicals like the Government of Canada, 1995; The US EPA PBT Program (cf. US EPA homepage on this programme), The Danish EPA (1996, 1997 1998 and 1999, cf. www.mst.dk), conventions and declarations (e.g. OSPARCOM, HELCOM, The Barcelona Convention, North Sea Declaration).

The concern caused by persistent and bioaccumulative chemicals e.g. for the marine environment or other presumed or realised especially vulnerable remote natural environments have also been raised in relation to ongoing development and implementation of risk assessment methodologies e.g. in the EU. In these cases reference has been made to the current limitations of the employed risk assessment paradigm and whether this fully encompasses the concern for marine pollutants, PBTs or POPs (e.g. Ahrens A., 1999). Concern has also been expressed on issues like:

- Lack of data and concept for assessment of the *long term* impact of PBT, because of limitations of the normally employed environmental risk assessment concepts, which for example do not include assessment of life time exposure to chemicals, the *total* impact of substances with the same kind of toxic mechanisms, with possible interactions between simultaneously occurring toxic chemicals or the possible impact of the chemicals at steady state, if current environmental release continues - the latter may furthermore include slow environmental release which is difficult to quantify, because this release includes substances accumulated in the technosphere, i.e. from products, goods and constructions both during and after use

- Lack of relevant and comprehensive data and evaluation concepts for integration of the variation in degradation capacity of the natural environment in different climatic regions caused by the huge natural differences of these environment (e.g. fresh water rivers in Central Europe compared with frozen lakes in the North or the Subarctic and Arctic Ocean)

- General lack of relevant and comprehensive data end evaluation concepts for the evaluation of bioaccumulation, including transfer through the food webs, which may result in build up of high concentrations of POPs especially in marine top predators of the long food webs of the marine environment.

- Generally lack of monitoring data, concepts and models concerning long-range environmental transport and distribution of chemicals. Monitoring data is generally only available on few chemicals including some well-known POPs. Concepts and data for modelling transport of POPs on airborne particles and via currents of the ocean are furthermore premature.

- That the current knowledge for the vast majority of chemicals including many PBTs generally is limited as regards possible harmful chronic and delayed toxic effects

- That, contrary to other types of hazardous chemicals, harmful effects caused by POPs may also be caused in remote environments, which are not well known, and where environmental fate processes and recovery of biota after harmful impacts may be slower and effects may be more pronounced

- That for POPs not only more widespread harmful effects may occur, these effects may also persist long time after control measures have decreased or removed environmental release

- That human populations remote from where POPs are utilised, may only be confronted with the harmful effects and not with the benefits of the use of these chemicals. The human populations in question furthermore are typically highly dependent on fishing and hunting and thereby extremely likely to be highly exposed to persistent bioaccumulative chemicals in game animals

Because of these characteristics and because POPs therefore are very difficult or impossible to control by other measures than ban or restrictions which significantly reduces the environmental releases, a special *precautionary approach* is needed. Experience in the past has shown that surprising and concerning environmentally harmful effects of such chemicals may later be revealed even though these chemicals at the time of marketing and use generally were regarded beneficial and harmless to man and the environment (e.g. PCBs, DDT and certain other persistent and bioaccumulative pesticides).

POP selection criteria.

During the development of the LRTAP POPs protocol various screening exercises were done for selection of POP candidates based on a rough evaluation of selected intrinsic properties which are thought to be for special relevance for the ability of substances to undergo long range atmospheric transport. Based on these exercises and separate discussions on establishment of screening criteria to be used for selection of POPs both in the context of the LRTAP protocol and the future UNEP convention, certain selection criteria have been established or suggested. (Cf. e.g. "Executive Body Decision 1998/2 on Information to be submitted and a procedure for adding Substances to Annexes I, II or II to the POPs protocol" (EB/AIR/WG.5/52, Annex II and other documents at the UNEP POP homepage on the internet); US EPA PBT Program internet homepage; Government of Canada, 1995; Cowan (1998); Rodan et al (1998)). The screening criteria for identification of POPs include use of evidence from monitoring at remote places, but concern mainly certain intrinsic environmental fate related properties such as persistency in water, soil and / or sediment and bioaccumulative properties (BCF) and in relation to airborne transport, also atmospheric persistency.

Finally the POPs criteria also quite unspecified refer to "high toxicity" or even "chronic toxicity". At present it is discussed in various fora, such as in the current UNEP POP working groups, to what extent evidence for chronic toxicity should be included in the initial screening or selection stage for identification of potential POPs. The reasons for not requiring such evidence are many, such as the general lack of experimental data on most chemicals regarding their long term effects in mammalian species (including scarce information of effect endpoints like cancer, mutagenicity and reproductive effects) and in species normally used in ecotoxicological tests. Furthermore knowledge about more subtle types of chronic toxic effects caused by chemicals like disturbance of the endocrine system, immunological effects and indirect effects affecting - e.g. the competitive abilities of wildlife - are generally very difficult to detect and prove, and are thus generally not available. Finally the

combined effects of the joint action of the many released and occurring chemicals - including possibilities of "synergistic effects" - are generally not known.

It should be noted that the criteria which have been established for identification of POPs and PBTs are somewhat arbitrary, but nevertheless fixed according to experience, when employed in screening exercises on a great number of substances. The criteria for persistency is often regarded to be the basic criteria, because this criterion addresses not only a potentially concerning property of the substance *after* having reached the target ecosystem, i.e. that it will stay for a long time in a remote environment also after having been transported long distances. Environmental persistency also gives indications on the stability within the mobile environmental transport media air and water themselves. Persistency in air and water gives therefore also indications on the inherent possibility of the substance undergoing long-range environmental transport (cf. below). Whether persistency in the atmosphere as a selection criteria for POPs, is also required, depends on whether the regulation only concerns releases of substances to the air, such as for the ECE Protocol on Long Range Transboundary Atmospheric Transport.

POP criteria on persistency
In relation to environmetal transport:

T ½ air *:

2 days: 1 day:

With wind speed of 5m/sec (18-km/h) *gentle breeze*
(Beaufort Force: 3)
=>
Transport distance for lowering the initial conc. to 3 %:

4320 km ***2160 km***

With wind speed of 10m/sec (36-km/h) *fresh breeze*
(Beaufort Force: 5)
=>
Transport distance for lowering the initial conc. to 3 %:

8640 km ***4320 km***

T ½ water *: 2 months

With speed of river water / ocean current of *0.5 km/h (0.3 knots)*: =>
Transport distance for lowering the initial conc. to 3 %:

3600 km

With speed of river water /ocean current of *1 km/h (0.7 knots)*: =>
Transport distance for lowering the initial conc. to 3 %:

7200 km

***: Relating to half-life under environmental conditions**

Various suggestions on how to include the above mentioned concerns and establish criteria for POPs and other substances of special concern have thus been put forward (cf. e.g. Government of Canada, 1995; Cowan, 1998; Rodan et al, 1998; the UNEP and USEPA homepage on POPs/ PBTs, Swedish Chemicals Inspectorate, 1994; Danish EPA, 1996; Blok H. et al 1998). Various selection procedures employing criteria for identification of POPs have been proposed. An overview and explanation of these criteria and a justification for them are included e.g. in Cowan, 1998; Rodan et al, 1998; Blok H. et al., 1998 and in documents on the USEPA PBT and UNEP POP home pages on the internet.

Use of QSARs and experimental data for selection and ranking of PBTs / POPs.

Generally there is a substantial lack of experimental test data and environmental release and monitoring data on chemicals. Even for high production volume data experience has shown a considerable lack of experimental test data on basic toxicological, ecotoxicological and environmental fate data. Cf. ***Appendix 1*** showing the availability of selected test data in the IUCLID (i.e. concerning substances produced in high tonnage in the EU). The reported percentages regarding the data availability in the IUCLID for various effect endpoints may be somewhat higher or lower than the actual availability of all existing test data. The reasons include factors like: not all available test data have actually been included in IUCLID, when the percentages were counted, the data were not evaluated: in some instances only a statement or a reference to another substance is mentioned in the database, actual experimental data for the substance in question is actually not included.

Some recent US studies (cf. e.g. US EPA, 1998 and US EPA OPPT homepage on the internet concerning the HPV Challenge Program) have also shown that for a very significant fraction of high tonnage substances basic experimental test data are not presently available.

Analogously: for many high tonnage substances environmental release and monitoring data do not exist.

For substances with a lower tonnage even less experimental data and environmental monitoring data are generally available, even though some of these substances may have a significant potential for environmental release, persistence and transport over long distances and also for exerting harmful effects to biota either directly or because of increase of the internal dose due to bioaccumulation.

Use of existing experimental effect data and monitoring evidence have often been employed for both selection/ ranking (nomination) and for risk assessment purposes. The shortcoming of requiring experimental data on individual substances for the selection / ranking (nomination) process is thus, that because of lack of such data far most substances cannot be evaluated.

The same kind of shortcoming applies if requiring monitoring evidence on substances individually, because consideration for evaluation then requires that monitoring of the substances has taken place. The monitoring approach is by nature retrospective and not a proactive approach. Furthermore, for monitoring to take place, it is a prerequisite that a suitable analytical method has been developed, that monitoring has been regarded necessary, and that the substance has actually been detected.

