

Miljøprojekt Nr. 635 2001

Rapport om Vejledende liste til selvklassificering af farlige stoffer

Miljøstyrelsen vil, når lejligheden gives, offentliggøre rapporter og indlæg vedrørende forsknings- og udviklingsprojekter inden for miljøsektoren, finansieret af Miljøstyrelsens undersøgelsesbevilling.

Det skal bemærkes, at en sådan offentliggørelse ikke nødvendigvis betyder, at det pågældende indlæg giver udtryk for Miljøstyrelsens synspunkter.

Offentliggørelsen betyder imidlertid, at Miljøstyrelsen finder, at indholdet udgør et væsentligt indlæg i debatten omkring den danske miljøpolitik.

Indhold

RESUMÉ	5
1 BAGGRUND FOR LISTEN, LISTENS INDHOLD OG BRUG AF LISTEN	7
1.1 BAGGRUND	7
1.2 QSAR MODELLER - EN ALTERNATIV METODE TIL FAREVURDERING	8
1.3 DEN VEJLEDENDE LISTE TIL SELVKLASSIFICERING AF FARLIGE STOFFER	9
1.4 PRODUCENTERS OG IMPORTØRERS PLIGT TIL AT SELVKLASSIFICERE	11
2 TEKNISK BESKRIVELSE AF LISTENS TILBLIVELSE OG DE ANVENDTE QSAR MODELLER	13
2.1 INDLEDNING	13
2.1.1 SAR / QSAR	13
2.1.2 Modellernes domæne	14
2.1.3 Nøjagtighed af modelforudsigelser	14
2.1.4 Software	15
2.2 METODOLOGIEN I UDARBEJDELSEN AF LISTEN	17
2.2.1 De udvalgte farlige egenskaber	17
2.2.2 De vurderede kemiske stoffer	17
2.2.3 Testresultater	18
2.2.4 Brug af QSAR modeller	18
2.2.5 Resultatet	19
2.3 AKUT DØDELIG VIRKNING VED INDTAGELSE	19
2.4 ALLERGIFREMKALDENDE EFFEKT VED HUDKONTAKT	23
2.5 SKADER PÅ ARVEANLÆGGENE	25
2.6 KRÆFTFREMKALDENDE EFFEKT	28
2.7 FARLIGHED FOR VANDMILJØET	32
3 REFERENCER	36

Resumé

Denne rapport indeholder en beskrivelse af Miljøstyrelsens Vejledende liste til selvklassificering af farlige stoffer. Stofferne er identificeret ved hjælp af QSAR modeller (QSAR: Quantitative Structure-Activity Relationship). Listen er et tilbud til producenter/ importører som hjælp til selvklassificering.

Rapportens Del I indeholder en beskrivelse af baggrunden for listen, listens indhold og brug af listen. Del II indeholder en teknisk beskrivelse af de anvendte QSAR modeller, listens tilblivelse og relationerne til kriterierne for klassificering for udvalgte farlige egenskaber. Selve listen kan findes på MST's hjemmeside (www.mst.dk) under "kemikalier".

Ved hjælp af QSAR modeller har Miljøstyrelsen gennemgået ca. 47.000 kemiske stoffer og fundet, at 20.624 stoffer vurderes at skulle klassificeres for en eller flere af følgende farlige egenskaber: Akut dødelig virkning ved indtagelse, allergifremkaldende effekt ved hudkontakt, skader på arveanlæggene, kræftfremkaldende effekt og farlighed for vandmiljøet.

Ifølge kriterierne for klassificering skal klassificering ske på baggrund af tilgængelig viden, hvilket oftest er resultater fra dyreforsøg. Imidlertid er det Miljøstyrelsens erfaring, at producenter / importører finder det svært at opfylde pligten til at vurdere om et stof, der ønskes markedsført, skal klassificeres. Der findes nemlig kun meget få oplysninger om de kemiske stoffers farlighed. Miljøstyrelsen vurderer, at der for ca. 90% af stofferne kun er få eller ingen tilgængelige testresultater fra dyreforsøg m.m. om stoffernes farlige egenskaber overfor mennesker eller miljø.

Kriterierne for klassificering giver udover testresultater fra dyreforsøg også mulighed for at benytte alternative metoder. Det kan f.eks. være undersøgelser, der ikke kræver anvendelse af forsøgsdyr, men som baseres på sammenligning med andre lignende kemikalier ved såkaldte struktur-aktivitetssammenhænge.

QSAR modellering er en sådan alternativ måde at vurdere kemiske stoffers farlighed på. Miljøstyrelsen har gennem flere år arbejdet med at udvikle og anvende QSAR modeller til at forudsige kemiske stoffers egenskaber med. De her anvendte modeller er nu så pålidelige, at de med ca. 70-85% statistisk sikkerhed kan forudsige om stofferne har en eller flere af de udvalgte egenskaber.

For nogle stoffer på den Vejledende liste til selvklassificering af farlige stoffer vil der, trods den generelle datamangel, kunne findes pålidelige oplysninger fra f.eks. egnede dyreforsøg om stoffernes farlige egenskaber. Disse oplysninger bør lægges til grund for selvklassificering fremfor denne listes anbefalinger.

Det skal understreges, at listen ikke har bindende karakter. Ansvar for klassificering påhviler stadig importører / producenter. Miljøstyrelsen opfordrer importører / producenter til at benytte denne Vejledende liste til selvklassificering af farlige stoffer som et redskab i forbindelse med deres

vurdering af kemiske stoffers farlige egenskaber i tilfælde, hvor der er utilstrækkelige eller ingen data for de udvalgte farlige egenskaber.

1 Baggrund for listen, listens indhold og brug af listen

1.1 BAGGRUND

Når kemiske stoffer skal klassificeres for deres farlighed, vurderer man stoffernes iboende egenskaber ud fra den tilgængelige viden om stofferne /1,57/. Ofte sker vurderingen ud fra resultater fra dyreforsøg. Vurderingen skal ske egenskab for egenskab og kan således betyde, at der for bare ét stof skal foreligge mange dyreforsøg. En kortlægning af alle de egenskaber, der i dag klassificeres for, kan således betyde op til tredive dyreforsøg for et enkelt stof.

Undersøgelser har vist, at der findes meget få oplysninger om farligheden for mennesker og miljø af kemiske stoffer på det europæiske marked. Kommissionen vurderede i 1999 omfanget af testdata for de ca. 2.700 stoffer, der er på markedet i stor mængde (over 1.000 tons per producent / importør per år i EU). De fandt, at der kun for 14% af stofferne var de oplysninger om stoffernes farlighed, der ifølge EU-regler som minimum er nødvendige for at risikovurdere industrikemikalier. For 21% af stofferne var ingen testdata for stoffernes giftighed overfor mennesker eller miljø indberettet /2/.

Miljøstyrelsen foretog i 2000 en undersøgelse af tilgængelige data om farlighed for de ca. 100.000 stofindgange på EU's fortegnelse over eksisterende stoffer (Einecs^{*})/3/, som fandtes i to af verdens største kilder til offentligt tilgængelige testdata (RTECS, 2000^{**}; AQUIRE, 1994^{***}). Undersøgelsen viste, at der på udvalgte effekttyper fandtes testdata for følgende procentdele af Einecs-stofferne /4/:

TABEL 1

Akut giftig virkning:	13,4 %	RTECS
Reproduktionsskadelig virkning:	2,5 %	RTECS
Skader på arveanlæg:	3,9 %	RTECS
Kræftfremkaldende egenskab:	1,8 %	RTECS
Vandmiljøeffekter:	3,5 %	AQUIRE

Miljøstyrelsen vurderer derfor, at der i dag stadig mangler oplysninger om kemiske stoffers farlige egenskaber for omkring 90% af stofferne på Einecs. Mange kemiske stoffer på det europæiske marked kan således have ukendte farlige egenskaber, selvom de har været brugt i mange år. Dyreetiske og økonomiske hensyn betyder, at der næppe indenfor en overskuelig

* Einecs: European Inventory of Existing Chemical Substances. Fortegnelse over de stoffer, der mellem 1971 og 1981 blev meldt ind af industrien som værende på det europæiske marked.

** RTECS: Registry of Toxic Effects of Chemical Substances, The National Institute of Occupational Safety and Health, Washington, D.C.

*** AQUIRE: AQUatic toxicity Information RETrieval, United States Environmental Protection Agency, Mid-Continent Ecology Division, Duluth, MN.

tidshorisont vil foreligge testdata for en stor del af disse stoffers farlige egenskaber.

Kriterierne for klassificering beskriver hvordan tilgængelige eksperimentelle testdata (fra dyreforsøg m.m.) skal anvendes i vurderingen og klassificeringen af stoffernes giftighed overfor mennesker og miljø. Kriterierne beskriver også, hvordan stoffernes farlighed kan vurderes ved at sammenligne dem med andre lignende stoffer, hvor giftigheden kendes (SARs, struktur-aktivitetssammenhænge). De giver endelig også mulighed for at benytte ekspertvurderinger, bl.a. praktiske erfaringer med brug af stoffet, som grundlag for klassificering /1,57/.

Der lægges fra dansk side og internationalt i EU og OECD vægt på at udvikle alternative metoder, som ikke er baseret på dyr. For nogle egenskaber benyttes allerede laverestående organismer som alger og bakterier. For visse egenskaber er der allerede i dag gode resultater med alternative tests i stedet for dyreforsøg. Der er for nyligt indføjet en testmetode for hudirritation uden brug af levende dyr i reglerne for klassificering /1,57/. Imidlertid er det for mange farlige egenskaber hidtil ikke lykkedes at finde egnede testmetoder, der ikke kræver brug af forsøgsdyr.

1.2 QSAR MODELLER - EN ALTERNATIV METODE TIL FAREVURDERING

Kvantitative struktur-aktivitetssammenhænge (QSAR modeller) kan benyttes til vurdering af farlige egenskaber som alternativ til dyreforsøg. Modellerne er relationer mellem en effekt og molekylære deskriptorer, der er fundet at være knyttet til denne effekt. De kan bruges til at forudsige effekt for stoffer uden testdata ved hjælp af information om de relevante molekylære deskriptorer. Ved hjælp af computeres kapacitet til at gennemgå store mængder af informationer er QSAR modeller i dette projekt blevet benyttet til at vurdere et stort antal stoffer.

Princippet bag struktur-aktivitetssammenhænge er, at stoffer med lignende struktur har lignende egenskaber. SARs og QSARs er velkendte værktøjer til vurdering af kemiske stoffer. Redskaberne benyttes både af USA's og EU's myndigheder samt i industrien til vurdering af fysisk-kemiske, toksikologiske og økotoksikologiske egenskaber og til forudsigelser af stoffers skæbne i miljøet.

Validerede struktur-aktivitetssammenhænge samt ekspertvurderinger er nævnt blandt de datakilder til klassificering, der er opgivet i kriterierne /1,57/. SARs og QSARs er blevet benyttet til klassificering for vandmiljøeffekter i tilfælde, hvor testdata om giftighed eller nedbrydning i vandmiljøet manglede. Med hensyn til klassificering for sundhedseffekter er SARs blevet brugt i konkrete tilfælde, og for nylig er redskabet blevet anvendt ved diskussionen for de to særlige egenskaber, narkotisk effekt og affedtende egenskaber.

1.3 DEN VEJLEDENDE LISTE TIL SELVKLASSIFICERING AF FARLIGE STOFFER

Miljøstyrelsen har arbejdet med QSAR modellering i flere år og området udvikler sig til stadighed. På nuværende tidspunkt råder Miljøstyrelsen over gode modeller for følgende farlige egenskaber udvalgt til dette projekt: Akut dødelig virkning ved indtagelse, allergifremkaldende effekt ved hudkontakt, skader på arveanlæggene, kræftfremkaldende effekt og farlighed for vandmiljøet. Disse modeller identificerer ifølge valideringsresultater stoffer for de respektive egenskaber med en statistisk sikkerhed på ca. 70-85%, afhængig af modellen.

Udgangspunktet for listen har været den europæiske fortegnelse over eksisterende stoffer, EINECS. QSAR modellerne kan af tekniske årsager kun vurdere stoffer, som har en éntydig kemisk struktur, såkaldte diskrete stoffer. Af EINECS-listens ca. 100.000 stofindgange er omkring halvdelen sådanne diskrete organiske stoffer. Miljøstyrelsen har benyttet QSAR modellerne til en systematisk vurdering af ca. 47.000 af disse stoffer, idet de ca. 7.000 kemiske stoffer, der allerede er klassificeret af EU's myndigheder*, ikke indgik i vurderingen.

Kriterierne for QSAR modellernes udvælgelse af stoffer for en given egenskab er blevet defineret så tæt som muligt på kriterierne for klassificering af kemiske stoffer /1/. Hvor der har været mulighed for fortolkning, er der foretaget en afstemning efter bedste skøn med henblik på at få en operativ liste ud. Listens tilblivelse er nærmere beskrevet i del II.

Resultatet af den computerbaserede vurdering er denne Vejledende liste til selvklassificering af farlige stoffer, der indeholder 20.624 kemiske stoffer med vejledende klassificeringer for en eller flere af de udvalgte farlige egenskaber. Listen kan ses på Miljøstyrelsens hjemmeside (www.mst.dk) under "kemikalier".

Miljøstyrelsen ønsker med den Vejledende liste til selvklassificering af farlige stoffer at tilbyde producenter / importører et værktøj, som kan bruges ved selvklassificering af kemiske stoffer for de farlige egenskaber, som listen omfatter. Virksomhederne opfordres til at tage de vejledende klassificeringer i denne liste med i deres vurdering af kemiske stoffer, for hvilke der ikke findes resultater fra dyreforsøg eller andre pålidelige data om de udvalgte farlige egenskaber.

De udvalgte farlige egenskaber og klassificeringer er angivet i tabel 2.

* Disse stoffer kan findes på listen over farlige stoffer /5/. Disse stoffer er så vidt muligt fjernet før vurderingen af stoffene til denne Vejledende liste til selvklassificering af farlige stoffer. Der kan dog forekomme stoffer, som også er på Listen over farlige stoffer. Dette skyldes, at der ikke findes officielle oversigter over, hvilke stoffer gruppeindgangene på Listen over farlige stoffer dækker. En anden grund er, at det samme kemiske stof kan høre under flere CAS numre. Forekommer der stoffer på denne Vejledende liste til selvklassificering af farlige stoffer, som er på Listen over farlige stoffer, skal der ses bort fra den her anførte vejledende klassificering, og stoffet skal klassificeres som angivet på Listen over farlige stoffer.

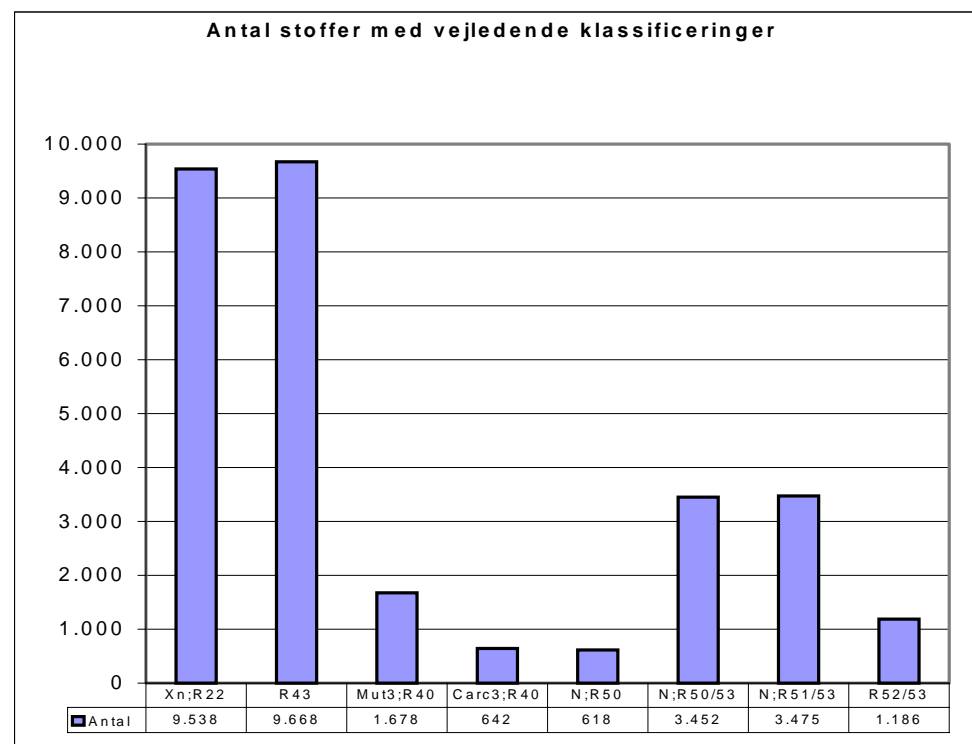
TABEL 2

Farlig egenskab	Klassificering	Ordlyd
Akut dødelig virkning ved indtagelse	Xn;R22	Sundhedsskadelig; Farlig ved indtagelse
Allergifremkaldende virkning ved hudkontakt	R43	Kan give overfølsomhed ved kontakt med huden
Skader på arveanlæggene	Mut3;R40	Mutagen kategori 3; Mulighed for varig skade på helbred
Kræftfremkaldende effekt	Carc3;R40	Kræftfremkaldende kategori 3; Mulighed for varig skade på helbred
Farlighed for vandmiljøet	N;R50	Miljøfarlig; Meget giftig for organismer, der lever i vand
	N;R50/53	Miljøfarlig; Meget giftig for organismer, der lever i vand; kan forårsage uønskede langtidsvirkninger i vandmiljøet
	N;R51/53	Miljøfarlig; Giftig for organismer, der lever i vand; kan forårsage langtidsvirkninger i vandmiljøet
	R52/53	Skadelig for organismer, der lever i vand; kan forårsage langtidsvirkninger i vandmiljøet

For hvert stof på listen er udover de vejledende klassificeringer også angivet EINECS-navn på dansk og engelsk samt EINECS-nummer og CAS nummer.

I figur 1 er vist, hvor mange af listens 20.624 stoffer, der har vejledende klassificeringer indenfor de forskellige farlige egenskaber.

FIGUR 1



Det skal bemærkes, at den Vejledende liste til selvklassificering af farlige stoffer ikke er en udtømmende liste over alle farlige stoffer på EINECS. De stoffer, der er blevet vurderet udgør som før nævnt omtrent halvdelen af stofferne på EINECS.

Vurderingen af stofferne har kun omfattet de farlige egenskaber, hvor Miljøstyrelsen vurderede at have QSAR modeller, der var tilstrækkeligt gode. Stofferne på listen kan således godt have andre farlige egenskaber.

Desuden er der indenfor hver af de udvalgte farlige egenskaber kun medtaget de mest pålidelige forudsigelser. Der vil være vurderede stoffer der, selvom de ikke er medtaget på listen, har en eller flere af de udvalgte farlige egenskaber.

Analogt med dette kan et stof på listen med en vejledende klassificering eksempelvis for allergifremkaldende virkning, men uden vejledende klassificering eksempelvis for kræftfremkaldende effekt, alligevel godt være kræftfremkaldende. En grund til at stoffet ikke står på listen med en vejledende klassificering for kræftfremkaldende effekt kan være, at modellerne for denne egenskab ikke er i stand til at give en god forudsigelse for dette stof.

At et stof ikke er medtaget på listen, eller at det er medtaget, men uden en eller flere vejledende klassificeringer, kan altså enten skyldes at stoffet ifølge modelforudsigelserne ikke besidder disse farlige egenskaber, eller det kan skyldes at modellerne ikke var i stand til at give en god forudsigelse for disse.

For de stoffer som modellerne dækker godt vil alle forudsigelser dog ikke være korrekte. For en vis del af stofferne vil modellerne fejlagtigt forudsige, at stofferne har en given egenskab, som de ikke har (falsk positive). For andre stoffer vil modellerne forudsige, at de ikke har en given egenskab, som de rent faktisk har (falsk negative). Denne liste kan således bruges til at identificere stoffer, som har farlige egenskaber, vel vidende at nogle af disse forudsigelser vil være falsk positive. Hvis listen havde indeholdt negative forudsigelser, ville en del af disse også være ukorrekte (falsk negative). Dermed ville stoffer som i virkeligheden har de farlige egenskaber blive anbefalet ikke at skulle klassificeres for disse.

Denne liste, som kun indeholder positive forudsigelser kan ikke bruges til at "frikende" stoffer for at have farlige egenskaber.

1.4 PRODUCENTERS OG IMPORTØRERS PLIGT TIL AT SELVKLASSIFICERE

Det er producentens eller importørens ansvar at undersøge de kemiske stoffers egenskaber samt at klassificere stofferne for deres iboende farlige egenskaber, før de markedsføres. Denne selvklassificering skal ske på baggrund af tilgængelige oplysninger om stofferne efter kriterierne i klassificeringsbekendtgørelsen /1,57/.

Stoffer, for hvilke der er vedtaget en harmoniseret klassificering, er angivet i Listen over farlige stoffer /5/. For disse ca. 7.000 stoffer skal den vedtagne klassificering følges. Det vil sige, at for de resterende omkring 93.000 ud af de ca. 100.000 stoffer på EU's fortegnelse over eksisterende stoffer /2/ er det importøren / producenten, der har pligt til at vurdere, om stoffet skal klassificeres som farligt (selvklassificering). Selvklassificeringen skal ske i overensstemmelse med kriterierne i klassificeringsbekendtgørelsens bilag 1.

Den vejledende liste til selvklassificering af farlige stoffer er tænkt som et værktøj til hjælp for producenter / importører til at opfylde deres pligt til at klassificere korrekt i de tilfælde, hvor oplysninger om et stof mangler eller er utilstrækkelige. Miljøstyrelsen har ikke i forbindelse med udarbejdelse af listen undersøgt, om der findes data i litteraturen for de enkelte stoffer. Pligten til at kortlægge tilgængelige oplysninger om stoffer, der skal selvklassificeres, påhviler producenter / importører. Pålidelige testresultater eller relevant ekspertviden om konkrete stoffer bør altid anvendes fremfor computerforudsigelser. Hvor sådanne oplysninger, der modsiger denne listes anbefalinger, er tilgængelige, skal de derfor anvendes fremfor denne listes vejledende klassificeringer. Samtidig skal det understreges, at denne vejledende liste kun omfatter nogle af de farlige egenskaber, som producenter / importører skal vurdere et stof for. De øvrige egenskaber indenfor brand- og eksplosionsfare, sundhedsfare og miljøfare skal producenter / importører naturligvis stadig vurdere.

Brug af listen

Listen anbefales at blive anvendt til selvklassificering af stoffer på følgende måde:

1. Undersøg om stoffet er på listen over farlige stoffer /5/. I så fald klassificeres stoffet som den lovpligtige klassificering angiver.
2. Hvis stoffet ikke er på listen over farlige stoffer, klassificeres det efter kriterierne i klassificeringsbekendtgørelsen /1/.
3. I tilfælde af et ellers utilstrækkeligt datagrundlag til denne selvklassificering anbefales at anvende den Vejledende liste til selvklassificering af farlige stoffer.

2 Teknisk beskrivelse af listens tilblivelse og de anvendte QSAR modeller

2.1 INDLEDNING

Indenfor et fagområde, der udvikler sig så hurtigt, som QSAR-området gør i dag, fremkommer der hele tiden bedre modeller, bedre valideringer og modeller for flere effekter – og følgelig opstår der aldrig det ”rette” tidspunkt til at udgive vejledende klassificeringer baseret på QSAR forudsigelser. Der er dog i dag opsamlet betydelig information, som vurderes at kunne være til hjælp i den ellers vanskelige opgave, det er at vurdere farlige egenskaber for mange tusinde ikke-testede kemiske stoffer. Denne viden vil også kunne bidrage i planlægningen af fremtidige teststrategier, så indsatsen kan målrettes til de områder, hvor det er allermest nødvendigt.

2.1.1 SAR / QSAR

Teorien om at kemiske stoffer med lignende strukturer vil have lignende egenskaber er ikke ny. Allerede i 1890'erne blev det eksempelvis opdaget, at den bedøvende (narkotiske) virkning af kemiske stoffer over for akvatiske organismer var relateret til deres forhold mellem opløseligheden i henholdsvis olie og vand. Denne sammenhæng førte til brugen af $\log K_{ow}$ (oktanol/vand fordelingskoefficienten) til forudsigelse af denne effekt. I dag er det velkendt, at alle kemiske stoffer vil udvise en minimums- eller ”basal” narkotisk effekt, som er relateret til deres absorption i cellemembraner, og som kan forudsiges ved deres lipofile (”fedtelskende”) egenskaber.

En SAR model eller en QSAR model ((Quantitative) Structure Activity Relationship) er en relation mellem kemiske stoffers struktur og en givet aktivitet / egenskab. Den modellerede egenskab kan eksempelvis være en fysisk-kemisk egenskab eller en biologisk aktivitet, herunder en toksicitetsparameter.

Ved udviklingen af en model sammenlignes en række kemiske stoffers struktur-egenskaber (molekylære deskriptorer) med målte værdier for den aktuelle egenskab for en gruppe stoffer kaldet et træningssæt. Målet er at fastlægge hvilke deskriptorer, der på en afgørende måde er forbundet med den undersøgte egenskab, og at sætte en relation op mellem disse deskriptorer og egenskaben. Deskriptorerne kan omfatte $\log K_{ow}$, molekylært index, kvantemekaniske egenskaber, form, størrelse, ladning, elektronfordelinger m.m. Sammenligningen udføres ofte ved brug af statistiske metoder.

En SAR model kan give en kvalitativ forudsigelse af, om et stof besidder en given egenskab eller ikke, og en QSAR model kan give en kvantitativ forudsigelse af denne egenskab, eksempelvis en LD_{50} -værdi. Det vil sige, at en QSAR model er en matematisk relation mellem kvantificerede deskriptorer for et kemisk stof og en mere eller mindre gradueret skala for den undersøgte

egenskab. En sådan model kan bruges til ud fra deskriptorer for andre kemiske stoffer at give en kvantitativ forudsigelse for den modellerede egenskab. I dag anvendes sædvanligvis computere til udvikling og brug af QSAR modeller.

2.1.2 Modellernes domæne

En QSAR models domæne begrænser brugen af modellen til den egenskab, der er modelleret, og den gruppe af stoffer, som den kan give pålidelige forudsigelser for. Domænet defineres ved udvælgelsen af træningssættet. De områder af "det kemiske univers", som deskriptorerne i træningssættet "dækker" definerer området af kemiske stoffer, som modellen kan give gode forudsigelser for.

2.1.3 Nøjagtighed af modelforudsigelser

Modeller skal valideres for at fastslå, hvor gode deres forudsigelser er. En validering er en undersøgelse af en models forudsigelser for en gruppe stoffer, som er uafhængig af træningssættet, men som er inden for modellens domæne. Modellens forudsigelser for disse stoffer sammenlignes med målte værdier for den modellerede egenskab for at fastlægge nøjagtigheden / usikkerheden af modelforudsigelserne.

Ideelt set skulle alle modeller valideres ved at blive afprøvet med en gruppe kemiske stoffer, der er uafhængigt af træningssættet. Dette er dog ikke altid simpelt. Dels fordi udelukkelsen af et sådant eksisterende datasæt fra selve modellens træningssæt kan betyde, at værdifuld information ikke bliver anvendt i modellen. Men også fordi det kan være uhyre vanskeligt at afgøre hvordan "eksterne" stoffer relaterer til modellens domæne – om de repræsenterer en vilkårlig udvælgelse der bredt dækker domænet, og dermed kan give et rimeligt billede af modellens præstationsevne.

Disse problemer søges ofte løst ved at anvende én af flere mulige metoder for statistisk krydsvalidering. Det er en meget vigtig metode til at bestemme præstationsevnen for modeller, og i nogle tilfælde, når der kun kan findes få eller ingen testresultater for stoffer, der ikke blev brugt til at udvikle modellen, er denne metode den eneste tilgængelige.