The reason for requiring experimental toxicological and ecotoxicological test data on individual substances to be considered for selection / ranking purposes is often implicitly, that such data are regarded as a true reflection of the toxicogical and ecotoxicological properties. Whereas availability of high quality experimental test data *is* extremely valuable also for selection/ ranking purposes, it is, in the majority of selection / ranking processes which have been employed, a problem that the

experimental data quality and reliability could not be evaluated in depth. The main reason is most often that too many substances have to be evaluated in relation to the required manpower for in-depth data quality assessment. Another, but not minor, problem is of course, that the majority of substances were left out for selection and ranking due to lack of experimental data. These two kinds of shortcomings by requiring employment of experimental test result in two main problems:

- lack of consideration for selection and ranking of the majority of substances because of lack of experimental data

- uncertain selection / priority setting because the available experimental data is of unknown quality (which concerns both reliability and representativity)

The latter problem can also be illustrated by an example. The spread of reported short term LC50 (fish) from the US EPA database, AQUIRE for a well investigated substance like DDT covers five orders of magnitude - this enormous difference in recorded effects concentration hardly represents the true smaller difference in fish species sensitivity of the short term toxicity of DDT. The reason for the huge difference in recorded effect concentrations for acute fish toxicity is also - and most likely - primarily caused by inclusion of both high quality and poor quality data. Very often existing experimental data for the same ecotoxicity or toxicity endpoint for the same substance vary significantly not only caused by differences in the biological test systems, but very often also because both high quality data and unreliable data are recorded in databases, reference handbooks and primary scientific data sources. If experimental data therefore should be used for selection / nomination of POPs a thorough data evaluation must take place. For practical reasons however this has seldomly taken place.

Lack of experimental data or lack of knowledge regarding data quality is however not the same as lack of information. Employment of SARs and QSARs have increasingly been discussed in relation to selection / ranking purposes as well as in the context of effects assessment (e.g. ECETOC, 1998; OECD 1999; US EPA; 1994; Zeeman et al, 1995); for environmental hazard classification (e.g. TemaNord, 1995; non published working papers of the OECD group on harmonisation of classification systems for the aquatic environment) and for risk assessment (cf. e.g. European Commission, 1996).

It is neither the objective of this paper to discuss the employment of QSARs in relation to these aspects and the limitations of the various QSARs nor to address the important issue regarding internal and external validation of QSARs. These aspects have extensively been addressed elsewhere (cf. e.g. OECD 1994; European Commission, 1996; ECETOC 1998) and development takes place all the time in relation to questions like:

- Which QSARs can be used ?
- How should they be used and in what regulatory context are they useful - e.g. for priority setting for testing or assessment, for hazard identification (for e.g. classification), for establishment of environmental quality criteria or objectives and / or for risk assessment (as a basis for limitation of use and ban) ?
- How should validation of QSARs take place ?
- and how does the predictive power of QSARs (e.g. as expressed in the relative number of predicted false positive and false negatives) relate to the desired (regulatory) purpose of using these QSARs?

In relation to selection and ranking purposes, QSARs have widely been employed e.g. in relation to identification of “potentially persistent bioaccumulators” (Zeeman et al 1995) or to selection of substances, which are regarded as “especially harmful” and therefore generally “undesirable” (e.g.

Danish EPA 1996, 1998 and 1999). In a recently issued report (Blok H et al , 1998) a POP selection exercise on 563 substances with high acute aquatic toxicity based on experimental data is presented by employing a combination of the experimental toxicity data and QSARs for environmental fate related properties on which international POP criteria have been established.

In this paper we have adopted a slightly different approach and have only employed QSARs for initial selection of POPs. By doing so it has been avoided that only a tiny fraction of existing substances, namely those with experimental data could be included in the exercise. Furthermore it was avoided by this procedure to go through a laborious data quality check on the available experimental data. Even though the selection based on QSARs also have some limitations - cf.-. below - a significantly greater number of substances could be included in this exercise than have been reported elsewhere. Hopefully thus this QSAR based selection of potential POPs may give a more realistic - but off course still rough - impression of the relative approximate number of substances which would be identified when employing different POP selection criteria. Furthermore this comprehensive selection exercise may also give a supplementary impression of which types of chemicals may be regarded as POP or PBT candidates.

Generally it may therefore be concluded that employment of QSARs may play a potential useful role in relation to selection / ranking. The advantage of a QSAR approach for this is that almost all discrete organic substances (substances with a defined molecular structure) can be included e.g. including those on the EU market (cf. EINECS containing 100.000 different existing chemicals on the EU market). Around half of the chemicals on EINECS are discrete organic substances (with a well-defined molecular structure). Virtually all these chemicals could be included in the employed QSAR based screening.

The QSARs employed for selection of candidate POPs are QSARs for biological degradation, bioaccumulation, hydrolysis and volatility. The latter two QSARs did not influence the number of selected substances significantly, but were employed for removing substances which are unlikely to behave like POPs, because they are degraded in water chemically or unlikely to enter soil/sediment or aquatic ecosystems, because they exclusively partition to the air compartment. For identification of POPs in relation to air pollution only (cf. ECE LRTAP POPs Protocol) also employment of QSARs for atmospheric stability was employed. Some QSARs have been used in an informative way, but not for the actual selection of the POP candidates. This is the case for the indicatively employed QSARs for aquatic and mammalian toxicity.

For certain of the employed QSARs below the estimations are generally believed to be reliable - this is the case for e.g. photodegradation in air (i.e. the estimation of persistency in the gaseous phase of the substances, cf. e.g. Güsten H., 1999), volatility and hydrolysis. However for the latter employment was only possible on a rather small fraction of the total number of substances due to limitations intrinsic to this QSAR model. This model was only able to handle 29.037 different substances out of the total number of 166.075 substances, i.e. on around 17 % of all of the substances. All other QSAR models could be employed on virtually all of the substances.

For some of the employed QSARs, external validation (validation on substances with experimental data, which have not been employed for establishing the QSAR) has not been possible or only been done to a limited extent or done but not yet published. The former is the case for the QSARs for carcinogenesis and reproductive toxicity, whereas the latter concerns the QSARs for biodegradation and mutagenesis, respectively. The QSAR model for prediction of in vitro mutagenisis (Ames test) was recently validated externally employing 140 different substances with a good result. (Niemelä J, 1999). One of the employed QSARs for biodegradation (BPP1, cf. later) have been externally validated several times with the result that this QSAR model gives reliable predictions regarding not ready biodegradability (cf. OECD, 1994; TemaNord, 1995; Roije et al, 1999).

In case of the aquatic toxicity only *minimum* estimates can be obtained by employing the chosen QSAR This is the case for the QSAR for prediction of "the acute baseline" fish toxicity extrapolated to long term minimum toxicity for fish by combination with one of the QSARs for bioaccumulation (BCF-Syracuse). Reliable QSARs may not presently cover some substances, such as substances exerting certain types of specific modes of toxic action towards aquatic organisms/ belonging to certain categories or "families" of chemicals for acute fish toxicity. These substances will typically exert excess toxicity than that estimated by employing available QSARs. These substances may thus "slip through" a QSAR based screening for long-term fish toxicity. Use of available experimental long-term aquatic toxicity data may thus be used in a supplementary selection procedure.

In other cases, such as for bioaccumulation in fish (e.g. if using log Kow estimates) present QSARs tend to make estimates which are somewhat more conservative than actual experimental test data indicate would be reasonable. Thus employment of such QSARs may result in selection of too many substances. QSAR-models like the BCF-Syracuse or BCF-Connell for direct prediction of BCFs in fish may be more realistic to employ. Alternatively or supplementary validated experimental bioaccumulation test data may be used. In this case industry would be encouraged to provide such data if "the burden of proof was changed" for substances which according to employment of QSARs were evaluated to be potentially resistant to biological / microbial degradation and in the same time bioaccumulative in fish.

For a further discussion of the employed QSARs and QSAR trigger values in relation to the POPs criteria, cf. the sections after the POPs selections in flow-chart.

Employed QSARs for selection of candidate POPs/ PBTs.

The Danish EPA QSAR database.

In this paper we report the results of a screening for POP / PBT candidates of global concern by employing selection criteria as close as possible to the ECE LRTAP and UNEP POP criteria on *all* organic substances with 3 D-structures in the Danish EPA QSAR database, i.e. on 166.075 organic substances.

For general information regarding the employed QSARs please refer to e.g. Tema Nord 1995: 581; Blok H. et al. (1998) and the references given in Meylan W. M. & P. H. Howard (1999). Some justification regarding the chosen cut off criteria compared with the POP criteria is addressed in Blok H. et al. (1998).

In the selection exercise data from the *in-house DK EPA QSAR database* was utilised. This database and the employed QSARs will be described in another paper by J. Niemelä, Danish EPA, but briefly the database contains data on *166 075 different organic substances* such as:

* 3D- structures, SIMLES (i.e. a code of the chemical structure which can be handled by the QSAR programmes), CAS no, and in most cases chemical names

* intrinsic properties like those included in the Syracuse EPIWIN program (version 3.02, including molecular mass, vapour pressure, Biodegradation Probability Program 1 to 4 (BPP1 to 4), BCF for fish (BCF-Syracuse), log Kow, water solubility, Henry's law constant, melting point, boiling point),

* baseline short term aquatic toxicity data on fish, Daphnia, algae (according to the preferred QSAR equations for minimum toxicity for non polar narcosis of the EU TGD for risk assessment in EU (European Commission, 1996) and the US EPA classification of the mode of toxic action in fish (US EPA, 1994; Clemens et al, 1995; Smrcheck et al, 1995),

* the estimated time for reaching 95 % equilibrium in a fish test according to OECD TG 305 (t95), and BCF according to Bintein (Bintein S. & J. Devillers, 1993; Devillers et al, 1996) and according to Connell (cf. the EU TGD), and Mackay level-I environmental distribution (cf. the EU TGD)

*TOPKAT (version 5.0) predictions for acute mammalian toxicity, skin sensitisation, subacute mammalian toxicity, Ames test mutagenicity, mammalian carcinogenicity (rat and mouse, male and female) and mammalian reproductive toxicity.