Den valideringsteknik, der oftest anvendes, er den såkaldte "drop one" eller "Q²" metode, hvor ét stof ad gangen fjernes fra træningssættet, og hvor en ny model udviklet på det reducerede træningssæt, bruges til at give en forudsigelse for stoffet. Dette gentages for alle stoffer i træningssættet. Skønt denne metode ofte anvendes, har den en tendens til at føre til en overvurdering af modellens kvalitet.

En mere robust teknik til validering er eksempelvis metoden "3*10% ud", hvor 10% af stofferne i træningssættet fjernes, hvorefter en ny model udviklet på de resterende 90% bruges til at give forudsigelser for de udeladte stoffer. Denne procedure udføres sædvanligvis tre gange. Den kan også udføres indtil alle stoffer har været udeladt fra træningssættet, men tre gange vil generelt være tilstrækkeligt til at bestemme nøjagtigheden af korrelationen /50/.

Resultatet af valideringen af en parametrisk model udtrykkes sædvanligvis som "sensitiviteten", "specificiteten" og "konkordansen" af modellen. Sensitiviteten er et mål for, hvor effektivt modellen "fanger" kemiske stoffer, som har den modellerede egenskab. En sensitivitet på 80% betyder, at modellen forudsiger 80% af de "ægte positive" i valideringssættet som

positive, og at de resterende 20% ikke-korrekt forudsiges at være negative (falsk negative). Specificiteten er et mål for hvor mange falsk positive forudsigelser modellen giver. En specificitet på 80% betyder, at modellen forudsiger 80% af de ”ægte negative” i valideringssættet som negative, og at de resterende 20% ikke-korrekt forudsiges at være positive (falsk positive). Konkordansen er et overordnet mål for korrektheden af forudsigelserne. En konkordans på 80% betyder, at modellen forudsiger 80% af stofferne i valideringssættet korrekt som positive eller negative, og at modellen giver ikke-korrekte forudsigelser for de resterende 20% (falsk negative og falsk positive).

Modellens evne til at give korrekte forudsigelser afhænger både af den anvendte metode og den modellerede egenskab. Generelt kan gode QSAR modeller give korrekte forudsigelser for 70-85% af de kemiske stoffer den adspørger om, under forudsætning af at de er indenfor modellens domæne /53,54/. Dette gælder også de modeller, der blev anvendt i dette projekt. Naturligvis kan en model aldrig blive mere nøjagtig end de testdata, som den blev baseret på. Derfor er det uhyre vigtigt at være bevidst om nøjagtigheden og reproducerbarheden af de testdata, der anvendes til at lave en model. Hvis en biologisk test giver forkerte resultater i 17% af tilfældene, vil den ”perfekte” model baseret på disse testdata også give forkerte forudsigelser i 17% af tilfældene.

Udover at vurdere modellens evne til at give korrekte forudsigelser er det også nødvendigt at overveje i hvilken sammenhæng, modellen skal anvendes. I nogle tilfælde vil et stort antal ”falsk positive” eller ”falsk negative” være acceptabelt, mens det i andre tilfælde vil være uacceptabelt. I dette projekt var der ingen bevidste forsøg på at justere modellernes vægt mod falske forudsigelser i nogen retning. Målet var opnå de mest korrekte forudsigelser for de undersøgte stoffer. Alternativet, når der ikke findes testdata eller andre relevante informationer om et stof, har ofte været at det ikke er blevet vurderet.

2.1.4 Software

I dag eksisterer der et stort antal computerprogrammer, som kan anvendes til at lave og bruge QSAR modeller for en lang række effekter fra biondbrydelighed til kræft. Blandt disse er ”fragment-baserede” statistiske systemer som TOPKAT og M-CASE, såvel som systemer til tre-dimensionel modellering af ”ligand docking”^{**} såsom COMFA (Comparative Molecular Field Analysis). OASIS, en sofistikeret programpakke af systemer, der er i stand til at estimere en lang række effekter ved brug af 3D- og kvantemekaniske parametre, skal også nævnes i denne sammenhæng /46,47/. Systemet anvendes for øjeblikket blandt andet til estimering af kemiske stoffers bindingsaffiniteter til østrogen receptorer /48/.

Den metode, som programmerne anvender i udviklingen og anvendelsen af QSAR modeller er i virkeligheden ikke ny. Programmerne grupperer ganske simpelt stoffer, der ligner hinanden mht. struktur og ”globale” (for hele molekylet) eller ”lokale” (for en del af molekylet) parametre såsom $\log K_{ow}$ og elektrofilicitet på en måde, som eksperter stort set ville kunne gøre det. Men

* I fragment-baserede programmer baseres forudsigelsen på forekomsten af molekyllære understrukturer.

** I en model for ”ligand docking” ønskes en forudsigelse af, hvor godt et kemiske stof passer ind i en bestemt rumlig struktur af et makromolekyle med biologisk betydning, såsom en receptor (f.eks. en hormonreceptor).

programmerne udfører det med meget høj hastighed, og de er i stand til at tage hensyn til et stort antal faktorer samtidig (såsom kritiske inter-atomiske afstande), og at gøre det fejlfrit. Dermed kan de hjælpe eksperter til at finde hidtil ikke erkendte sammenhænge. Derudover kan eksempelvis de nedenfor beskrevne programmer TOPKAT og M-CASE efterligne en anden menneskelig egenskab, nemlig at afvise at give estimater for et stof, hvor der simpelthen ikke er nok information i modellen til at give en pålidelig forudsigelse - dvs. at stoffet er udenfor modellens domæne. Programmerne udfører blot dette ved iterative statistiske metoder i stedet for ved menneskelig intelligens eller intuition.

M-CASE

M-CASE er et erfarings-baseret kunstigt intelligenssystem, der er i stand til at lære direkte fra data. Modeller sat op i dette program kan forudsige de modellerede toksiske egenskaber på basis af entydige fragmenter i stoffernes kemiske struktur. Disse fragmenter er i udviklingen af modellen fundet at være statistisk relevante for en specifik biologisk aktivitet, enten ved at forøge eller nedsætte denne aktivitet. Programmet kan dermed give en "kemisk" forklaring på observerede biologiske egenskaber. Antagelsen i programmet er, at tilstedeværelsen af fragmenter, som i træningssættet er overrepræsenterede i de aktive stoffer indikerer, at fragmentet potentielt er relateret til aktiviteten. Denne fragment-baserede metode antages at være en fornuftig basis til vurdering af aktiviteten af andre kemiske stoffer. På basis af tilstedeværelsen af fragmenter i den kemiske struktur for et stof vil programmet beregne en værdi for den egenskab, der er modelleret. Dette sker ved hjælp af "lokale QSARs" for de forskellige fragmenter, samt ved eventuelle "globale QSARs", hvis det er fundet, at der eksisterer en generel relation mellem en parameter og den modellerede egenskab, som det eksempelvis er tilfældet for $\log K_{ow}$ og toksiciteten overfor akvatiske organismer. Programmet giver en advarsel, hvis der er fragmenter i det forespurgte stof, som ikke findes i træningssættet for modellen. Med advarslen indikerer programmet, at det forespurgte stof er udenfor modellens domæne /38,43/. Forudsigelser for stoffer, der er indenfor modellens domæne, og som modellen kan give pålidelige forudsigelser for, benævnes i denne rapport som AOKs ("All OK stoffer").

TOPKAT

TOPKAT vurderer toksiciteten af kemiske stoffer ud fra deres kemiske struktur ved brug af QSTRs (Quantitative Structure Toxicity Relationships) modeller for specifikke skadelige sundhedseffekter /56/. Når en model i programmet forespørges ved at give modellen stoffets kemiske struktur, vil programmet først bestemme hvilken kemisk klasse stoffet tilhører. Dernæst vil programmet beregne de deskriptorer, der er nødvendige for beregningerne i en undermodel for denne kemiske klasse. Deskriptorerne kan for eksempel være elektro-topografisk tilstand, kappa index, molekylvægt, og symmetriindex. Programmet kontrollerer om alle de fragmenter, der er til stede i det forespurgte kemiske stof, var til stede med tilstrækkelig frekvens i træningssættet for modellen. Hvis der ikke er nogen fragmenter, der ikke er dækket tilstrækkeligt af modellen, checker programmet dernæst om forespørgslen er indenfor modellens domæne mht. den specifikke ligning. Hvis det er tilfældet gennemsøges den aktuelle lignings træningssæt for de stoffer, der mht. struktur ligner det forespurgte kemiske stof mest. Programmet kan bestemme overensstemmelsen mellem faktiske testdata og modellens forudsigelser for de stoffer, der kommer tættest på det forespurgte stof /45/. Hvis der er overensstemmelse for de fire mest statistisk relevante

stoffer, accepteres modellens forudsigelse og benævnes i denne rapport som AOK analogt med forudsigelserne fra M-CASE.

Epiwin

Denne gruppe programmer udviklet af Syracuse Research Corporation blev i dette projekt anvendt til at beregne tre økotoksikologiske parametre: Bionedbrydning, $\log K_{ow}$ og biokoncentration. I modsætning til TOPKAT og M-CASE, forsøger Epiwin ikke at vurdere om stofferne er inden for eller uden for modellernes domæne, og alle forudsigelser blev anvendt som givet af programmet.

Chem-X

I dette program er det muligt at estimere et stort antal forskellige fysisk-kemiske egenskaber, udvikle 2D og 3D-QSARs samt at lagre store mængder data og kemiske strukturer i databaser.

Miljøstyrelsen har opbygget en database i Chem-X, som indeholder QSAR forudsigelser for omkring 166.000 kemiske stoffer /55/, dækkende næsten alle de organiske stoffer på EINECS med entydig kemisk struktur, i alt ca. 47.000 stoffer. Databasen indeholder forudsigelser for en række forskellige egenskaber, der dækker både sundheds- og miljø-effekter. Disse forudsigelser har dannet baggrund for de vejledende klassificeringer på den Vejledende liste til selvklassificering af farlige stoffer. Programmet giver detaljerede muligheder for at foretage søgninger, definere fremvisningen af data, samt at håndtere kemiske strukturer. Dette værktøj blev i stort omfang anvendt til at sammenligne testdata, forudsigelser og udvalgte understrukturer i ekspertvurderingerne af QSAR modellerne.

Mulighederne for udbredelse af denne database med de specifikke QSAR forudsigelser er på nuværende tidspunkt uklare grundet spørgsmål om copyright.

2.2 METODOLOGIEN I UDARBEJDELSEN AF LISTEN

2.2.1 De udvalgte farlige egenskaber

Følgende farlige egenskaber blev udvalgt til evaluering ved QSAR modeller:

- Akut dødelig virkning ved indtagelse
- Allergifremkaldende effekt ved hudkontakt
- Skader på arveanlæggene
- Kræftfremkaldende effekt
- Farlighed for vandmiljøet

2.2.2 De vurderede kemiske stoffer

Det overordnede formål med projektet var at vurdere så mange som muligt af de kemiske stoffer på EINECS (European Inventory of Existing Commercial Chemical Substances) /2/. EINECS består af 100.116 indgange indeholdende organiske eller uorganiske stoffer i enten enkeltstofindgange eller i indgange, der dækker grupper af stoffer.

Evalueringen måtte af tekniske grunde begrænses til at dække organiske stoffer med en entydig kemisk struktur, da det ikke er praktisk muligt at lave/bruge modeller, når man ikke kender stoffets struktur. Det vil sige, at UVCB'er

(Unknown, Variable Composition and Biologicals) og andre dårligt definerede strukturer eller blandinger måtte udelades. Undtagelser blev dog foretaget, hvor det blev vurderet rimeligt. Eksempelvis indgik C12-C16 n-alkoholer i vurderingen som C14 og hydrogenchlorid-salte indgik som modermolekylet osv.

Uorganiske stoffer er ligeledes ikke blevet vurderet. Disse stoffer kan ofte vurderes ved simple metoder for tilgængeligheden af de respektive an- og kationer med velkendte toksicitetsprofiler. "Organo-metaller" blev også udeladt fra vurderingen som dårlige kandidater for modellering. Endelig omfattede vurderingen, på grund af ressourcer, kun kemiske stoffer, hvor oplysninger om de tredimensionelle strukturer var tilgængelige [7].

Så vidt muligt blev alle stoffer, der allerede er klassificerede på den formelle liste i EU (Listen over farlige stoffer), ved brug af CAS nummer sammenligninger udeladt fra vurderingen, idet disse stoffer aldrig skal undergå selvklassificering.

Dette resulterede i en "startliste" på 46.707 kemiske stoffer ud af Einecs fortegnelsens 100.116 stofindgange, som kunne undergå en systematisk vurdering ved hjælp af QSAR modeller.

2.2.3 Testresultater

For langt størstedelen af de vurderede kemiske stoffer var eksperimentelle testresultater ikke tilgængelige i vurderingen. Men hvor de var tilgængelige som en del af de anvendte modeller, blev de generelt brugt fremfor modellernes forudsigelser.

Det er vigtigt at understrege, at der ikke blev gjort nogen forsøg på at søge testresultater for stofferne i alverdens publicerede og ikke-publicerede datasamlinger over toksikologiske oplysninger, for at fastlægge om det var nødvendigt at foretage QSAR forudsigelser. Kortlægning af eksperimentelle testresultater er en opgave, som påhviler producenter / importører af de kemiske stoffer.

2.2.4 Brug af QSAR modeller

De tekniske specifikationer for de anvendte QSAR modeller og en beskrivelse af hvordan modellerne er blevet brugt til en systematisk vurdering af stofferne, er givet i de følgende fem kapitler for de udvalgte fem farlige egenskaber.

Det skal betones, at de anvendte modeller ikke direkte giver forudsigelser af "klassificeringer", men at de giver forudsigelser om biologiske aktiviteter, som kan føre til en klassificering. Derfor blev der på basis af de formelle kriterier for klassificering og de tilgængelige modeller for de individuelle farlige egenskaber opsat "regler" for, hvordan modelforudsigelserne for biologisk aktivitet skulle anvendes til fastsættelse af vejledende klassificeringer. Anvendelsen af fastsatte "regler" gjorde det muligt at foretage en systematisk vurdering af det store antal stoffer på startlisten. Sådanne regler er ikke perfekte, men dybest set adskiller processen sig ikke fra den, der pålægges eksperter, når de skal selvklassificere et stof, hvor der ikke findes de ønskede testresultater.