The database also contains data like whether the substance is present on EINECS; EU High Production Substance List; the preliminary EU Medium Production Volume List; occurrence in cosmetics in the EU; occurrence in products on the Danish market (and number of different preparations if so); occurrence in products on the market in one or more of the four Nordic countries Denmark, Sweden, Norway or Finland, and the official EU classification and labelling (according to presence on the EU "List of Dangerous Substances").

The database makes it possible to make searches by setting conditions to the mentioned types of data, including possibilities for searches based on fragments of the chemical structures (because the structural information of the substances is included in the CHEM-X-system).

The employment of QSARs has been based on the availability of SMILES codes for the organic structures under consideration. The total number of SMILES in the database is around 180.000. The SMILES codes originates from various sources such as SMILES lists from the European Chemicals Bureau, US EPA, Syracuse Inc., National Cancer Institute (US) and Danish EPA.

Therefore the employed selection procedure has *not* included organic substances without a precisely defined chemical structure. In this way plant-, animal- and mineral-oil derived substances have been excluded from consideration, i.e., complex hydrocarbons like oil fractions, and oil derived substances, like certain types of detergents and chlorinated hydrocarbons have not been included in the selection exercise. (This includes for examples short chain chlorinated paraffins, which have been considered as a POP candidate in the development of the ECE LRTAP POPs protocol). It may later to some extend be possible to include complex substances in an exercise by allocating an average MW for some of these substances and then define a SMILES code by hand. Such an extension will however be quite laborious (cf. also a discussion about this in Danish EPA, 1996 and Blok H. et al., 1998). Another limitation of the current selection procedure was that only those chemicals, for which a 3 D chemical structure were available in our database, were included for consideration at this stage. Therefore acknowledged POPs like chlordane, endrin, dieldrin and heptachlor do not yet appear in the current version of the Danish QSAR database. QSAR predictions based on the QSARs employed in the selection procedures below show however, that all of these POPs would have been identified, if they had been included in the exercise. The current limitations of the Danish EPA QSAR database reduced the total number of chemicals which was included in the screening from approximately 180 000 on which SMILES codes are available to around 166 000 different chemicals.

The selection procedures.

The selection procedures below were followed by applying the screening criteria sequentially roughly in the same order as proposed to be used in the LRTAP or draft UNEP POP criteria.

The flow charts below also indicate the number of identified substances at each stage of the various selection processes. Attempts have been made to investigating the number of organic substances selected by employing QSAR calculated data which represent proposed or adopted criteria regarding POPs and which are also expressed in regulatory frameworks like the marine conventions. The main difference between the selection procedures is whether or not atmospheric transport is a prerequisite (as it is in the context of the ECE Protocol on POPs undergoing Long Range Transboundary Atmospheric Transport). Other differences include parameters and cut offs chosen for bioaccumulation. Cf. the following pages:

Selection 1 to 6 is selections of POPs according to the draft ***UNEP POP criteria***. The main selection procedure consists of selection stages where *persistent; non-volatile and non-hydrolysable substances are identified. The further selection then depends on the method for selecting bioaccumulative substances*: They are identified by either using two alternative cut off values for BCF in two different estimation methods for BCF in fish or by using estimated logKow values and employing the two alternatively proposed log Kow cut off values.

Selection 7, 8 and 9 are selection of POPs according to the ***ECE criteria for POPs which undergo long range atmospheric transport***. Thus it is <u>also</u> *a prerequisite that these substances fulfill the criterion for persistency in air* besides being chemically and biologically recalcitrant or persistent, i.e. persistent in water/soil/sediment.

Some background information on the chosen QSARs and selection parameters in relation to the POPs criteria appears after the presentation of the outcome of the selection procedure. This includes the number of selected substances at each stage, i.e.:

- in total (amongst the 166075 substances in the database),
- amongst substances on EINECS (existing substances on the EU market, EINECS),
- amongst the High production Volume Substances in EU (produced in > 1000 t per manufacturer or importer per year (1990-93),(HPVC))
- amongst the preliminary (until spring 1999) reported Medium Production Volume Chemicals in EU (produced in >10 t and < 1000 t per year per manufacturer, around half the number of the medium production volume chemicals are thought to have been reported until now; pMPVC)
- amongst substances known to be marketed in one or more of the four Nordic countries Denmark, Finland, Norway or Sweden according to information from Product Registries (NRPR)
- amongst HPVC or pMPVC or NPR (cf. above)

Selection 1: POPs selected according to the draft UNEP POP criteria, i.e. selection of chemicals which may be transported long distances *either* via air *or* by water in rivers or marine currents. The employed criterion for bioaccumulation (*BCF-Syracuse (fish)<u>≥</u> 1000*) is slightly less than the current draft UNEP cut off value for aquatic animals (BCF <u>≥</u> 5000), because BCFs of certain other aquatic animals, like mussels, are generally higher than BCFs in fish. Furthermore chemicals having "significantly less bioaccumulation" are also regarded as fulfilling the criterion, if they are "highly toxic" or "bio-monitoring evidence for bioaccumulation exists".

Selection 2: POPs equivalent to the draft UNEP POP criteria, i.e. chemicals which may be transported long distances either via air or by water in rivers or marine currents. However the most stringent criterion for bioaccumulation in aquatic animals by taking *BCF-Syracuse (fish) <u>≥</u>5000* is employed, i.e. only extremely bioaccumulative substances are selected.

Selection 3: POPs equivalent to the draft UNEP POP criteria, i.e. chemicals which may be transported long distances either via air or by water in rivers or marine currents. However the *BCF-Connell ≥1000* is used as cut off for identifying candidate POPs.

Selection 4: POPs equivalent to the draft UNEP POP criteria, i.e. chemicals which may be transported long distances either via air or by water in rivers or marine currents. However the BCF-*Connell ≥5000* used as cut off for selection of candidate POPs.

Selection 5: POPs equivalent to the draft UNEP POP criteria, i.e. chemicals which may be transported long distances either via air or by water in rivers or marine currents. However the lower of the proposed cut off values, i.e. log *Kow value of ≥ 4* is used as an indicator for selection of candidate POPs.

Selection 6: POPs equivalent to the draft UNEP POP criteria, i.e. chemicals which may be transported long distances either via air or by water in rivers or marine currents. However the higher of the proposed log *Kow value of ≥ 5* are used as an indicator for selection of candidate POPs.

Selection 7: Selection of POPs equivalent to the ECE LRTAP POP criteria. The employed criterion of *atmospheric half-life of 1 day* is less than the adopted guiding cut off value of 2 day of the ECE LRTAP criterion. Arguments for investigating a selection employing a lower atmospheric half life includes that some of the agreed POPs do not fulfill the ECE LRTAP half life criterion for atmospheric stability in the gaseous phase, that some chemicals may be protected from atmospheric degradation due to transport in an sorbed state on airborne particles, and that the ECE LRTAP criterion for atmospheric stability includes use of monitoring data alternatively for indicating sufficient atmospheric stability to be considered for being of concern as regards long range air transport. The selection may therefore be of relevance in a global framework but may also be of special relevance for atmospherically transported substances at a continental/regional scale. The employed criterion for bioaccumulation (*BCF-Syracuse (fish) ≥ 1000*) is slightly less than the adopted ECE LRTAP criterion BCF cut off (BCF ≥ 5000), because BCFs in certain other aquatic animals, like mussels, are generally higher than BCFs in fish. Furthermore chemicals having "significantly less bioaccumulation" are also regarded as fulfilling this POP criterion, if the chemicals are "highly toxic" or "bio-monitoring evidence for bioaccumulation exists".

Selection 8: Selection of POPs equivalent to the ECE LRTAP POP criteria. The criterion for *atmospheric* stability is that of the ECE LRTAP protocol, i.e. a *half life of 2 days*.. The employed criterion for atmospheric half live fulfill the ECE LRTAP criterion. The employed criterion for bioaccumulation (*BCF-Syracuse (fish) ≥ 1000*) is slightly less than the adopted ECE LRTAP criterion BCF cut off (BCF ≥ 5000), because BCFs in certain other aquatic animals, like mussels, are generally higher than BCFs in fish. Furthermore chemicals having "significantly less bioaccumulation" are also regarded as fulfilling this POP criterion, if the chemicals are "highly toxic" or "bio-monitoring evidence for bioaccumulation exists".