Kun de modelforudsigelser der opfyldte følgende kriterier blev anvendt:

For TOPKAT modeller skulle forudsigelserne være indenfor modellens domæne og de fire mest statistisk relevante stoffer refereret til af modellen skulle have en acceptabel overensstemmelse mellem testresultat-værdi og forudsigelse. De forudsigelser, der opfylder dette kriterie benævnes AOKs.

For M-CASE modeller skulle forudsigelserne være indenfor modellernes domæne, hvilket betyder, at der ikke må være fragmenter for det undersøgte stof, som modellen ikke kender. Modellen skulle være i stand til at give en pålidelig forudsigelse på baggrund af kendskab til alle stoffets fragmenter.

Som beskrevet i kapitlerne for de individuelle farlige egenskaber er der så vidt muligt indenfor projektets rammer foretaget ekspert evaluering af sandsynligheden for, at QSAR modellernes forudsigelser var korrekte. Evalueringen omfattede undersøgelse af forudsigelsernes overensstemmelse med biologiske aktiviteter og kemiske egenskaber. Der er ikke blevet foretaget dybdegående toksikologiske vurderinger af de individuelle kemiske stoffer. Tvivlsomme forudsigelser indenfor en af de udvalgte farlige egenskaber medførte, at stoffet ikke blev tildelt vejledende klassificering for den pågældende farlige egenskab.

Omfanget af ekspert evalueringerne varierede for de forskellige farlige egenskaber. Generelt blev den største indsats lagt på en vurdering af forudsigelserne for skader på arveanlæggene og kræftfremkaldende effekt, og den mindste indsats blev lagt på vurderinger af forudsigelser om allergifremkaldende virkning og farlighed for vandmiljøet.

2.2.5 Resultatet

Det er vigtigt at understrege, at resultaterne af den systematiske vurdering som angivet på den Vejledende liste til selvklassificering af farlige stoffer, kun repræsenterer POSITIVE QSAR forudsigelser. Der er i udarbejdelsen af listen ikke skelnet mellem, om der for et stof var en negativ forudsigelse eller en upålidelig forudsigelse (en ikke-AOK forudsigelse), som blev udeladt.

At et QSAR vurderet stof ikke er med på listen, eller at det er angivet på listen men uden vejledende klassificeringer for en eller flere af de udvalgte farlige egenskaber, kan altså enten skyldes, at modelforudsigelserne angav, at stoffet ikke havde den / disse farlige egenskaber, eller det kan skyldes, at modellerne ikke var i stand til at give en pålidelig forudsigelse for, om stoffet havde den / de farlige egenskaber.

Listen kan derfor ikke bruges til at konkludere, at et specifikt stof ifølge QSAR modellerne ikke besidder de farlige egenskaber. Afhængigt af hvilken farlig egenskab, der blev vurderet, var mellem 5% og 65% af forudsigelserne upålidelige.

2.3 AKUT DØDELIG VIRKNING VED INDTAGELSE

EU's kriterie for klassificering

Det formaliserede kriterie for klassificering for akut dødelig virkning ved indtagelse indeholder forskellige muligheder for tests, herunder test med den kritiske dosis (fastdosismetoden), og for tolkning af de forskellige kilder til oplysninger om akut dødelig virkning ved indtagelse. Klassificeringen baseres dog ofte på akutte LD₅₀ testresultater i rotter, for hvilke følgende klassificeringskriterier anvendes:

TABEL 3

Klassificeringskriterie	Klassificering
LD ₅₀ oral, rotte ≤ 25 mg/kg	Tx;R28 (meget giftig; meget giftig ved indtagelse)
25 mg/kg < LD ₅₀ oral, rotte ≤ 200 mg/kg	T;R25 (giftig; giftig ved indtagelse)
200 mg/kg < LD ₅₀ oral, rotte ≤ 2000 mg/kg	Xn;R22 (sundhedsskadelig;farlig ved indtagelse)

Vurdering baseret på modelforudsigelser

Stoffer med forudsigelser eller testresultater, der viste at LD₅₀ oral, rotte ≤ 2000 mg/kg, blev medtaget på listen med den vejledende klassificering Xn;R22. Der blev ikke gjort forsøg på at differentiere mellem de tre niveauer for farlighed, og det er vigtigt at bemærke, at den vejledende klassificering derfor ofte vil være mindre stringent end hvis klassificeringen var baseret direkte på testresultater.

Hvis testresultater var lettilgængelige (havde været brugt til at udvikle modellen) blev disse anvendt frem for modelforudsigelser.

Da testresultater for akut dødelig virkning ved indtagelse i nogle tilfælde var tilgængelige fra forsøg med mus, hvor stofferne var indtaget gennem andre veje end oralt, blev disse brugt til at forudsige rotte oral LD₅₀ ved brug af følgende QSARs angivet efter præference /8,9/:

TABEL 4

1. $\text{Log LD}_{50} \text{ oral, rotte} = 0,731 + 0,841 * (\text{Log LD}_{50} \text{ oral, mus})$ RTECS data 1989, n=3919, R ² = 0,750, Q ² = 0,749
2. $\text{Log LD}_{50} \text{ oral, mus} = 0,682 + 0,373 * (\text{Log LD}_{50} \text{ iv, mus}) + 0,518 * (\text{Log LD}_{50} \text{ ip, mus})$ RTECS data 1994, n = 286, R ² = 0,766, Q ² = 0,764
3. $\text{Log LD}_{50} \text{ oral, mus} = 0,731 * (\text{Log LD}_{50} \text{ ip, mus})$ RTECS data 1994, n=286, R ² = 0,724, Q ² = 0,724
4. $\text{Log LD}_{50} \text{ oral, mus} = 0,945 + 0,802 * (\text{Log LD}_{50} \text{ iv, mus})$ RTECS data 1994, n=286, R ² = 0,689, Q ² = 0,688

iv: Intravenøst (i vene)

ip: Intraperitoneal (i bughinde)

Testresultater for LD₅₀ i enten mus eller rotter var tilgængelige for lige over 10% af de vurderede kemiske stoffer. For de stoffer uden sådanne testresultater blev rotte, oral LD₅₀ beregnet ved en TOPKAT model (v. 5.01). Ifølge TOPKAT indeholder modellen ca. 4000 stoffer, og deres egen krydsvalidering af denne model viser, at 86-100% af modellens forudsigelser falder indenfor en faktor fem, når de sammenlignes med testresultater /10/.

Miljøstyrelsens egen eksterne evaluering af modellen, hvor der blev brugt 1.840 kemiske stoffer, som ikke var del af modellens træningssæt, gav et noget ringere resultat; R² = 0,31. Ifølge denne vurdering falder 86% af modellens forudsigelser indenfor en faktor ti, når de sammenlignes med testresultater /11/. Fordelingen kan ses i tabel 5.

TABEL 5

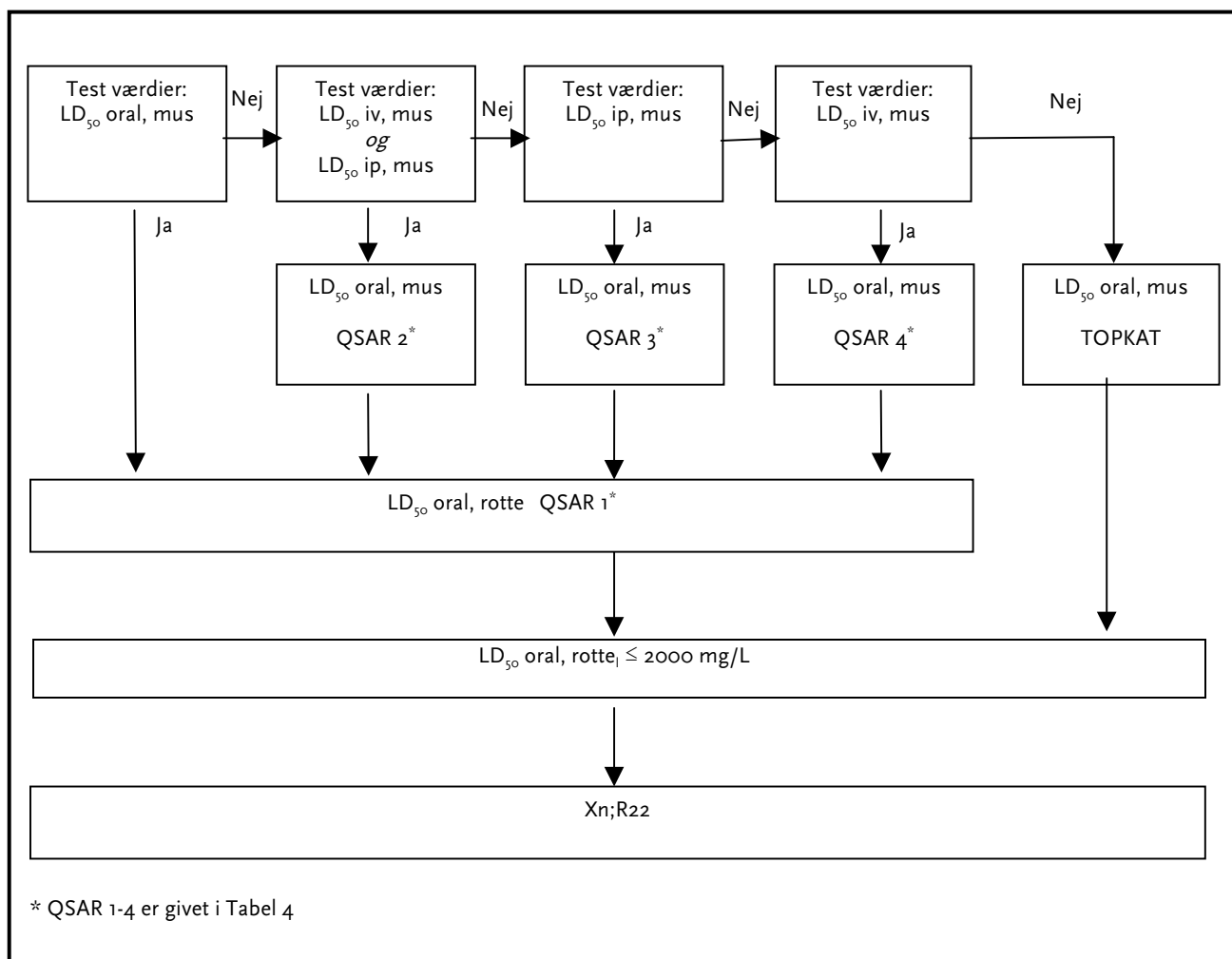
Modelforudsigelser indenfor en faktor:	%	N (kumulativ)
2	42	671
4	67	1069
6	78	1235
8	83	1323
10	86	1368

I moderne LD₅₀ forsøg med anvendelse af få forsøgsdyr er den statistiske variation ofte i størrelsesordenen 2-4, og imellem forskellige laboratorier er en faktor 10 i forskel i de fundne LD₅₀ værdier ikke ualmindelig /12/. Selvom TOPKAT modellen, som det fremgår, ikke er perfekt, anses dens forudsigelser at være tilstrækkeligt tilnærmede til at give den mindst strikte klassificering for akut dødelig virkning ved indtagelse, Xn;R22. Derimod vurderes modellens nøjagtighed ikke at være tilstrækkelig til at differentiere mellem de tre klassificeringsniveauer ("sundhedsskadelig", "giftig" og "meget giftig"). Det er derfor vigtigt at bemærke, at der vil være nogle stoffer med den vejledende klassificering Xn;R22 på listen, som udfra eksempelvis dyreforsøg vil vise sig at skulle klassificeres som T;R25 eller Tx;R28.

Ud af de stoffer, hvor TOPKAT modellen var i stand til at give en pålidelig forudsigelse (AOK), blev 57% af stofferne forudsagt at have en akut LD₅₀ oral, rotte på ≤ 2000 mg/kg. Procentdelen af stoffer med en akut LD₅₀ oral, rotte på ≤ 2000 mg/kg ud af 12.632 stoffer fundet i Registeret for toksiske effekter af kemiske stoffer, RTECS (RTECS 1998) /52/ var 61%. At procentdelen for TOPKATs forudsigelser og resultaterne i RTECS er så tæt på hinanden, er ikke overraskende, idet RTECS data var den primære kilde til testresultater til træningssettet for TOPKAT modellen.

En skematisk oversigt over den systematiske vurdering er givet i figur 2.

FIGUR 2 DEN SYSTEMATISKE VURDERING



iv: Intravenøst (i vene)
ip: Intraperitoneal (i bughinde)

Ca. 10.200 stoffer havde ifølge beregningerne en akut LD₅₀ oral, rotte på mindre end eller lig med 2.000 mg/kg*. Omkring 700 ud af disse blev ikke medtaget på den Vejledende liste til selvklassificering af farlige stoffer. Dette skyldtes, at en ekspertvurdering fandt, at det eksempelvis var aminosyre- og proteinstoffer, som sandsynligvis bliver nedbrudt af mavesyre før optagelse, eller at det var stoffer, hvor det forventes, at optagelsen vil være ringe. Dette resulterede i, at 9.538 stoffer blev medtaget på listen med den vejledende klassificering Xn;R22.

* TOPKAT beregninger blev også udført for rotters kroniske LOAEL (lavest observerede skadelige effekt-niveau). Nøjagtigheden bestemt ved krydsvalidering var i samme størrelsesorden som for den anvendte model for akut giftighed, med 95% af forudsigelserne indenfor en faktor 3-5 af de målte værdier /13,44/. Der er dog ingen EU klassificeringer, der er relaterede specifikt til denne farlige egenskab (men snarere til "alvorlig effekter på morfologi eller gifteffekter efter gentagen dosering", R48). Der blev derfor ikke foreslået nogen klassificering for denne farlige egenskab.

2.4 ALLERGI-FREM-KALDENDE EFFEKT VED HUDKONTAKT

EU's kriterie for klassificering

Klassificering som allergifremkaldende ved hudkontakt, R43 ("Kan give overfølsomhed ved kontakt med huden"), er baseret enten på dyrestudier eller på praktisk erfaring, eller på en kombination af begge dele. Dyrekriteriet er baseret på test med eller uden hjælpestof.