Selection 9: Selection of POPs equivalent to the ECE LRTAP POPs criteria. The cut offs for persistency and bioaccumulation of the ECE LRTAP are employed by taking the *atmospheric half-life (for the gaseous phase) of ≥ 2 days and BCF (in fish) of ≥5000.*

Selection 1 to 6:

POPs with potential for transport via air *or* rivers/ marine currents:

Organic substances :				
		166 075		
EINECS: 46708	**HPVC: 1212**	**NRPR: 3113**	**pMPVC: 3772**	**HPVC/NRPR/pMPVC: 6165**

bio-persistence: ⇓ **BPP1< 0.15 & BPP3< 2.2**

Biologically persistent substances :				
		15490		
EINECS: 4075	**HPVC: 54**	**NRPR: 227**	**pMPVC: 340**	**HPVC/NRPR/pMPVC: 494**

volatility : ⇓ **VP < 50 Pa**

non-volatile biologically persistent substances :				
		14343		
EINECS: 3499	**HPVC: 46**	**NRPR: 214**	**pMPVC: 307**	**HPVC/NRPR/pMPVC: 448**

hydrolysis: ⇓ **$t_{1/2}$ > 30 days**

non-volatile, non-hydrolysable biologically persistent substances :				
		13118		
EINECS: 3165	**HPVC: 44**	**NRPR: 199**	**pMPVC: 285**	**HPVC/NRPR/pMPVC: 418**

⇓

selection for bioaccumulative potential

cf. the table next page

The table below presents the ***number of*** different substances out of those selected in the above flow chart by employing different alternative QSARs for identifying bioaccumulative substances. Thus the number indicates how many PBTs***- or POPs*** - candidates which are being identified by applying the ***selection procedure no. 1 to 6.***

Selection 1: BCF-Syracuse ≥ 1000
Selection 2: BCF-Syracuse ≥ 5000
Selection 3: BCF-Connell ≥ 1000
Selection 4: BCF-Connell ≥ 5000
Selection 5: logKow ≥ 4
Selection 6: log Kow ≥ 5

Selection:	Total:	EINECS:	HPVC:	NRPR:	pMPVC:	HMNRC:
1	2888	520	7*	46	48	76**
2	1543	255***	5	22	22	38
3	5865	1370	13	85	122	172
4	4195	986	9	70	90	131
5	6185	1498	18	107	132	195
6	4109	1033	13	86	90	143

- **Total**: in total (amongst the 166075 substances in the database),
- **EINECS**: amongst substances on EINECS (existing substances on the EU market),
- **HPVC**: amongst the 2700 High production Volume Substances in EU (produced in > 1000 t per manufacturer or importer per year (1990-93),
- **pMPVC**: amongst the preliminary (until spring 1999) reported Medium Production Volume Chemicals in EU (produced in >10 t and < 1000 t per year per manufacturer, around half the number of the medium production volume chemicals are thought to have been reported until now)
- **NRPR**: amongst substances known to be marketed in one or more of the four Nordic countries Denmark, Finland, Norway or Sweden according to information from Product Registries
- **HPVC/pMPVC/NRPR = HMNRC**: amongst HPVC or pMPVC or NPR (cf. above), i.e. commercially widely occurring substances

Footnote:

*** & **:** These substances have been listed - cf. Appendix 2

*****:** If hydrolysable and volatile chemicals had not been excluded, this number of chemicals would be 343

It is not surprising that the number of selected candidate POPs are less the high the trigger value for the bioaccumulation potential. Also not surprising is that the *number of selected candidate POPs decrease according to the employed QSAR in the following order: log Kow > BCF-Connell > BCF-Syracuse.*
The number of candidate POPs selected by BCF-Syracuse is a third to half the number of candidate POPs selected by BCF-Connell (cf. the section "3) Bioaccumulation" with a brief description of these BCF QSARs)

Selection 7:

Potential for transport via air: POPs of potential *global* concern

(Initial selection of biologically persistent substances - cf. previous flow chart)

⇓

⇒ **Biologically persistent substances**: **15 490**

atmospheric persistence: ⇓ **AOP : $t_{1/2} \geq 1$ day**

atmosphericly & bio-persistent substances : 4786

EINECS: 1591	HPVC: 32	NRPR: 63	pMPVC: 132	HPVC/NRPR/pMPVC: 187

bioaccumulation: ⇓ **$BCF_{Syracuse} \geq 1000$**

Bioaccumulative atmospherically & bio-persistent substances: 1022

EINECS: 291	HPVC: 7	NRPR: 15	pMPVC: 25	HPVC/NRPR/pMPVC: 35

volatility: ⇓ **$VP_{Syracuse} < 50$ Pa**

non-volatile, bioaccumulative, atmospherically & bio-persistent substances 882

EINECS: 204	HPVC: 7	NRPR: 13	pMPVC: 18	HPVC/NRPR/pMPVC: 28

hydrolysis: ⇓ **$t_{1/2}$ (hydrolysis) > 30 days**

candidate POPs with potential for long range air transport 792

EINECS: 191	HPVC: 7	NRPR: 12	pMPVC: 17	HPVC/NRPR/pMPVC: 27

Selection 8:

Potential for transport via air: POPs of potential *global* concern

(initial selection of biologically persistent substances - cf. previous flow chart)

⇓

⇒ **Biologically persistent substances**: **15 490**

atmospheric persistence: ⇓ **AOP : $t_{1/2} \geq 2$ *days***

atmospherically bio-persistent substances : 3768

EINECS: 1336 HPVC: 32 NRPR: 49 pMPVC: 108 HPVC/NRPR/pMPVC: 157

bioaccumulation: ⇓ **$BCF_{Syracuse} \geq 1000$**

Bioaccumulative atmospherically & bio-persistent substances : 767

EINECS: 237 HPVC: 7 NRPR: 10 pMPVC: 18 HPVC/NRPR/pMPVC: 27

volatility: ⇓ **$VP_{Syracuse} < 50$ Pa**

non-volatile, bioaccumulative, atmospherically & bio-persistent substances 647

EINECS: 159 HPVC: 7 NRPR: 8 pMPVC: 11 HPVC/NRPR/pMPVC: 20

hydrolysis: ⇓ **$t_{1/2}$ (hydrolysis) > 30 days**

candidate POPs with potential for long range air transport 583

EINECS: 149 HPVC: 7 NRPR: 7 pMPVC: 10 HPVC/NRPR/pMPVC: 19

Selection 9:

Potential for transport via air: POPs of potential *global* concern

(initial selection of biologically persistent substances - cf. previous flow chart)

⇓ ⇒ **Biologically persistent substances**: **15 490**

photo-persistence: ⇓ **AOP : $t_{1/2} \geq 2$ *days***

atmospherically bio-persistent substances : 3768

EINECS: 1336 **HPVC: 32** **NRPR: 49** **pMPVC: 108** **HPVC/NRPR/pMPVC: 157**

bioaccumulation: ⇓ **$BCF_{Syracuse} \geq 5000$**

Bioaccumulative atmospherically & bio-persistent substances : 413

EINECS: 107 **HPVC: 5** **NRPR: 3** **pMPVC: 10** **HPVC/NRPR/pMPVC: 14**

volatility: ⇓ **$VP_{Syracuse} < 50$ Pa**

non-volatile, bioaccumulative, atmospherically & bio-persistent substances 363

EINECS: 71 **HPVC: 5** **NRPR: 3** **pMPVC: 5** **HPVC/NRPR/pMPVC: 9**

hydrolysis: ⇓ **$t_{1/2}$ (hydrolysis) > 30 days**

candidate POPs with potential for long range air transport 338

EINECS: 68 **HPVC: 5** **NRPR: 3** **pMPVC: 5** **HPVC/NRPR/pMPVC: 9**

Time has not yet allowed for a thorough analysis of kind and number of chemicals selected according to the various selection procedures employed. Some general remarks can however be made at this stage. According to the stringency of the QSARs employed the more chemicals are selected. It is also interesting to note that rather few high production substances, medium production volume substances or generally occurring in chemicals in marketed products are identified as candidate POPs, regardless of the selection criteria employed.

Generally it is noted that a high persistency increases the potential of the substance to:

- stay in the environment for so long time that significant long range transport may occur, and
- to stay in the receiving remote environment for so long time that bioaccumulation and harmful effects in biota may occur.

If a substance is stable in *both* air and water, then the substance has a potential for significant long-range transport via the atmosphere a*nd* the aquatic environment (i.e. by rivers and currents of the ocean). Such a substance is relevant for consideration in relation to POPs undergoing long range atmospheric transport (e.g. LRTAP POPs protocol) but also for POPs undergoing long range transport via the hydrosphere (i.e. UNEP POPs protocol and Marine Conventions).

If a substance is stable only in water, it has only a potential for significant long-range transport via the aquatic compartment, i.e. rivers and currents of the oceans. Such a substance is not relevant for consideration of POPs in relation to long range atmospheric transport (e.g. LRTAP POPs protocol), but *is* relevant in relation to long range transport via the hydrosphere (e.g. UNEP POPs protocol and Marine Conventions).

If the substance is stable in the atmosphere but is degradable in water, then the substance may be transported long distances via air to a significant extent, but the substance will not stay in the receiving environments, because it will be degraded in water / sediment / soil. A prerequisite is of course that the substances are degradable in the ecosystem of the receiving environment.

The employed QSARs in relation to POPs criteria.

1) Persistence in water, soil and sediment

When persistency in water/soil/sediment are evaluated by use of existing experimental test data, it is necessary to evaluate hydrolysis and especially biodegradation of either simulation tests or such data from standard laboratory tests like those generated by employment of OECD Test Guidelines (cf. OECD, 1998; Steinhaeuser K-G, 1999). But even for experimental data extrapolation and / or interpretation have to take place for comparing such data with the half-life criteria for persistency of POPs. It is however obvious that the application of QSARs for biodegradation is necessary, if vast numbers of chemicals have to be evaluated, because even standard laboratory test data on biodegradability are not available for the majority of chemicals. For screening purposes however we think however that it is possible to use expert judgement for identifying substances which are potentially resistant to biological degradation in the manner explained below, especially if available experimental data for POP candidates later on are considered.

Before applying a QSAR approach for identifying chemicals fulfilling the POP criteria for persistency in water/soil/sediment, it was necessary to interpret how such QSAR predictions could be employed relative to the POP criteria. This was done by selection of ***biologically sufficiently persistent*** *chemicals* by employing the QSAR programmes included in the EPIWIN programme package, i.e. the BPP1 and BPP3 programme, on all of the 166 075 different organic

substances (on which SMILES and a 3 D structure are available), and then identifying candidate POPs by requiring that appropriate cut off values for the values predicted by each of these programmes were fulfilled. For a discussion on the validation of BPP1 (< 0.5) for prediction of substances being not readily biodegradable cf. e.g. OECD, 1994; TemaNord 1995, Roije et al, 1999. The criteria employed here are almost the same as those employed in Blok H. et al (1998), who used the criteria BPP < 0.1 and BPP< 2. Our criteria for biological persistency however employed "a benchmark approach", when deciding in appropriate trigger cut offs of BPP1 and BPP3, which were chosen to exactly also include "the inherently biodegradable" trichlorobenzenes (TCBs) as being of sufficient persistency for being identified, i.e. the criteria employed here were: *BPP1 < 0.15 and BPP3 <2.2.* There is a very high possibility, that chemicals selected by employing these two biodegradation QSARs with the respectively indicated cut off values are not readily biodegradable *and* have a half life for ultimate biodegradation in the order of a month or more.