Der eksisterer flere forskellige metoder med hjælpestof til undersøgelse for hudsensibiliserende virkning, men Magnusson-Kligmann's metode (GPMT: Guinea Pig Maximization Test) er den foretrukne. For denne test resulterer respons i 30% af forsøgsdyrene i klassificering. For test uden hjælpestoffer (eksempelvis Büehler test) anses respons i 15% af forsøgsdyrene som et positivt resultat. Humane testresultater kan stamme fra lappetest, studier af enkeltestudier eller epidemiologiske studier.

Vurdering baseret på modelforudsigelser

Der blev anvendt to fremgangsmåder til vurdering af hudsensibilisering /14,15/.

I den første fremgangsmåde blev to TOPKAT QSTR modeller anvendt. Den første model blev anvendt til at forudsige om stoffet har allergifremkaldende effekt ved hudkontakt, og for de stoffer, hvor forudsigelserne var positive, blev den anden model anvendt til at forudsige "Stærk eller svag/moderat allergifremkaldende effekt". Modellerne er primært udviklet over testresultater fra Magnusson-Kligmann's metode (GPMT). Kun de stoffer, der ifølge modelforudsigelserne var stærkt allergifremkaldende, blev anset for med sandsynlighed at opfylde EU's kriterie for klassificering med R43.

I den anden fremgangsmåde blev en M-CASE model anvendt til forudsigelse af hudsensibiliserende virkning. Træningssættet, som er brugt til at udvikle M-CASE modellen, adskiller sig en del fra træningssættet til TOPKAT modellerne, da der i M-CASE modellen er anvendt testresultater fra både GPMT og humane resultater. For M-CASE modellen blev kun positive forudsigelser af allergifremkaldende effekt med en M-CASE score på ≥ 40 (svarende til "meget aktiv") anset for med sandsynlighed at opfylde EU's kriterie for klassificering med R43.

TABEL 6 DE ANVENDTE MODELLER

Model	Tekniske specifikationer
TOPKAT (v. 5.01 1998) No sensitization v. any	n=389 GPMT Krydsvalidering med "Q ² " metoden viste /14/: Sensitivitet 84-94% Specificitet 87-96%
TOPKAT (v. 5.01 1998) Strong v. weak/Moderate	n=266 GPMT Krydsvalidering med "Q ² " metoden viste /14/: Sensitivitet 88-96% Specificitet 88-98%
M-CASE (v. 3.320 1999) Model A33: Allergic contact dermatitis	n=1034 GPMT eller resultater fra human erfaring Krydsvalidering med "3*10% ud" metoden viste /15/: Sensitivitet 69 – 89% Specificitet 89– 94% Chi ² > 50, p < 0,0001

Eksterne valideringer af både TOPKAT og M-CASE modellerne blev forsøgt ved brug af fortrolige testresultater fra EU's program for nye kemiske stoffer. I valideringen af den todelte TOPKAT model (n = 64 AOK forudsigelser) blev

67% af de hudsensibiliserende stoffer og 77% af de ikke-hudsensibiliserende stoffer korrekt identificeret. For M-CASE (n = 75 AOK forudsigelser) blev 45% af de hudsensibiliserende stoffer og 81% af de ikke hudsensibiliserende stoffer korrekt identificeret /16/.

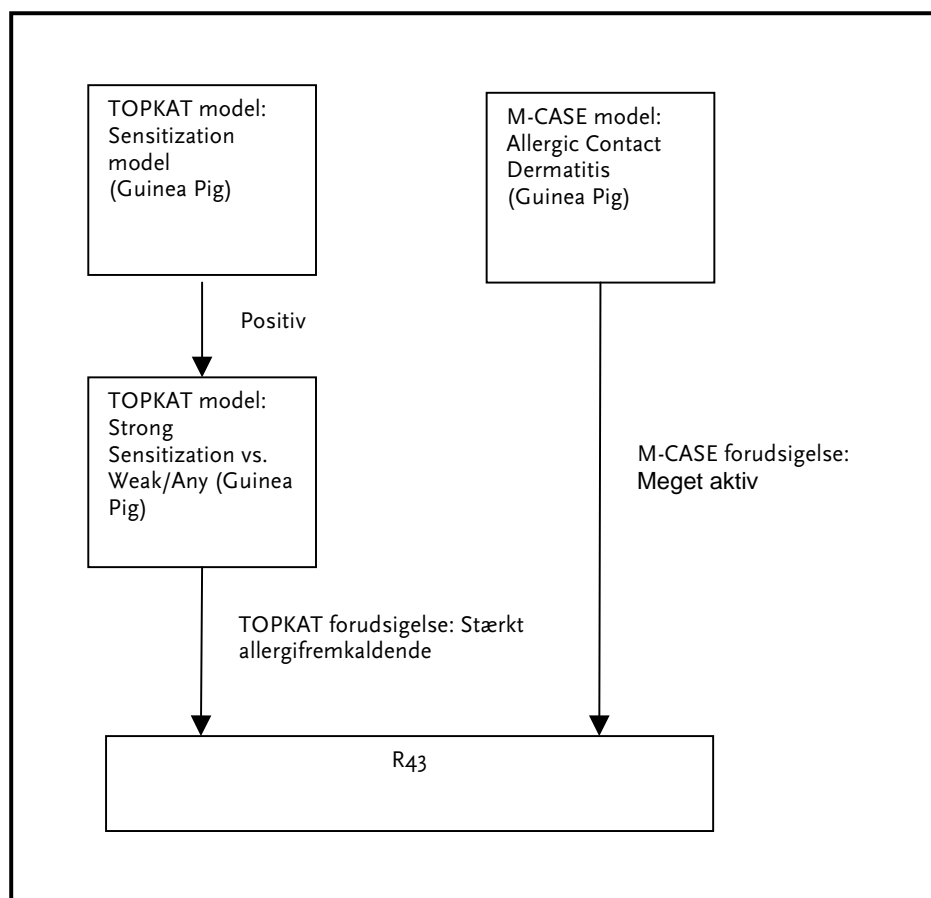
Det er vanskeligt at afgøre hvor repræsentative de nye kemiske stoffer er for ”kemikalieuniverset”. Generelt har nye kemiske stoffer mere komplekse strukturer og højere molekylvægt. Den måske mest overraskende side af denne valideringsøvelse var opdagelsen af, at der ud af mere end tre tusinde nye anmeldte kemiske stoffer kun var en lille procentdel af stofferne, der havde brugbare testresultater for allergifremkaldende effekt ved hudkontakt.

Kun stoffer, der havde AOK forudsigelser i de anvendte modeller, blev taget i betragtning i den regelbaserede vurdering. Det blev overvejet at anvende et kriterium om, at stofferne skulle have positive forudsigelser fra både TOPKAT og M-CASE modeller for at forøge sikkerheden i vurderingen, men dette viste sig ikke at virke efter hensigten. Ikke fordi der ikke var overensstemmelse mellem modelforudsigelserne, men fordi modellernes domæner var væsentligt forskellige, og der derfor var et stort antal stoffer, der kun havde en AOK forudsigelse i en af modellerne.

Der blev ikke gjort nogen forsøg på systematisk at ekspertvurdere forudsigelserne af allergifremkaldende effekt ved hudkontakt.

En skematisk oversigt over den systematiske vurdering er givet i figur 3.

FIGUR 3 DEN SYSTEMATISKE VURDERING



9.668 kemiske stoffer blev i den systematiske vurdering efter de beskrevne kriterier fundet at have allergifremkaldende effekt ved hudkontakt, og de blev medtaget på den Vejledende liste til selvklassificering af farlige stoffer med R43. Dette antal kan virke temmelig stort for mange eksperter, og selv om de anvendte modeller repræsenterer "state-of-the-art", kan det være udtryk for, at de er meget følsomme. Det er dog meget svært at finde nogen pålidelige indikationer af, hvor mange af de eksisterende kemiske stoffer, der ville være hudsensibiliserende, hvis de blev testede på dyr eller mennesker. Forskellige vurderinger af hvor stor en procentdel af stofferne på EINECS, der er allergener, har strakt sig fra 5-25% med en vis præference for en procentdel på 10%. Ser man på Annex I stoffer (Listen over farlige stoffer) er der i dag ca. 10% af stofferne, der er klassificerede med R43. Det er dog ikke muligt at vurdere, om denne procentdel er repræsentativ for alle de eksisterende stoffer. Positiv bias kan være tilstede, fordi kemiske stoffer, der ved test er fundet at være positive, er overrepræsenterede. Negativ bias kan være forårsaget af, at de fleste af stofferne på Annex I aldrig er blevet testet for denne effekt. Spørgsmålet står åbent.

2.5 SKADER PÅ ARVEANLÆGGENE

EU's kriterie for klassificering

Indenfor denne farlige egenskab er klassificeringen opdelt i tre kategorier:

Klassificering som mutagen, kategori 1 (mut1;R46, Kan forårsage arvelige genetiske skader) er baseret på bevis på en årsagssammenhæng mellem udsættelse for stoffet og forekomst af arvelige skader på det genetiske materiale hos mennesker.

Klassificering som mutagen, kategori 2 (mut2; R46, Kan forårsage arvelige genetiske skader) er baseret på dyreforsøg, der viser mutagenicitet i kønsceller enten fra forsøg på kønsceller eller fra forsøg, hvor der er demonstreret mutagene effekter i somatiske celler *in vivo* eller *in vitro* såvel som etableret bevis for, at stoffet vil nå kønscellerne.

Klassificering som mutagen, kategori 3 (mut3;R40, Mulighed for varig skade på helbred) er baseret på enten *in vivo* mutagenicitetstest eller testresultater for cellulære påvirkninger med *in vitro* forsøg som understøttende bevis. For denne klassificering er det ikke nødvendigt at demonstrere kønscellemutationer.

Vurdering baseret på modelforudsigelser

Der blev anvendt en række modeller til vurdering af mutagenicitet. De forskellige modeller dækker et antal forskellige genotoksiske effekter. Forudsigelse af induktion af mikronuclei i *in vivo* forsøg blev anvendt i den systematiske vurdering til at demonstrere kromosomskader i somatiske celler *in vivo*. De resterende anvendte modeller giver forudsigelser for genotoksicitet *in vitro*. Positive testresultater i *in vitro* forsøg for genotoksicitet ville sædvanligvis føre til en klassificering for mutagenicitet, kategori 3. I den systematiske vurdering blev modelforudsigelserne for *in vitro* genotoksicitet anvendt som understøttende bevis for *in vivo* forudsigelserne af induktion af mikronuclei.

TABEL 7 DE ANVENDETE MODELLER

Model	Tekniske specifikationer
M-CASE (v. 3.320 1999) Model A2E: Structural Alerts for DNA Reactivity	n=784 Krydsvalidering ved "3*10% ud" metoden viser /24/: Sensitivitet 85-98% Specificitet 60-69% Chi ² >22, p < 0,0001
M-CASE (v. 3.320 1999) Model A62: Induction of Micronuclei	n=238 GeneTox testresultater Krydsvalidering ved "3*10% ud" metoden viser /30/: Sensitivitet 80-100% Specificitet 50-70% Chi ² >4, p < 0,05
TOPKAT (v. 3.01, 1998) Salmonella (Ames) Mutagenicity,	n=1866 Krydsvalidering ved "Q" metoden viser /25/ at de 10 sub-moduler har sensitiviteter og specificiteter på 75-100%. Ekstern evaluering (Miljøstyrelsen 1998, n=118) viste /26/: 82% korrekte negative forudsigelser og 76% korrekte positive forudsigelser
M-CASE (v. 3.320 1999) Model A2H: Salmonella (Ames) Mutagenicity	n=2034 NTP eller GeneTox testresultater Krydsvalidering ved "3*10% ud" metoden viser /27/: Sensitivitet 75-78,5% Specificitet 78,2-90% Chi ² >150, p < 0,0001
M-CASE (v. 3.320 1999) Model A61: Chromosomal Aberrations	n=233 NTP testresultater i dyrkede CHO celler Krydsvalidering ved "3*10% ud" metoden viser /28/: Sensitivitet 44-80% Specificitet 50-80% Chi ² < 2, p > 0,15 (Yderligere validering under udarbejdelse)
M-CASE (v. 3.320 1999) Model A2F: Mutations in Mouse Lymphoma	n=210 NTP thymidine kinase i L5178Y celler Krydsvalidering ved "3*10% ud" metoden viser /29/: Sensitivitet 64-100% Specificitet (not determined) Chi ² ≈ 2, p = 0,15 (yderligere validering under udarbejdelse)

For klassificering baseret på målte testresultater er et positivt resultat for induktion af mikronuclei i et *in vivo* forsøg tilstrækkeligt grundlag til at klassificere som mutagen i kategori 3. Da grundlaget i denne systematiske vurdering er modelforudsigelser, blev forudsigelserne af induktion af mikronuclei understøttet ved at inkludere modeller for en række andre indikatorer for genotoksicitet.

Dette skal ikke ses som et udtryk for, at det foreslås, at positive *in vivo* testresultater skal understøttes af *in vitro* testresultater i klassificeringskriterierne. Men det er et udtryk for, at det blev vurderet, at de vejledende klassificeringer skulle baseres på mere end modelforudsigelser om induktion af mikronuclei alene. Disse blev derfor understøttet af forudsigelser fra andre relevante QSAR modeller for at forøge sandsynligheden for, at de positive forudsigelser var korrekte.