For comparison the persistence criteria included in the LRTAP and draft UNEP POP protocol are half-lives in water greater than 2 months or half-lives in sediment or soil greater than 6 months. These half lives are generally understood as referring to ultimate and primarily microbial degradation and not merely removal (including translocation) from the environmental compartment in question.

In the EU TGD on risk assessment (European Commission, 1996) test data from standard tests on ready and inherent biodegradability are extrapolated to estimated and generalised half lives for ultimate degradation in sewage treatment plants, surface water, sediment and soil to be employed as input parameters in a Mackay level III type of model for making it possible to perform a regional exposure assessment typical for the EU. The extrapolated half-lives for surface water extrapolated from standard laboratory test data are indicated below:

Estimated aquatic biodegradation half lives according to results in standard tests (after EU TGD, cf. European Commission, and 1996) compared with POP criteria half-life in surface water:

readily biodegradable and fulfilling the 10 d time window :	**15**
readily biodegradable not fulfilling the 10 d time window:	**50**
inherently biodegradable:	**150**
not inherently biodegradable:	**∞**
LRTAP /UNEP POP criteria:	**≥60**

These half lives in surface water are in the context of environmental risk assessment in the EU extrapolated to estimated half lives in sediment and soil taking the partitioning of the substance (i.e. Kd, Koc or log Kow) into account. If using experimental data from ready or inherent

standard biodegradability tests in relation to the evaluation of persistence of POPs, it may be considered to estimate half-lives for employment in the same manner as done in the EU TGD.

In soil and sediment the degradation rate of sorptive (lipophilic) substances is decreased compared to less sorptive substances. The sorption of substances to particulate matter is described by the solids-water partition coefficient (K_d) which can be estimated by log Kow. The degradation half-life in soil and sediment can roughly be estimated from half-lives in water.

Biodegradation half-lives for readily and inherently degradable substances in soil and aerobic sediment layers (after European Commission, 1996).

K_d [L/kg]	Corresponding log K_{ow} (*)	estimated half-life [days] readily degradable	inherently degradable
≤ 10	< 4.4	30	300
100 - ≤ 1000	4.4. - 5.7	300	3000
1000 - ≤ 10000	5.7 - 6.9	3000	30000
LRTAP/ UNEP POP	≥4 or 5 (or less**)	≥180	

*) assuming $K_d = f_{oc} \cdot K_{oc}$, $\log K_{oc} = 0.81 \log K_{ow} + 0.10$ (QSAR for soil and sediment sorption predominantly based on experimental data on hydrophobic chemicals) and $f_{oc} = 0.02$.

**) if presence of other indicators (high toxicity or high bioaccumulation according to biomonitoring).

According to the values used in the EU Technical Guidance for Risk Assessment (European Commission, 1996, cf. upper part of the table above), a two times lower degradation rate is estimated in soil than in surface water for readily degradable substances with a low solid-water partition coefficient (e.g. log Kow < 4). An increase of the partition coefficient will result in a similar increase of the half-life.

Federle et al. (1997) investigated the relationship between degradation rates in ready tests and in a soil mineralisation test. The degradation rates of 9 substances were more or less the same in soil as in the ready test with an average ratio (k_{soil}/k_{RBT}) of 1.2 (0.3-2.8) although no statistical relationship was found, not even when the physico-chemical properties of the substances were taken into account. Also Boethling et al. (1995) have made such comparisons with almost similar conclusions. In average, the degradation half-life was 0.4-1.9 times longer in fresh water than in surface soil. Boethling et al. (1995) concluded that "for screening purposes it is probably acceptable to assume that degradation will occur at roughly comparable rates".

Thus, in conclusion more or less the same biodegradation rate may be assumed for surface soil as for aerobic aquatic environments unless for sorptive substances, i.e. substances with a log K_{ow} > 4. For sorptive substances - like POPs - however, a lower degradation rate may be foreseen in soil and sediment. This may justify that the LRTAP and draft UNEP POP criteria (cf. table

above, lower part), which specifically address lipophilic substances, operate with cut offs for half lives which are three time longer in sediment and soil than in water.

According to the current EU risk assessment procedure readily biodegradable substances with a high lipophilicity - as POPs - and inherently degradable substances are estimated to have half lives in sediment and soil which are significantly longer than the proposed half life POP criteria for soil and sediment.

Generally it may be concluded based on the previous considerations, that *substances fulfilling currently proposed half life criteria for POPs in water (≥ 2 months), soil and sediment (≥ 6 months) in relation to test data obtained by standard test methods are not readily biodegradable, but either "inherently biodegradable" or "persistent".*

As described previously in this section *the employed QSARs* for half life in water, sediment and soil are related to biodegradability and *employ the same general concept* in relation to the POPs criteria than that concluded in relation to use of experimental test data.

For most substances, but especially industrial substances, experimental data do not exist. When such data however do exist, they do normally not refer to tests, where half lives for biological degradation under environmentally realistic conditions are obtained such as low density of micro-organisms, very low concentration of the chemical in question and low temperatures as those prevailing in temperate, sub-arctic and arctic conditions. In relation to this it should be noted that a biodegradation half life of a substance is not only substance specific, but also extremely dependent on the environmental conditions, which often are attempted to be employed under simulation test conditions (cf. e.g. the ISO draft surface water simulation test, ISO (1998)) Application of the POPs criteria for degradation relative to either experimental standard or simulation test data will therefore in any case also require expert judgement (cf. e.g. OECD, 1998; Steinhaueser K-G, 1999).

Selected in our screening procedures utilising the biodegradability predictions on all of the 166 075 substances were 15 490 *different substances*.

2) Persistence in air.

Substances which are also ***persistent in air*** were selected in *selections 7 to 9* by employing the EPIWIN QSAR (AOP) for atmospheric half-life by reaction with OH - radicals. The chosen *half-life was either ≥ 1 or ≥ 2 days*.

Generally the AOP predicted half-lives are regarded to be quite reliable i.e. generally within an order of a factor 3 from that obtained by experimental determination (Gütsen H, 1999) and such QSAR predictions are thus even referred to and used in the context of international risk assessment programmes on existing substances such as the current EU or the OECD programme.

A less strict criterion for atmospheric half life of 1 day instead of the half live cut off of 2 of the draft UNEP and ECE LRTAP protocol has recently been suggested based on the fact that experimental data have shown that the atmospheric half life is much longer, when the substance is adsorped onto particles, which actually may be the case with lipophilic substances. A selection procedure employing this criterion may be highly relevant for substances transported in the atmosphere in relation to shorter distances, e.g. in relation to more regional / sub-continental approaches.

Non of the selected substances would have a tropospheric degradation half-life caused by reaction with ozone of less than 2 days according to the relevant EPIWIN QSAR model.

3) Bioaccumulation.

The ***bioaccumulative substances*** were identified by employing the BCF-predictive model from EPIWIN, i.e. BCF-fish (Syracuse), the BCF-fish QSAR preferred by the EU TGD (EU, 1996) i.e. the BCF-Connell model or by using the LOGKOW calculation method of the EPIWIN programme.

The employed criteria was BCFs of 1000, 5000 or a log Kow of 4 or 5 cf. the *selection procedure 1 to 6.*

Two types of comments can be made to the selection by employing calculated log Kow values. First of all fragmentation based QSARs for log Kow and especially a method such as the LOGKOW (Syracuse) have been validated and is generally shown to be quite reliable (cf. e.g. TemaNord, 1995; Devillers et al, 1995, or references in Meylan W.M. & P.H.Howard, 1999). Secondly logKow is only a good predictor of BCF for fish for those substances which do not undergo significant metabolisation, but which are bioaccumulated in fish because of simple partitioning. Uptake of some extremely lipophilic substances may be stericly hindered, resulting in significantly less bioaccumulation than predicted by their lipophilicity (log Kow).

The BCF (Syracuse) is an empirically based fragmentation model, which takes into account that certain structural and molecular factors influence bioaccumulation. Some chemicals may undergo metabolism or may be stericly hindered for uptake in fish, e.g. because of large molecular size. Such substances may therefore not bioconcentrate to an extent predicted by their log Kow. Another advantage by using BCF-Syracuse instead of log Kow is, that the latter is only a fairly good predictor for bioconcentration in fish for substances with a log Kow value below 6 (cf. e.g. European Commission, 1996). Generally the BCF-Syracuse by employing its fragmentation methodology seems to identify the same types of substances which are also identified by using the bilinear BCF equation by Bintein (Bintein et al 1993, Devillers et al, 1996), if the same BCF cut off criterion of BCF > 1000 is chosen. (This BCF criterion is equivalent to a lipophilicity criterion of log Kow being between approximately 4.2 and 7.15). The BCF-Connell is a QSAR model, which is preferred in the context of risk assessment of existing chemicals in EU (Cf. European Commission, 1996). This QSAR model predicts BCF in a more precautionary way, because it predicts a higher BCF value for chemicals with a log Kow value above 6 (ECETOC, 1999).

In *selection 1* it was noted that the trichlorobenzenes were removed in the selection by employing the above mentioned BCF-Syracuse criterion of 1000 because the predicted BCF (Syracuse) for TCB is less than 1000. The same would have been the case, if either the estimated BCF according to the Bintein equation or the estimated log Kow (for TCB = 3.92) had been used. These predicted BCF values for trichlorobenzenes may be compared with the rather conservative log Kow and BCF values of 4.05 and 2000, respectively, which have been chosen in the draft EU Risk Assessment Report on 1,2,4-TCB (Danish EPA, 1999) based on evaluation of a range of experimental values and including consideration of biomonitoring data. 1,2,4-TCB is also according to a recent modelling exercise (van Pul et al, 1998) evaluated to fulfil the initial screening criteria for POPs with a long-range atmospheric transport potential. TCBs have furthermore during the development of the LRTAP POP criteria been evaluated, and were in that exercise rated to be of quite high concern, although not between the 16 finally selected POPs. This example illustrates the general problem of letting the BCF criterion being too decisive, when screening for POPs. This was also mentioned by different participants during the most recent discussions regarding the future UNEP POPs protocol (pers. comm. from participants of CEG1 meeting of the UNEP POP protocol in Bangkok Oct. 1998). The example may also illustrate that it may for some "borderline substances" be important to supplement the selection

done by employing QSARs with employment of evaluated high quality data from experimental tests.

4) Volatility

Volatile substances were removed by applying a volatility criterion *vapour pressure < 50 Pa* and thereby identifying *persistent, bioaccumulative and semi-volatile substances*. The chosen volatility criterion is somewhat lower that the volatility criterion of the ECE LRTAP POPs convention which is < 1000 Pa. The latter criterion has however recently been criticised and a substantially lower range of vapour pressures has been proposed for consideration as a modified criterion. We chose the above mentioned vapour pressure by comparing the types of substances, which were selected by changing the vapour pressure criterion between 1 and 1000 Pa. The difference in the number of substances was remarkably small. By selecting the above mentioned vapour pressure cut off value mainly longer chained flourinated aliphatics were however removed, which seemed to be reasonable. In the same time substances like TCBs would not have been removed, because their vapour pressure is just below 50 Pa, i.e. according to the EPIWIN model: 33 Pa, and according to evaluated experimental data in the draft EU Risk Assessment Report: 46 Pa (Danish EPA, 1999).

As regards the selections of candidate POPs with a long range transport potential in *either* air or water, it seems reasonable also to remove extremely volatile substances by applying the volatility criterion chosen because substances with such high volatility may tend to stay in the atmospheric compartment and be unlikely to be deposited and enter aquatic or terrestrial ecosystems remote from their site of release.

5) Hydrolysis

Finally ***substances, which are likely to hydrolyse,*** were removed in the selections. This was done by running the *EPIWIN model for hydrolysis* and then removing all substances showing a potential for hydrolysis by having a predicted $t_{1/2} < 30$ *days* giving a final selection of substances fulfilling the screening environmental fate criteria for being considered as candidate POPs with a potential for long range transport. It was only possible to apply this QSAR for hydrolysis for around 17 % of the chemicals due to intrinsic limitations of the EPIWIN model.

Substances removed by employing the QSAR for hydrolysis were for example various peroxides, esters of tetrachloro- and pentachlorophenol, esters of chlorinated benzoic acids, some anhydrids and octa- and nonachlorobutadien. (It is noted that PCP itself, which has been discussed for inclusion within the ECE LRTAP POP protocol, was not identified as a candidate POP, because it has a predicted BCF below 1000 - cf. the discussion above cf. point 3. In addition to this it could be added, that PCP is ionised less and thus significantly more bioaccumulative at low pHs prevailing in acidified environments, such as those in many lakes in the sub-arctic region, and that the BCF prediction employed above does not take such a phenomenon into account.)

Concluding remarks about the QSAR based screenings for POP candidates.

It is by inspection of some of the resulting lists of selected substances according to *selection 1 to 9* likely that quite many substances of such selections may be grouped within classes or "chemical families" (e.g. PCBs, dioxins, PAHs etc.).

Manually "fine-tuning" some of the selections may probably reduce the final number of substances selected further.

The potential of the selected substances of being of concern for aquatic toxicity or for chronic mammalian toxicity was however investigated by utilising the content of the Danish EPA QSAR database. This was however because of time constrains only done for the substances selected by the procedure described in *selection 1*. However the employed procedure below gives an indication on how the concept more generally could be employed:

The potential of the selected POP candidates of being of concern for aquatic toxicity and for chronic mammalian toxicity.

The $NOEC_{fish,\ predicted}$ was estimated by estimating the baseline (*minimum)* toxicity towards fish by assuming that this can be expressed by a predictive model which takes into account the lethal body burden of non polar narcotics at equilibrium (cf. ECETOC, 1995), the potential for bioaccumulation in fish (by employing the BCF-Syracuse predictions and which presumes that the long term NOEC for fish can be extrapolated from the short term fish data by employing a factor of 10 for extrapolation from lethal toxicity to sub-lethal toxicity. By employing this model:

$NOEC_{equilibrium} = 0.1 \times LC50_{equilibr.} = 0.1 \times 8{,}15 \times MW / BCF_{Syracuse}$ (ppm)

On all of the selected substances, the following frequency distribution occurred:

no of substances:	estimated range for $NOEC_{equilibrium}$ for non-polar narcosis (ppm)
0	< 0.001
407	< 0.01
1927	< 0.1
2888	< 1.0

This may be compared with the numbers of substances selected in the flow chart in *selection 1*. (Cf. flow chart and table)

Based on this it may be concluded that ***all of the selected substances are predicted to have a high baseline (minimum) toxicity to fish***.

The potential of the selected substances in *selection 1* to be mutagenic (i.e. the in vitro assay the Ames Salmonella typhimurium / microsome test), of being a mammalian carcinogen or reproductive toxicant was investigated by using the TOPKAT predictions. The TOPKAT QSARs for cancer are based on the US National Cancer Institute Cancer models. Interpretation of the TOPKAT predictions for carcinogenicity may be quite difficult, because the models for this endpoint predict the probability for carcinogenicity in male and female rat and mouse, and because it is not obvious how predictions on the same substance should be regarded, if positive in one species and / or sex

but negative or equivocal in another. Below therefore is only presented the number of substances which have been tested and the number of “reliable predictions” (this concept is defined in the TOPKAT system and thus introduced in the Danish EPA QSAR database by “reliability codes”)

TOPKAT QSAR predictions of mutagenicity, carcinogenicity and reproductive toxicity on POP candidates *(selection 1).*

endpoint:	positive:	negative:	equivocal:	reliable predictions:
cancer rat male	**381**	1106	107	**1594**
cancer rat female	**132**	426	34	**592**
cancer mouse male	**374**	370	32	**772**
cancer mouse female	**773**	423	46	**1242**
mutagenicity	**616**	629	30	**1275**
reproductive tox.	**727**	339	20	**1086**

Number of substances indicated to be mutagenic, toxic for reproduction or carcinogenic according to the TOPKAT ***QSAR*** models (ver. 5.0). ***Total number of substances were 2888 candidate POPs selected by employing selection procedure 1.*** Note that only the number of substances identified with reliability codes "reliable" are included and that the models were generally, but depending of the model, only able to perform reliable predictions of around half or less of the selected substances.

Experimental data on POP candidates (*selection 1)* concerning mutagenicity, carcinogenicity and reproductive toxicity.

endpoint:	positive:	negative:	equivocal:	number of experimental test data:
cancer rat male	**2**	7	0	**9**
cancer rat female	**1**	8	1	**10**
cancer mouse male	**2**	5	0	**7**
cancer mouse female	**1**	6	0	**7**
mutagenicity	**52**	20	0	**72**
reproductive tox.	**7**	2	0	**9**

Number of substances indicated where ***experimental test data*** are available according to the TOPKAT database. The total number of POP candidates was the 2888 selected POP candidates according to ***selection 1***. Note that the relative number of data were 2.4 % for in vitro mutagenicity and between 0.24% and 0.34 % for the cancer and reproductive endpoints. Data from other sources than the TOPKAT database may increase these very small figures slightly.

It is obvious that the existing experimental data is scarce relative to the data which can be generated by use of QSARs (compare the number of chemicals in the two above tables).

Even though evaluation of the above QSAR based predictions may not be so easy, it may based on this preliminary analysis be concluded that

- 10 to 20 % of the selected candidate POPs show some evidence of being mutagenic or carcinogenic. Of the selected substances which could be predicted reliably according to these QSAR models more than half were predicted to be in vitro mutagens whereas between 20 % and more than half (depending on the QSAR model applied) were predicted to have a carcinogenic potential, and

- around 25 % of the selected candidate POPs and more than two third of those which could be predicted by the model for reproductive toxicity show some evidence of being toxic for reproduction to mammalian species

Thus experimental data for POP candidates regarding chronic mammalian toxicity, including mutagenicity, carcinogenicity and reproductive toxicity, seems to be scarce but according to QSAR evidence a fairly high percentage of the selected candidate POPs have a mutagenic/ carcinogenic potential and an even higher percentage may be toxic for reproduction in mammalian species.

Based on a more comprehensive analysis of the data in the Danish EPA QSAR database and ongoing QSAR work within the Danish EPA it may later be possible to make some further evaluations based on e.g. the potential for mammalian chronic toxicity (e.g. to include predicted $NOAELs_{28\ days\ repeated\ dose\ for\ rat}$) and possibly also to expand the analysis for genotoxicity and cancer.

Exposure related data on the selected substances.

A thorough manual inspection in depth of many of the lists of substances selected by employing *selection procedure 1 to 9* has not yet been performed. When doing so focus could easily be put on substances on the EU market (EINECS) or on substances, which are widely occurring (i.e. appears on the EU HPVC list, the pMPVC list or on the Nordic Market (NRPR)). After having done this and for example grouped chemicals of the same structural families and possibly removed some irrelevant substances, an investigation of the availability of experimental test data and information on use profile, volume, environmental release and monitoring of these substances could take place.