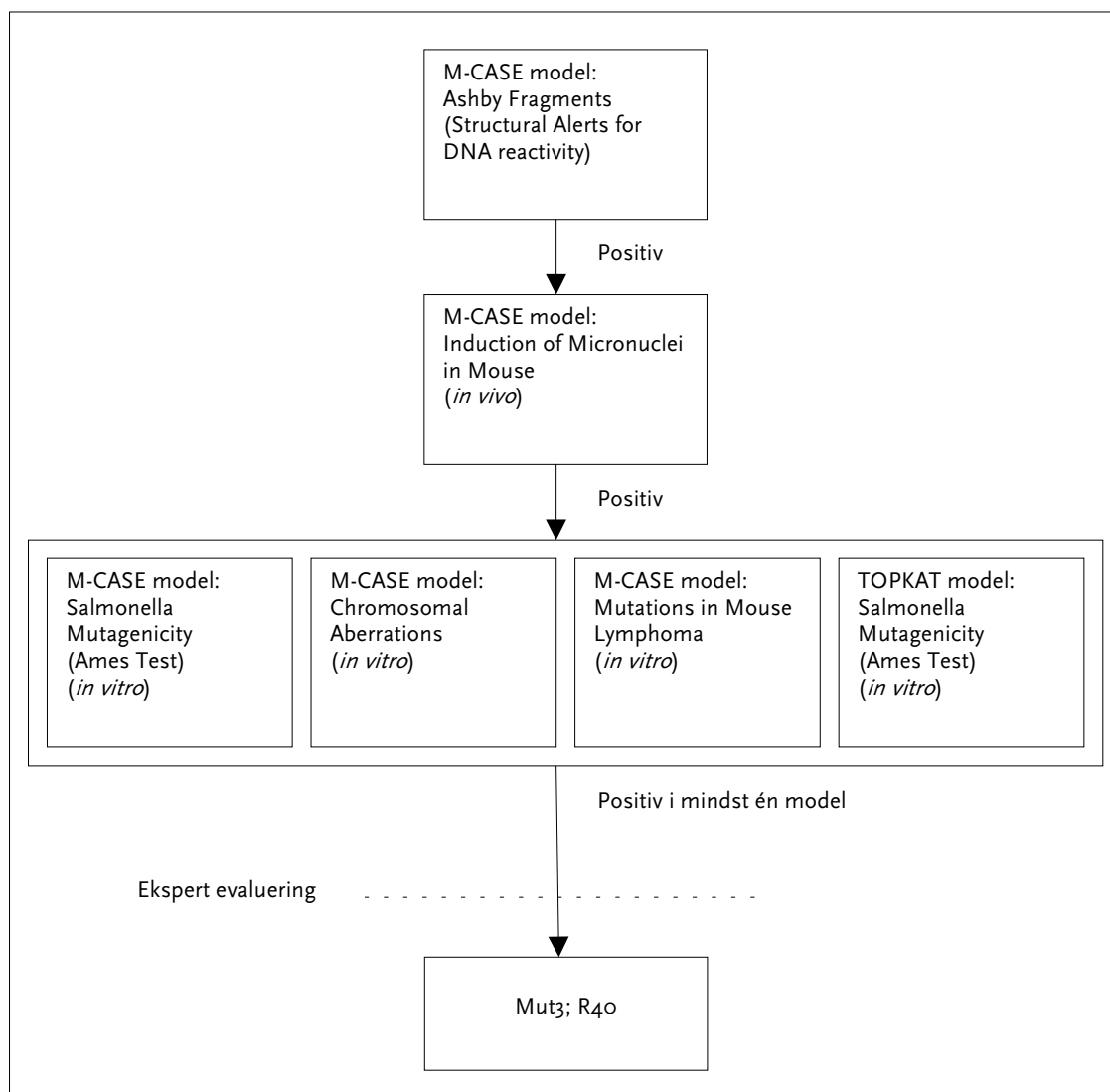
Kriteriet for mutagenicitet i den systematiske vurdering var, at det vurderede stof skulle have positive modelforudsigelser for strukturelle indikationer på DNA reaktivitet* (Structural Alert) og induktion af mikronuclei, og mindst to positive forudsigelser for de resterende genotoksiske effekter.

* Forudsigelser fra M-CASE modellen for DNA reaktivitet blev accepterede selvom de havde "ét ukendt fragment".

Der blev anvendt to modeller for Salmonella (Ames test) mutagenicitet, henholdsvis en TOPKAT og en M-CASE model. Dette skyldes, at de to modeller adskilte sig med hensyn til deres domæner, og at der for mange stoffer kun var pålidelige forudsigelser fra den ene model. Hvis pålidelige modelforudsigelser var tilgængelige fra begge modeller, men de var i modstrid med hinanden, blev det vurderet i hvert tilfælde i den sidste evaluering af stofferne.

En skematisk oversigt over den systematiske vurdering er givet i figur 4.

FIGUR 4 DEN SYSTEMATISKE VURDERING



2.272 stoffer opfyldte kriterierne for den systematiske vurdering.

Da ingen af de anvendte modeller identificerede mutagenicitet i kønsceller, kunne de dermed ikke anvendes til en diskrimination mellem EU's formelle klassificeringskriterier for mutagenicitet i de tre kategorier. Den laveste klassificering, mutagen kategori 3, blev derfor givet som vejledende klassificering for de stoffer, der i denne vurdering blev fundet at være mutagene.

Der blev foretaget ekspertvurderinger af pålideligheden af forudsigelserne for de 2.272 stoffer, der opfyldte de opsatte kriterier for klassificering for mutagenicitet. Denne proces omfattede undersøgelse af den to- eller tre-dimensionelle struktur, og visuel sammenligning med testresultater indenfor grupper af strukturelt lignende stoffer. Hvis der i denne undersøgelse blev rejst nogen tvivl om pålideligheden, blev stoffet ført over på en liste over stoffer, der skal undersøges nærmere. Dette førte til en endelig udvælgelse af 1.678 stoffer med den vejledende klassificering mut3;R40.

2.6 KRÆFTFREMKALENDE EFFEKT

EU's kriterie for klassificering

Indenfor denne farlige egenskab er klassificeringen opdelt i tre kategorier:

Klassificering som kræftfremkaldende kategori 1 (carc1;R45, Giftig; kan fremkalde kræft eller carc1;R49, Giftig; kan fremkalde kræft ved indånding) er baseret på beviser for en stærk årsagssammenhæng mellem udsættelse for stoffet og udvikling af kræft hos mennesker.

Klassificering som kræftfremkaldende kategori 2 (carc2;R45, Giftig; kan fremkalde kræft eller carc2;R49, Giftig; kan fremkalde kræft ved indånding) er baseret på baggrund af egnede langtidforsøg i dyr for mindst to dyrearter eller for én art med understøttende beviser, såsom genotoksiske effekter *in vitro* eller *in vivo*.

Klassificering som kræftfremkaldende kategori 3 (carc3;R40, Sundhedsskadelig; mulighed for varig skade for helbred) er opdelt i:

- a) Velundersøgte stoffer med begrænset kræftfremkaldende effekt. Det baseres sædvanligvis på entydige data om svulstdannelser i én dyreart. Data for skader på arveanlæggene fra *in vitro* eller *in vivo* forsøg kan anvendes som understøttende bevis.
- b) Ikke-velundersøgte stoffer, som giver anledning til bekymring for mennesker.

Vurdering baseret på modelforudsigelser

Mange ikke-genotoksiske kræftfremkaldende stoffer virker gennem en række ofte ukendte mekanismer. Der blev i dette projekt fokuseret på de kemiske stoffer, som mistænkes for at fremkalde kræft gennem genotoksiske mekanismer.

Der blev derfor sat et genotoksisk udvælgelseskriterie op. Kun de stoffer, der havde en positiv forudsigelse for strukturelle indikationer på DNA reaktivitet (AOK forudsigelser eller ét ukendt fragment) samt to positive forudsigelser fra fem modeller for genotoksicitet, blev udvalgt til at blive vurderet ved hjælp af kræftmodellerne. De tekniske specifikationer for de anvendte genotoksicitets-modeller er givet i kapitlet om skader på arveanlæggene.

I modsætning til kriterierne for skader på arveanlæggene var der i det genotoksiske kriterie ikke krav om en positiv forudsigelse for dannelse af mikronuclei i mus, idet genotoksiske kræftfremkaldende stoffer ikke nødvendigvis er klastogene (forårsager tab, addition eller ændring af dele af

kromosomer). Der var 3.362 Einescs stoffer, der opfyldte det genotoksiske udvælgelseskriterie.

Der blev anvendt ti modeller og fire undermodeller for kræftfremkaldende effekt.

TABEL 8 DE ANVENDTE MODELLER

Model	Tekniske specifikationer
TOPKAT (v. 3.01 1998) NTP* Carcinogenicity: Male Rat	366 NTP studier på gnavnere Krydsvalideringer med "Q ² " metoden viser /32/: Sensitiviteter 82-87% Specifcitet 82-88%
TOPKAT (v. 3.01 1998) NTP Carcinogenicity: Female Rat	
TOPKAT (v. 3.01 1998) NTP Carcinogenicity: Male Mouse	
TOPKAT (v. 3.01 1998) NTP Carcinogenicity: Female Mouse	
TOPKAT (v. 5.01 1998) FDA** Carcinogenicity: Male Rat	n=384 Krydsvalidering ved "Q ² " metoden viser /33/: Sensitivitet 91% Specifcitet 90%
Sub-model: Single v. multiple organ tumors	n= 131 Krydsvalidering ved "Q ² " metoden viser /33/: Sensitivitet 91% Specifcitet 96%
TOPKAT (v. 5.0 feb. 1998) FDA Carcinogenicity: Female Rat	n=383 Krydsvalidering ved "Q ² " metoden viser /33/: Sensitivitet 84% Specifcitet 89%
Sub-model: Single v. multiple organ tumors	n= 125 Krydsvalidering ved "Q ² " metoden viser /33/: Sensitivitet 92% Specifcitet 96%
TOPKAT (v. 5.0 feb. 1998) FDA Carcinogenicity: Male Mouse	n=316 Krydsvalidering ved "Q ² " metoden viser /33/: Sensitivitet 82% Specifcitet 90%
Sub-model: Single v. multiple organ tumors	n=93 Krydsvalidering ved "Q ² " metoden viser /33/: Sensitivitet 93% Specifcitet 94%
TOPKAT (v. 5.0 feb. 1998) FDA Carcinogenicity: Female Mouse	n=312 Krydsvalidering ved "Q ² " metoden viser /33/: Sensitivitet 86% Specifcitet 89%
Sub-model: Single v. multiple organ tumors	n=100 Krydsvalidering ved "Q ² " metoden viser /33/: Sensitivitet 95% Specifcitet 95%
M-CASE (v. 3.320 1999) Carcinogenic Potency Database (CPDB) model: Rat (Danish EPA version of AoD, feb. 2000)	n=870 kemiske stoffer fra CPDB Krydsvalidering ved "3*10% ud" metoden viser: Sensitivitet 52-67% Specifcitet 63-68% Chi ² ≈ 6, p<0,014 (Yderligere validering under udarbejdelse)
M-CASE (V. 3.320 1999) Carcinogenic Potency Database model: Mouse (Danish EPA version of AoE, jan. 2000)	n=720 kemiske stoffer fra CPDB Krydsvalidering ved "3*10% ud" metoden viser: Sensitivitet 45-50% Specifcitet 64-72% Chi ² ≈ 2, P=0,15 (Yderligere validering under udarbejdelse)

* NTP: National Toxicology Program, USA

** FDA: Food and Drug Administration, USA

Nøjagtigheden af forudsigelserne fra disse modeller kan være svær at bestemme ved eksterne valideringer, idet der kun findes meget få uafhængige testresultater, som ikke allerede er anvendt i modellernes træningssæt. Dette gælder især TOPKAT modellerne, hvor de eneste bestemmelser af nøjagtigheden fra producenten er baserede på valideringer ved "Q²" metoden ("one out"). For M-CASE modellerne er der blevet anvendt andre statistiske metoder til bestemmelse af nøjagtigheden.

I et længerevarende projekt, hvor adskillige modeller for kræftfremkaldende effekt gav forudsigelser for 45 kemiske stoffer, som endnu ikke var blevet testede i USA's NTP-program, var den generelle konklusion ved efterfølgende test af stofferne, at nøjagtigheden af modellerne var omkring 70% for de stoffer, der klart kunne siges at være enten kræftfremkaldende eller ikke-kræftfremkaldende /31/. På grund af det begrænsede antal stoffer i denne eksterne validering er det svært at vurdere hvor meget vægt, der kan lægges på denne konklusion.

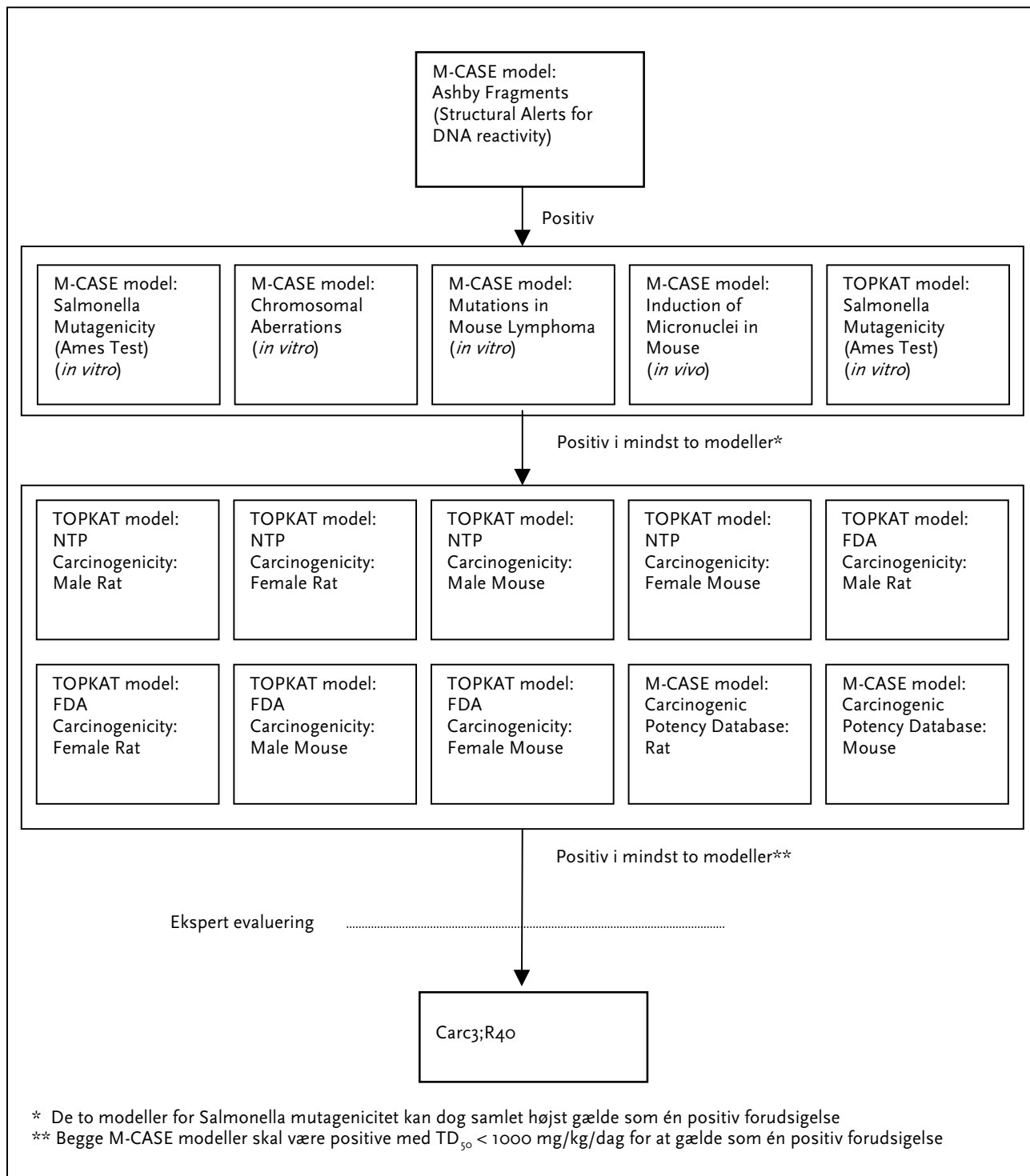
De 3.362 stoffer, der passede det genotoksiske udvælgelseskræterier, og derfor efterfølgende blev vurderet for kræftfremkaldende effekt, skulle opfylde følgende kriterier for udvælgelse:

Der skulle mindst være positive forudsigelser fra to af modellerne (undermodellerne ikke medregnet). En undtagelse blev gjort for de to M-CASE CPDB modeller. Da træningssættene for disse modeller er mindre homogene skulle begge modeller give en positiv forudsigelse for at kunne gælde som én positiv forudsigelse. Derudover var kravet til den estimerede styrke af den kræftfremkaldende effekt, at TD₅₀ skulle være mindre end 1.000 mg/kg/dag. Disse to modeller blev udviklet af Miljøstyrelsen ved brug af M-CASE metodologien, som er beskrevet for de her anvendte datasæt i referencerne /34,35,40/.

Hvis testresultater for kræftfremkaldende effekt var tilgængelige for de vurderede kemiske stoffer (indgik i modellernes træningssæt), blev disse anvendt fremfor modelforudsigelserne og resulterede i en direkte udvælgelse af stofferne til den Vejledende liste for denne farlige egenskab. Selvom dette i de fleste tilfælde ikke ændrede på hvilke stoffer, der blev udvalgt til listen (fordi modellerne har en stor tilbøjelighed til at give korrekte forudsigelser for de stoffer der indgår i træningssættet), blev det vurderet at være mere korrekt at basere de vejledende klassificeringer på modelforudsigelser, når der var testresultater til stede.

En skematisk oversigt over den systematiske vurdering er givet i figur 5.

FIGUR 5 DEN SYSTEMATISKE VURDERING



Ifølge disse kriterier blev 1.272 efter den systematiske vurdering fundet at skulle klassificeres for kræftfremkaldende effekt. Der blev udført ekspertvurderinger på QSAR-modellerne. I denne proces blev alle data gennemgået, herunder også forudsigelser fra de fire TOPKAT FDA modeller, sandsynligheder for hurtig metabolisme eller udskillelse blev vurderet, og hvor det var passende vurderinger af aryl hydroxylase aktiviteter /37/. Hvor der på denne måde blev rest tvivl om forudsigelsen af kræftfremkaldende effekt, blev stofferne ført over på en liste over stoffer, der skal undersøges nærmere, og de blev ikke inkluderet i denne første udgave af den Vejledende liste til selvklassificering af farlige stoffer.

Dette resulterede i, at 652 stoffer blev medtaget på listen med vejledende klassificeringer for kræftfremkaldende effekt. Da det blev vurderet, at de anvendte modeller ikke kunne danne grundlag for en skelnen mellem klassificering i de tre kategorier, blev den laveste klassificering Carc3;R40 anbefalet i alle tilfælde.

2.7 FARLIGHED FOR VANDMILJØET

EU's kriterie for klassificering

Klassificeringskriteriet er sammensat af tre hovedelementer: Bionedbrydning, biokoncentreringspotentiale og giftig virkning overfor akvatiske organismer. Klassificeringer baseret på farlighed for akvatiske organismer gives i overensstemmelse med kriterierne i nedenstående tabel.

TABEL 9 KLASSIFICERINGSKRITERIER

Klassificering	Kriterie for akut giftighed for akvatiske organismer* (LC ₅₀ / IC ₅₀ / EC ₅₀)
N;R50 (Miljøfarlig; Meget giftig for organismer, der lever i vand)	Akut giftighed ≤ 1,0 mg/l
N;R50/53 (Miljøfarlig; Meget giftig for organismer, der lever i vand; kan forårsage uønskede langtidsvirkninger i vandmiljøet)	Akut giftighed ≤ 1,0 mg/l og ikke let-nedbrydelig eller BCF** ≥ 100
N;R51/53 (Miljøfarlig; Giftig for organismer, der lever i vand; kan forårsage langtidsvirkninger i vandmiljøet)	Akut giftighed ≤ 10 mg/l og ikke let-nedbrydelig eller BCF ≥ 100
R52/53 (Skadelig for organismer, der lever i vand; kan forårsage langtidsvirkninger i vandmiljøet)	Akut giftighed ≤ 100 mg/l og ikke let-nedbrydelig
R53 (Skadelig for organismer, der lever i vand)	Opløselighed i vand < 1 mg/l og ikke let-nedbrydelig og BCF ≥ 100

* Den laveste effektkoncentration for fisk, dafnier eller alger anvendes

** BCF: BioConcentration Factor, biokoncentreringsfaktor

Vurdering baseret på modelforudsigelser

Vejledende klassificeringer blev givet til stoffer, der på basis af computerforudsigelser for bionedbrydning, biokoncentrerung og akut giftighed for fisk opfyldte kriterierne for klassificering angivet i tabel 9. Der blev på baggrund af den systematiske vurdering ikke givet vejledende klassificeringer med R53 alene, da den stærke sammenhæng mellem vandopløselighed og biokoncentreringsfaktoren gjorde denne klassificering overflødig.

Bionedbrydning

Bionedbrydning blev estimeret ved brug af Syracuse's BIOWIN program (v. 3.02) /17,41/. Der blev kun anvendt den lineære ligning for hurtig / ikke-

hurtig bionedbrydning. Tidligere validering af denne QSAR ved brug af såkaldte ”ready / not-ready” testresultater fra MITI-programmet viste, at mens der var en del ”not-ready” (ikke let-nedbrydelige) stoffer, der ikke blev fanget, var der ud af modellens forudsigelser af ”not-ready” 93% som var korrekte /18/. Med andre ord vil denne model ikke fange alle de ikke let-nedbrydelige stoffer, men antallet af falsk positive forudsigelser for mangel på nedbrydelighed er acceptabelt lille. Et totalantal på ca. 14.000 EINECS stoffer blev med denne model fundet at være ”ikke-letnedbrydelige” ifølge kriteriet /51*.

Biokoncentrering

I EU’s klassificeringskriterie foretrækkes målte testresultater for biokoncentrering i fisk, men da disse sjældent er tilstede eller er af mindre god kvalitet anvendes i stedet, når dette er tilfældet, $\log K_{ow}$ større end eller lig med 3,0 som indikation for, at BCF vil være 100 eller derover. Dette er i overensstemmelse med den lineære ligning fra Vieth og Kosian /41/. Selvom dette er en god tommelfingerregel, vil den føre til både over- og underestimeringer af BCF for mange klasser af stoffer. Derudover tager den ikke højde for det faktum, at biokoncentrering er en bilinear funktion af $\log K_{ow}$, der aftager når $\log K_{ow}$ er tilstrækkeligt høj.

Syracuse’s BCFWIN (v. 2.13) blev af denne grund anvendt til at give forudsigelser for biokoncentrationsfaktoren i fisk. Denne metode er baseret på en kombination af $\log K_{ow}$ QSARs og strukturelle fragmentkategorier. Ifølge en evaluering udført af skaberne af metoden er resultatet af den statistiske evaluering at $R^2 = 0,74$ ($n = 694$, S.D. = 0,65, S.E. = 0,47), hvilket er en signifikant forbedring i forhold til standardligningen fra Vieth og Kosian ($\log BCF = 0,85 * \log K_{ow} - 0,70$), hvor forudsigelserne for de samme 694 stoffer havde en statistisk nøjagtighed på $R^2 = 0,32$ (S.D. = 1,62, S.E. = 1,12) /20/. Omkring 11.000 af de vurderede EINECS stoffer har ifølge den anvendte model BCF værdier på 100 eller derover.

Der blev ikke gjort nogen forsøg på yderligere at fastlægge bioakkumuleringspotentialet dækkende den del, der henhører til optagelsen fra føden, på grund af manglende egnede modeller til forudsigelse af denne effekt.

Akut giftighed

I klassificeringen for farlighed for vandmiljøet anbefales det at anvende værdier for akut giftighed ($(L(E)C_{50})$) overfor både fisk, dafnier og alger, selvom det generelt er sjældent, at der for EINECS stoffer er værdier for alle tre arter. I denne systematiske vurdering med QSAR modeller blev det besluttet kun at anvende forudsigelser for giftigheden for fisk, på grund af deres pålidelighed og tilgængeligheden af testresultater af høj kvalitet til modeludvikling.

* Miljøstyrelsen har efter projektet med den Vejledende liste til selvklassificering af farlige stoffer udviklet en M-CASE model for bionedbrydning baseret på nye MITI data, som ifølge de første evalueringer synes at give betydeligt bedre forudsigelser. Resultatet af en validering ved ”3*10% ud” metoden gav 81% korrekte forudsigelser af ”not-ready” og 76% for ”ready”. En ekstern validering med brug af 72 ”not ready” stoffer, som ikke var i modellens træningssæt gav 89% korrekte forudsigelser. Yderligere analyse og finjustering af denne model er under udarbejdelse /19/.

Til at beregne den akutte giftighed overfor fisk udviklede Miljøstyrelsen en M-CASE model. Der blev anvendt et træningssæt med 569 96-timers LC_{50} testresultater fra Duluth Fathead minnow databasen /22/. Intern verifikation af modellen gav $R^2 = 0,85$. Krydsvalidering ved ”3*10% ud” metoden gav $R^2 = 0,735$. En beskrivelse af M-CASE metodologien anvendt på Fathead minnow testresultater er givet i referencerne /21,42/. Kun AOK forudsigelser blev anvendt i den systematiske vurdering.

Da der ikke var tilstrækkelig mange meget lipofile stoffer i træningssættet til modellen, blev den kun anvendt til vurdering af stoffer med $\log K_{ow}$ på 6 eller derunder. For stoffer med $\log K_{ow}$ over 6 blev der anvendt en anden QSAR. For disse stoffer blev det antaget, at de virkede ved ikke-polær narkotisk virkning og toksiciteten ved ligevægt blev beregnet ved følgende ligning ud fra den beregnede biokoncentrerings-faktor:

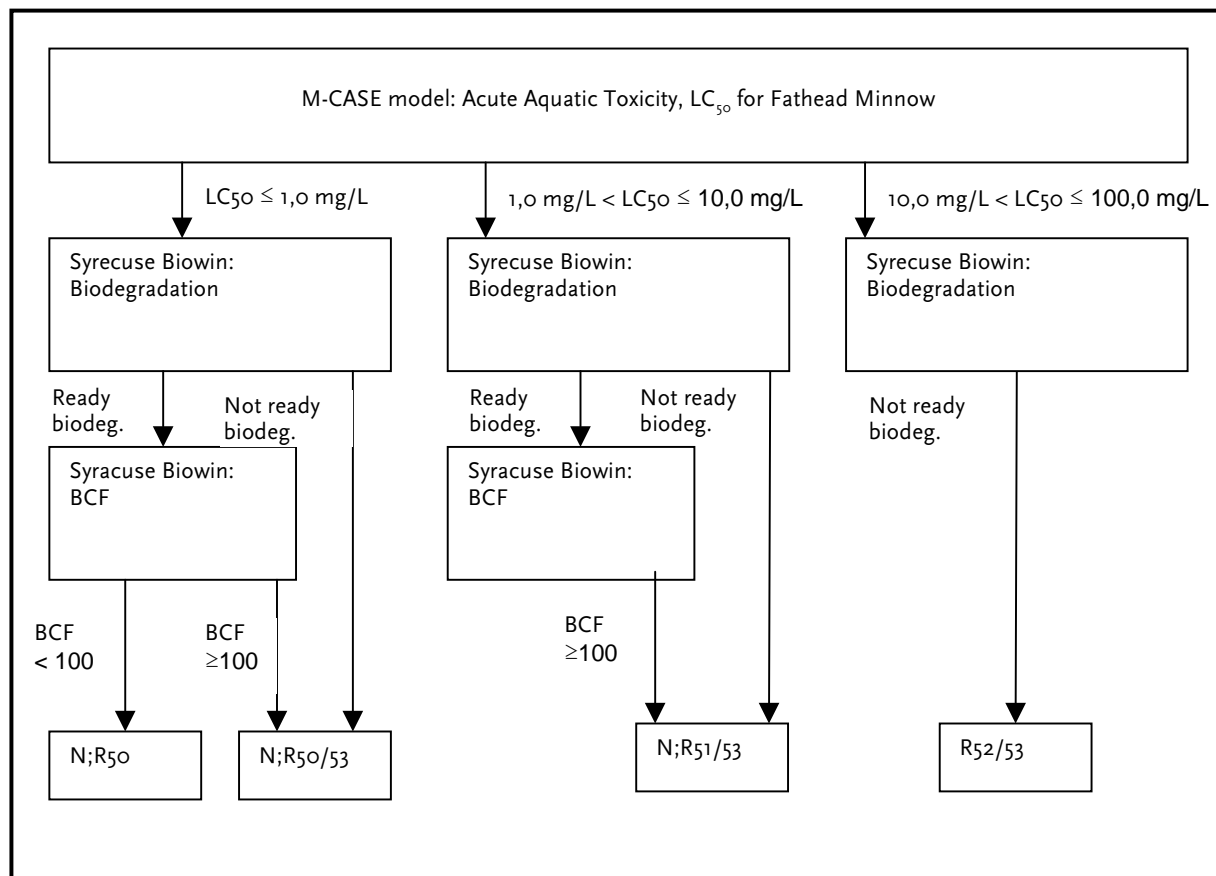
$$LC_{50} \text{ (ligevægt)} = 8,15 \text{ mmol} / BCF$$

Anvendelsen af denne ligning er i overensstemmelse med det teoretiske niveau til inducering af akvatiske effekter, angivet ved QSAR'en for ikke-polær narkotisk virkning på fisk, som er anbefalet i EU's Technical Guidance Document (til risikovurdering af kemiske stoffer) /41/. Dødelige koncentrationer (LBB: Lethal Body Burden) i fisk ud fra ikke-polær narkotisk virkning antages generelt omtrent at ligge i intervallet 2-8 mmol/kg /23,58/.