Concluding remarks.

Further analysis should be performed. It will be considered to include supplementary or alternatively in the analysis use of *experimental data* and data regarding potential for or evidence *about use category, environmetal release, exposure and occurrence* of chemicals. In this respect we have available quite many computerised databases which may be utilised. Care has however to be taken when using such data because of the questions about data quality, relevance and representatively. Maybe utilising such data should mainly be done *alternatively* for comparison with the current purely QSAR based selection procedure as a kind of "validation exercise". Another possibility may be to *supplement* the present exercise at some stage with selection procedures utilising available data on experimental effects or fate data and/ or information regarding environmental release and occurrence of organic substances.

Acknowledgement.

Thanks for valuable comments to drafts of this paper to Chris Cowan (ICCA), Henrik S. Larsen (Danish EPA), Sylvain Bintein (French EPA), Robert Boethling (US EPA) & Finn Pedersen (VKI).

References:

1) Ahrens A (1999):”Revision of the EU Risk Assessment Methodology with Regard to Persistent, Bioaccumulative Toxic Substances. On behalf of WWF, Germany”, working paper to the participants of the EU Technical Meeting on Risk Assessment of Existing Substances, June 1999.
2) AMAP (1997): "Arctic Pollution Issues: A State of the Arctic Environment Report", pp. 188.
3) Blok J, P.C. Okkermann, F. Balk:”Selection of PTBs phase 2". RIVM/BHK, 1998
4) Bintein S. & J. Devillers (1993):”Nonlinear dependence of Fish Bioconcentration on n-Octanol/Water Partition Coefficient", SAR and QSAR in Environmental Research 1, 29-39.
5) Boethling R.S. , P.H. Howard, J.A. Beauman & M.E. Larosche (1995)”Factors for Intermedia Extrapolation in Biodegradability Assessment", Chemosphere 30 (4), 741-52.
6) Clemens R.G., J.V. Nabholtz, M.G. Zeeman & C.M. Auer (1995):”The application of Structure-Activity relationssships (SARs) in the Aquatic Toxicity Evaluation of Discrete Organic Chemicals", SAR and QSAR in Environmental Research, 3, 203-215.
7) Cowan C.E, D.V. Versteg, R.J. Larson, P.J. Kloepper-Sams (1995)”Integrated Approach of Environmental Assessment of New and Existing Substances", Regulatory Toxicol. And Pharmacol., 21, 3-31.
8) Cowan C.E. (1998): ICCA Contribution to 7/97 request from UNEP and IFCS (revised 29 April 1998):”Procedure for Identifying Further POPs Candidate Substances for International Action".
9) Danish EPA (1996): “Kriterier for udvælgelse af uønskede stoffer” (in Danish only, “Criteria for selection of undesirable substances”), pp. 57.
10) Danish EPA (1996): “Excerpts from a discussion paper from Danish EPA: Chemicals - Status and Perspectives, Excerpts from the discussion paper"(Status og Perspektiver for kemikalieområdet. Et debatoplæg) Dec. 1996 - "English Summary, List of Undesirable Substances", pp. 34.
11) Danish EPA(1997): Arbejdsrapport 1, Effektlisten 1997, pp. 122
12) Danish EPA (1998): "Environmental review: List of Undesirable substances", pp.85
13) Danish EPA (1999): "A Strategy for intensified efforts in the fields of chemicals in Denmark, in the EU and globally", Ministry of Environment and Energy January 1999, pp.37 (cf. http://www.mst.dk/activ/01040000.htm).
14) Danish EPA (1999): “Draft Risk Assessment 1,2,4-Trichlorobenzene, CAS No. 120-82-1, EINECS No.: 204-428-0, Draft of May 26, 1999”
15) Deveillers J, S. Bintein & D. Domine (1996):”Comparison of BCF Models Based on log P”, Chemosphere 33 (6), 1047-65.
16) Devillers J, D. Domine & W. Karcher (1995):”Estimating n-Octanol/Water Partition Coefficients from the Autocorrelation method", SAR and QSAR in Environmental Research 3, 301-306
17) ECETOC (1995) Technical Report No. 67: “The Role of Bioaccumulation in Environmental Risk Assessment: The Aquatic Environment and related Food Webs”,
18) ECETOC (1998) Technical report No. 74: “QSARs in the Assessment of the Environmental fate and Effects of Chemicals”
19) ECETOC (1999)”ECETOC Task force on Persistent Organic pollutants (POPs) Response to UNEP/INC/CEG-1 Annex 1” June 1999.
20) European Commission (1996): “Technical Guidance Document in Support of Commission Directive 93/67/EEC on Risk assessment for New Notified Substances and Commission regulation (EC) No. 1488/94 on Risk Assessment for Existing Substances”, Part I to IV. (EU TGD)
21) Government of Canada (1995):”Toxic Substances management Policy”
22) Government of Canada (1995):”Toxic Substances Management policy. Persistence and Bioaccumulation Criteria”.

23) Güsten H (1999):"Predicting the Abiotic degradability of Organic Pollutants in the Trophosphere", Chemosphere 38 (6), 1361-70.
24) ISO (1998): ISO/CD 14592 Part 1/TR:"Water quality- Evaluation of the aerobic biodegradability of organic compounds at low concentrations - Part 1: Shake flask batch test with surface water or surface water/sediment suspensions"
25) Meylan W.M. & P.H. Howard (1999): User´s Guide for EPIWIN© april 1999, Syracuse research Corporation, USA.
26) Meylan W.M, P.H.Howard, R.S. Boethling, D. Aronson (1999):"Improved Method for Estimating Bioconcentration/Bioaccumulation factor from Octanol/water Partition Coefficient, Environ. Toxicol. Chem., 18 (4), 664-72.
27) Niemelä J (1999):"Computer Predictions of Ames test Mutagenicity (progress report, May 1999), Not yet published (contact point: jrn@mst.dk)
28) OECD (1994): OECD Environment Monograph No. 88. "US EPA/EC Joint Project on the Evaluation of (Quantitative) Structure Activity Relationships", Paris.
29) OECD (1998): OECD Guidelines for Testing of Chemicals. Including the 10th Addendum", Paris, Hard copy or CD ROM.
30) van Pul W.A.J., F.A.A.M. de Leuw, J.A. van Jaarsveld, M.A. van der Gaag & C.J. Sliggers (1998):"The Potential for Long-Range Transboundary Atmospheric Transport", Chemosphere 37 (1), 113-41.
31) Rodan B.:"Review of Screening Critreria for Persistent Organic Pollutants" Presentation to UNECE-LRTAP WGS, october 20, 1997
32) Rodan B., N. Eckley & R.S. Boethling (1998): "International Action on Persistent Organic Pollutants (POPs): developing Science-Based Screening Criteria" (working paper distributed at UNEP CEG Meeting 1 in Bangkok, oct. 1998)
33) Roije E., H. Loonen. M. Müller, G. Klopman & W.J. G.M. Peijnenburg (1999): "Evaluation and Application of Models for the Prediction of ready Biodegradability in the MITI-I test", Chemosphere 38 (6), 1409-17.
34) Smrcheck, M. Zeeman & R. Clemens (1995): "Ecotoxicology and the Assessment of Chemicals at the US EPA´s Office of Pollution Prevention and Toxics:Future Activities and Future Needs", p. 127-58, in "Making Environment Science", eds. J.R. Pratt, N. Bowers & J. R. Stauffer, pp. 271.
35) Steinhaeuser K-G (1999): "Test Methods for POP Screening Criteria", pp. 13, Working paper, UBA / Berlin, June 1999
36) Swedish Chemicals Inspectorate (1994): "Selecting Multi-problem Chemicals for Risk reduction - a presentation of the Swedish Sunset project", KemI Report 13/94
37) Swedish Chemicals Inspectorate (1994): "Chemicals Substance Lists - Guide to the Lists Used in the Swedish Sunset Project - KemI report 10/94" (Supplement to report 13/94)
38) TemaNord (1995):"Environmental Hazard Classification - data collection and interpretation guide (2nd edition)", Nordic Council of Ministers, Copenhagen, pp. 166
39) US EPA (1994): "Estimating Toxicity of Industrial Chemicals to Aquatic Organisms Using Structure Activity relationships", 2nd. Ed. (EPA- R- 93-001, febr. 1994)
40) US EPA (1998): "Chemical Hazard data Availability Study. What Do We Really Know About the Safety of High Production Volume Chemicals ?", EPA´s 1998 Baseline of Hazard Information that is Readily Available to the Public, Prepared by EPA´s Office of Pollution Prevention and Toxics (April 1998).
41) Webster E, D. Mackay & F. Waina (1998):"Evaluating Environmental Persistence", Environm. Toxicol. And Chem. 17 (11), 2148-58.
42) Waina F & D. Mackay (1996) "Tracking the Distribution of Persistent Organic Pollutants, Environmental Science and Technology News", 30 (9), 390A-96A.
43) Zeeman M, C.M.Auer, R.G. Clements, J.V. Nabholtz & R.S. Boethling "U.S.EPA Regulatory Perspectives on the Use of QSAR for New and Existing Chemicals Evaluations", SAR and QSAR in Environmental Research 3, 179-201.

Appendix 1.

Availability of data in IUCLID data on 2700 high tonnage substances:

water solubility:	**36 %**
vapour pressure:	**36 %**
log Kow:	**31 %**
biodeg.	**30 %**
BCF(fish):	**15 %**
LC_{50} fish:	**51 %**
EC_{50} daphnia:	**44 %**
EC_{50} algae:	**30 %**
EC_{50} microorg.:	**21 %**
LC_{50} soil organisms:	**3 %**
NOEC fish:	**6 %**
NOEC Daphnia:	**10 %**
LD_{50} oral rat:	**71 %**
LC_{50} inhalation rat:	**30 %**
LD_{50} dermal rat or rabbit:	**43 %**
Repeated dose mammalians:	**53 %**
genotox in vitro:	**62 %**
genotox in vivo:	**32 %**
reproductive tox mammalians:	**20 %**
developmental tox mammalians:	**28 %**
cancer mamalians: approx.:	**20 %**

Ref: Bjorn Hansen / ECB/JRC (*pers. comm.)*.

High tonnage substance: here defined as a substance with a imported or produced volume > 1000 t / year/ importer or manufacture within EU

Appendix 2.

The EU HPVC selected by selection procedure 1 includes:

- **DDT (CAS no. 50-29-3)***
- **hexachlorocyclopentadiene (CAS no. 77-47-4)**
- **a dihydroxylated and tetrabrominated biphenyl derivative (CAS no. 79-94-7)**
- **dicofol (CAS no. 115-32-2)**
- **hexachlorobenzene (CAS no. 118-74-1)***
- **pentabrominated diphenyl ether (CAS no. 32534-81-9) ****
- **pentachlorothiophenol (CAS no. 133-49-3)**

***: included in the UNEP POPs**
****: currently under EU risk assessment**

The candidate POP list selected by employing *selection procedure 1* and focusing on substances which occur either on the EU HPVC list, on the preliminary EU Medium High Production List or in chemicals products on the market in one or more of the Nordic countries (according to Product Register information) is currently available in the Danish EPA. This list of substances includes 76 different chemicals mainly PAHs, chloro- or bromo-aromatics and certain pesticides. Well-known commercially available POPs like PCBs, PBBs, mirex and toxaphene are selected by the employed selection procedures when these are used on chemicals on EINECS.

Lists according to any of the employed set selection criteria - or other selection criteria employing the chosen QSARs - can be made.

Initial selection of marine pollutants for OSPAR have recently been performed by applying slightly other PBT criteria cf. the working paper by Danish EPA and VROM/NL: "DYNAMEC initial selection: preliminary lists and summary statistics, Danish EPA, 27. May, 1999 " (contact point hel@mst.dk). In this exercise a QSAR approach employing the Danish EPA QSAR database was utilised as well as an exercise based on a Nordic database with experimental data. With the latter approach it was necessary to extrapolate especially the existing standard test data on biodegradability to estimated half-lives in marine water. The availability of experimental test data on EINECS substances was limited even though the Nordic Database was created by merging relevant information from different databases such as AQUIRE, IUCLID, Swedish Sunset Chemicals databases, German Frauenhof report database etc. The resulting number of identified potential marine pollutants selected by employing existing experimental test data was therefore only around 20 % of the number of substances identified by employing QSARs, because the latter could be employed on all discrete organic EINECS chemicals (around 50.000 chemicals).

Appendix 3.

Number of organic substances in the entire DK EPA database according to predicted persistency and bioaccumulation.

Total number of substances: 166.075

BCF-Syracuse:	regardless of biodegradation:	BPP1 < 0.5 *	BPP1 < 0.15 & BPP3 < 2.2 **
> 100	41.949	16.481	6.522
> 500	22.309	9.439	4.255
> 1000	15.449	6.898	3.351***
> 5000	6.609	3.217	1.775***

*: predicted biodegradation: not rapid (equivalent with "not readily biodegradable" according to standard OECD Test Guidelines, cf. OECD, 1994; TemaNord 1995; Roije et al, 1999).
**: predicted biodegradation: very slow biodegradability or persistent (half lives in the order of a month or more)
***: Note that these numbers are a bit higher than the equivalent values *of selection 1 and 2*, respectively, because they include some hydrolysable and volatile substances

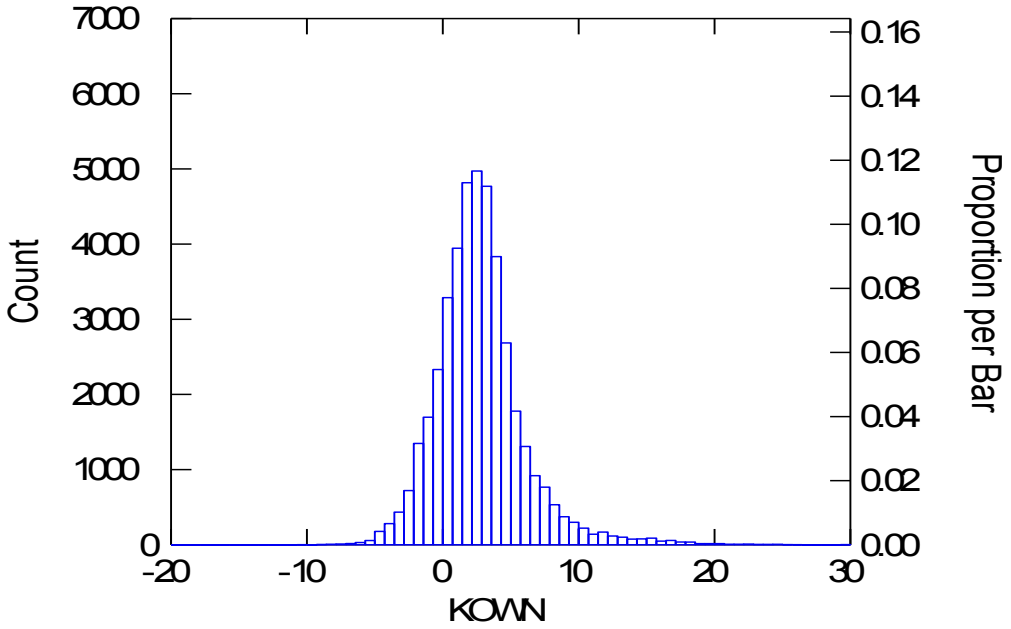
Number of substances with a predicted atmospheric half-life > 2 days: 20.573, and

BCF-Syracuse:	regardless of biodegradation:	BPP1 < 0.5 *	BPP1 < 0.15 & BPP3 < 2.2 **
> 100	3.356	2.730	1.653
> 500	1.621	1.395	1.002
> 1000	1.124	997	759 ***
> 5000	532	505	413

*: predicted biodegradation: not rapid (equivalent with "not readily biodegradable" according to standard OECD Test Guidelines; cf. OECD, 1994; TemaNord, 1995; Roije et al, 1999)
**: predicted biodegradation: very slow biodegradability or persistent (half lives in the order of a month or more)
***: this number of selected substances is slightly higher than the equivalent number by applying *selection 8,* because the selection of this table did not also include substances with a BCF = 1000.

Number of substances with a predicted atmospheric half-life > 1 day (regardless of predicted biodegradability and bioaccumulation): 34.213

Figure of log Kow distribution of EINECS substances (total 42633) which does NOT fulfill the QSAR POP persistency criteria: BPP1< 0.15 & BPP3 < 2.2

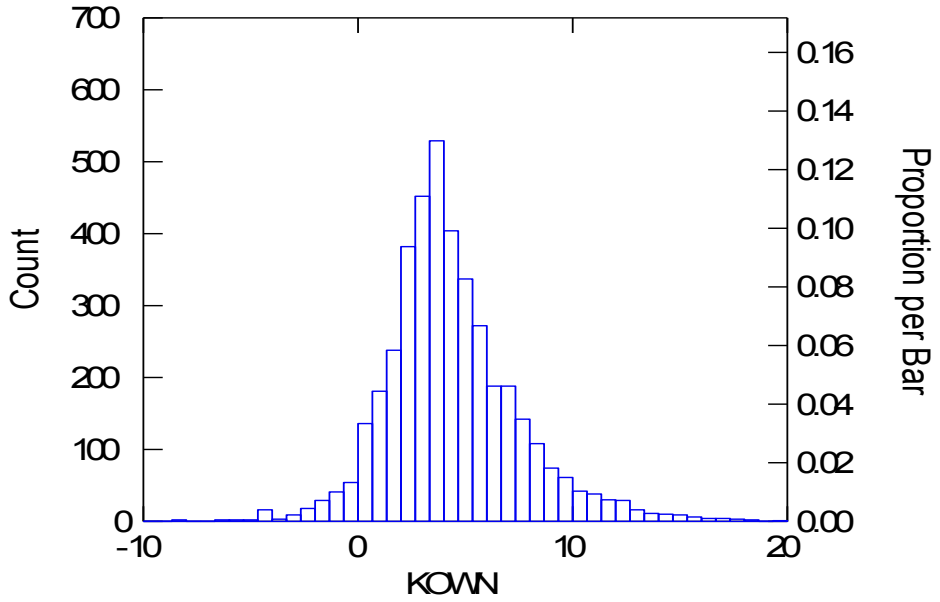


KOWN	
N of cases	42633
Minimum	-10.160
Maximum	29.810
Range	39.970
Sum	115259.960
Median	2.460
Mean	2.704
95% CI Upper	2.734
95% CI Lower	2.673
Std. Error	0.016
Standard Dev	3.258
Variance	10.613
C.V.	1.205
Skewness(G1)	1.233
SE Skewness	0.012
Kurtosis(G2)	4.424
SE Kurtosis	0.024

i.e. median log Kow for not persistent EINECS substances: 2.5

.

Figure of log Kow distribution for EINECS substances (total 4075) which fulfill the employed QSAR POP persistency criteria: BPP1 < 0.15 & BPP3 < 2.2.



KOWN	
N of cases	4075
Minimum	-8.220
Maximum	19.720
Mean	4.359
Standard Dev	3.078

i.e. median log Kow for persistent EINECS substances: 4.4