Selvom simple $\log K_{ow}$ QSARs eksisterer til forudsigelse af ikke-polær narkotisk virkning for fisk, dafnier og alger /41/, er disse ikke i stand til at forudsige, om der er specifik toksicitet, der er unik indenfor hver af de tre arter. Det er derfor ikke en fordel at anvende disse QSARs frem for at anvende fiske-modellen alene, da denne også giver fyldestgørende forudsigelser af ikke-polær narkotiske virkning såvel for fisk som for dafnier og alger, idet det er en rimelig antagelse, at ikke-polær narkotisk virkning giver effekt ved det samme koncentrationsniveau for alle tre arter /18/.

Ved brug af M-CASE modellen og den angivne QSAR blev omtrent 10.000 Einesc stoffer ud fra den systematiske vurdering fundet at være akut giftige for fisk med $LC_{50} \leq 100 \text{ mg/l}$.

FIGUR 6 DEN SYSTEMATISKE VURDERING



I alt 8.731 stoffer blev efter den systematiske vurdering angivet på den Vejledende liste til selvklassificeringer af farlige stoffer med vejledende klassificeringer for en af de fire udvalgte klassificerings-kategorier for farlighed for vandmiljøet. Taget i betragtning at der var AOK forudsigelser for giftigheden overfor fisk for lidt under halvdelen af antallet af de vurderede kemiske stoffer, synes antallet af stoffer, der er farlige for vandmiljøet, at være i overensstemmelse med hvad der kan forventes for Eines stoffer.

Fordelen ved at kunne forudsige giftige effekter, som er specifikke for både fisk, dafnier og alger er åbenlyse, og dette kan forhåbentlig opnås i fremtiden. Miljøstyrelsen har for nylig udviklet en M-CASE model for akut giftighed overfor dafnier ($n = 574$, $R^2 = 8,826$, ved "3*10% ud" metoden er $R^2 = 0,692$). Denne model undergår stadig videreudvikling. En M-CASE model for akut giftighed overfor alger er ligeledes under udvikling.

3 Referencer

1. Miljø- og Energiministeriets bekendtgørelse nr. 1065 af 30. november 2000 om klassificering, mærkning, salg og opbevaring af kemiske stoffer og produkter.
2. Einecs (European Inventory of Existing Commercial Chemical Substances). O.J. C146A, 15.6.1990, p.1.
3. Allanou et al., 1999: Remi Allanou, Bjørn Hansen and Yvonne van der Bilt (1999): Public Availability of Data on EU High Production Volume Chemicals (EUR 18996 EN).
4. Niemelä, 1992: EU project; Priority Setting for the Purpose of future Classification and Labelling of Dangerous Substances (Contract No. B91/B4.3044/12200), Danish Environmental Protection Agency, Copenhagen, November 1992 og Jay Niemelä, 1994, Miljøstyrelsens arbejdsdokument med opdatering af datagrundlaget i EU-projekt om i prioriteringsgrundlaget for klassificering og mærkning.
5. Miljø- og Energiministeriets bekendtgørelse nr. 733 af 31. juli 2000 af listen over farlige stoffer.
6. European Commission: Commission communication 90/C146A/01 pursuant to Article 13 of Council Directive 67/548/EEC of 27 June 1967 on the approximation of the laws, regulations and administrative provisions relating to the classification, packaging and labeling of dangerous substances, as amended by Directive 79/831/EEC –
7. Oxford Molecular, Chem-X version July 1998, SMILES to 3-D conversion module.
8. Niemela, J., “Non-Structural Activity Coefficients for Acute Oral Toxicity in the Mouse and Rat”, Danish EPA, working document, 1992 (available on request)
9. Niemela, J., “Acute Toxicity versus route of Administration in Mice,” Danish EPA, working document, 1995 (available on request)
10. Health Designs Inc., “TOPKAT 3.0 QSAR Module Summary, Rat Oral LD₅₀”, v3ld50.d94, Rochester, New York, U.S.A., 1997
11. Niemela, J., “Computer Predictions of LD₅₀ in the Rat using TOPKAT v. 5.01, Danish EPA, working document, 1999 (available on request).
12. Hunter, W.J., et. al., “Intercamparison Study on the Determination of Single Administration Toxicity in Rats,” J. ASSOC. AFF. ANAL. CHEM., Vol. 62, no. 4, 1979.
13. Health Designs Inc., “TOPKAT 3.0 QSAR Model Summary, Rat Chronic LOAL,” v3loal.595, Rochester, New York, U.S.A., 1997.

14. Health Designs, Inc., "New Skin Sensitization Model," Computational Toxicology News, no. 21, 1998
15. M-CASE; Model A33 Allergic contact dermatitis, July, 1998
16. Niemela, J., "QSAR's for the Estimation of Sensitization Potential," Danish EPA, working document, 1999 (available on request).
17. Howard, P.H., et. al., "Predictive Model for aerobic biodegradability developed from a file of evaluated biodegradation data," Environ. Toxicol. Chem, 11; 593-603, 1992. (see also BIOWIN Biodegradation Probability Program for Microsoft Windows 3.1, Syracuse Research Corporation, Syracuse, NY, U.S.A., 1994.
18. Pedersen, F., Tyle, H., Niemela, J., Guttman, B., Lander, L., and Wedebrand, A., 1995, "Environmental Hazard Classification – data collection and interpretation guide for substances to be evaluated for classification as dangerous to the environment. Appendix 9; Validation of the BIODEG Probability Program." TemaNord Report 589, 153-156.
19. Niemelä, J., and Wedebye, E.B., "MITIMOD, M-CASE Model for Aerobic Biodegradation, Danish EPA, internal working document, April, 2000.
20. Meylan, W.M., et. al., "Improved Method for Estimating Bioconcentration Factor (BCF) from Octanol-Water Partition Coefficient," Third Update Report, Syracuse Research Corporation, Syracuse, NY, U.S.A. August, 1997.
21. Klopman, Gilles, Saiakov, R., and Rozenkranz, S., "Multiple Computer-Automated Structure Evaluation Study of Aquatic Toxicity II. Fathead minnow," Environmental Toxicology and Chemistry, Vol. 19, No 2, pp. 441-447.
22. Brooke, L.T., et. al., "Acute Toxicities of Organic Chemicals to Fathead Minnows (*Pimephales promelas*)," Center for Lake Superior Environmental Studies, University of Wisconsin – Superior, 1988.
23. Hermens, J.L., et. al., "Aquatic Toxicity of Polar Narcotic Pollutants," University of Utrecht, Utrecht, Netherlands, 1998
24. M-CASE Structural Alerts for DNA Reactivity ("Ashby fragments", model A2E, July 1998.
25. Health Designs Inc., "TOPKAT 3.0 QSAR Module Summary: Mutagenicity (Ames), V3mut.d94, New York, U.S.A., 1997
26. Niemela, J., "Computer Prediction of Ames Test Mutagenicity (progress report, May 1998)," Danish EPA, working document, 1998 (available on request)
27. M-CASE Salmonella (Ames) Mutagenicity, model A2H, data evaluation, July 1999.

28. M-CASE Chromosomal Aberrations, model A61, data evaluation, July 1998
29. M-CASE Mutations in Mouse Lymphoma, model A2F, data evaluation, July 1998.
30. M-CASE Induction of Micronuclei, model A62, data evaluation, July 1998.
31. Zhang, Y.P., et. al., "Development of Methods to Ascertain the Predictivity and Consistency of SAR Models: Application to the U.S. National Toxicology Program Rodent Carcinogenicity Bioassays," *Quant. Struct-Act. Relat*, 16, 290-295, 1997.
32. Health Designs Inc., "TOPKAT 3.0 QSAR Module Summary: Rodent Carcinogenicity, v3carc.995, NY, U.S.A., 1997.
33. Health Designs Inc., "Computational Toxicology News, "MTA with FDA produces new Carcinogenicity QSTR's," no. 21, February 1998.
34. Cunningham, A.R., Klopman, G., Rosenkranz, H.S., "Identification of structural features and associated mechanisms of action for carcinogens in rats", *Mutation Research*, 405; 9-28, 1998.
35. Cunningham, A.R., Rosenkranz, H.S., Zhang, Y.P., Klopman, G., "Identification of "genotoxic" and "non-genotoxic" alerts for cancer in mice: The Carcinogenic potency database," *Mutation Research*, 398; 1-17, 1998.
36. Gold, L.S., et. al., "The Carcinogenic Potency Database," National Institute of Environmental Health Sciences, NIEHS Center, University of California, Berkeley, 1999
37. M-CASE, Binding to Aryl Hydrocarbon Hydroxylase, Model A68, July 1998.
38. Klopman, G., "MULTICASE 1. A Hierarchical Computer Automated Structure Evaluation Program," *Quant. Struct.-Act. Relat.*, 11; 176-184, 1992.
39. Oxford Molecular Group, Inc., Health Designs, Inc., *Computational Toxicology, "TOPKAT 5.0 Reference Manual," 1997.*
40. Mathews, E.J., Contrera, J.F., "A New Highly Specific Method for Predicting the Carcinogenic Potential of Pharmaceuticals in Rodents Using Enhanced MCASE QSAR-ES Software", *Reg.Tox. and Pharm.*, 28; 242-264, 1998.
41. European Commission, "Technical Guidance Document in Support of commission directive 93/67EEC on Risk Assessment for New Notified Substances and Commission Regulation (EC) No. 1488/94 on Risk Assessment for Existing Substances," Part III, ISBN 92-827-8013-9, 1996, p. 554.

42. Klopman, G., Saiakov, R, Rosenkranz, HS, Hermans, JLM, "M-CASE study of aquatic toxicity, I. Guppy," *Environ. Toxicol. Chem.* 18: 2497-2505.
43. G. Klopman and O.T. Macina, "Drug Design Based on Artificial Intelligence," *Computer Aided Innovation of New Materials II*, Elsevier Science Publishers, 1992.pp 1135-1140.
44. Mumtaz, M.M., et. Al. "Assessment of effect levels of chemicals from quantitative structure-activity (QSAR= models. I. Chronic lowest-observed-adverse-effect (LOAEL)," *Toxicology Letters*, 79 (1995), 131-143
45. Gomber, V. K., "Quantitative Structure-Activity Relationships in Toxicology; From Fundamentals to Applications," Health Designs Inc., 183 East Main Street, Rochester, NY 14604, USA.
46. O. Mekenyan, S. Karabunarliev and D. Bonchev, The OASIS Concept for Predicting the Biological Activity of Chemical Compounds, *J. Math. Chem.*, 4, 207-215 (1990).
47. O. Mekenyan, S. Karabunarliev and D. Bonchev, The Microcomputer OASIS System for Predicting the Biological Activity of Chemical Compounds, *Comput. & Chem.*, 14, 193-200 (1990).
48. Mekenyan, et. al. "COREPA Method Used for Evaluation of Reactivity Profile for High ER Binding Affinity," *Quant. Struct.-Act. Relat.*, 18:139-153, 1999.
49. Moudgal, C. J., Lipscomb, R. M. Bruce, "Potential health effects of drinking water disinfection by-products using quantitative structure toxicity relationships, *Toxicology*, vol 147, 2000, pp. 109-131.
50. Eriksson, L, Johansson, E., Keteth-Wold, N., and Wold, S. "Introduction to Multi- and Megavariate Analysis using Projection Methods (PCA & PLS," *Umetrics*, Sweden, 1999 pp. 251-260.
51. OECD Environmental Health and Safety Publications Series on Testing and Assessment, No. 24, "Revised Draft Guidance Document on the use of the Harmonized System of Classification of Chemicals which are Hazardous for the Aquatic Environment," Environment Directorate, Organization for Economic Cooperation and Development, Draft, Oct. 2000 (ENV/JM/HCL (2000) 15/REV2).
52. Registry of Toxic Effects of Chemical Substances (RTECS), The national Institute of Occupational Safety and Health, Washington, D.C, 1998 (Silver Platter electronic version).
53. M-CASE, Users Guide, Version 3.30 (Rev. 1.0), Multicase inc., 25825 Science Drive Park, suite 100, Cleveland, OH, 44122.
54. Rosenkranz, H.S., et.al, "Development, Characterization and Application of Predictive Toxicology Models," *SAR and QSAR in Environmental Research*, Vol. 10, pp. 277-298, 1999.

55. Wedebye, E.B., and Niemela, J. R., "Database with QSAR Estimates for more than 166,000 Chemicals," QSAR 2000, Ninth International Workshop on Quantitative Structure Activity Relationships in Environmental Sciences, September 16-20 2000, Bourgas, Bulgaria. Abstracts, PV.2
56. Enslein, Kurt, Health Designs, Inc., 183 E. Main Street, Rochester, NY 14604: "QSTR applications in acute, chronic, and developmental toxicity, and carcinogenicity.
57. Rådets direktiv 67/548/EØF om tilnærmelse af lovgivning om klassificering, emballering, og etikettering af farlige stoffer.
58. ECETOC Technical Report No. 67, "The Role of Bioaccumulation in Environmental Risk Assessment: The Aquatic Environment and Related Food Webs", 1995.