



**Miljø- og
Fødevareministeriet**
Departementet

Kemikalier

Den 28. januar 2019

Notat 1 fra arbejdsgruppen vedr. børneleukæmi og pesticider:

Beskrivelse af muligheder for at bruge befolkningsundersøgelser i vurdering af pesticider og hvad det betyder for vurdering af risikoen for gravide - herunder begrænsninger ved denne type undersøgelser og behov for yderligere undersøgelser.

Indhold

1.	Indledning.....	2
2.	Risikovurdering af pesticider ift. menneskers sundhed	2
2.1.	Risikovurderingsprocessen.....	4
2.2.	Myndighedskrav til data ifm. risikovurdering af pesticider	5
3.	Introduktion til epidemiologiske undersøgelser.....	7
4.	Befolkningsundersøgelser af pesticiders mulige helbredseffekter	10
5.	Begrænsninger ved befolkningsundersøgelser af pesticiders mulige helbredseffekter	10
6.	Anvendelse af befolkningsundersøgelser ved risikovurdering af pesticider	12
6.1.	Integrering af befolkningsundersøgelser i risikovurderingsprocessen	12
6.2.	Vurdering af pålideligheden ("reliabiliteten") af befolkningsundersøgelser....	13
6.3.	Relevansvurdering af befolkningsundersøgelser.....	16
6.4.	Vurdering af den biologiske relevans af en statistisk signifikant sammenhæng	18
7.	Forbedring af pesticid-epidemiologiske undersøgelser ift. regulatorisk anvendelse	19
8.	Konklusion.....	20
9.	Referencer	21
10.	Ordlister og forkortelser.....	22

1. Indledning

Forskere fra National Cancer Institute, The National Institute of Health i USA præsenterede den 29. august 2018 foreløbige resultater fra en epidemiologisk undersøgelse (en type befolkningsundersøgelser), der pegede på en mulig sammenhæng mellem gravides bopæl nær dyrkede marker samt eksponering for visse ukrudtsmidler og udvikling af leukæmi hos deres børn. Undersøgelsen er baseret på danske data og blev præsenteret på konferencen ISES-ISEE Joint Annual Meeting i Canada ved foredraget "Residential Proximity to Agricultural Herbicides During Pregnancy and Childhood Leukemia in the Danish National Birth Cohort"¹. Forskere fra Statens Serum Institut har bidraget med data fra den nationale fødselskohorte "Bedre sundhed i generationer" (BSIG), hvor man har registreret en række data om ca. 100.000 børn og deres mødre, som fødte i perioden 1996-2003. Desuden har forskerne brugt Cancerregisteret til at identificere kræfttilfælde hos børnene indtil de er 15 år gamle. Forskere fra Aarhus Universitet, Nationalt Center for Miljø og Energi, DCE, har bidraget med data om markbeliagenhed, afgrødetyper og pesticidforbrug.

På baggrund af undersøgelsens foreløbige resultater besluttede miljø- og fødevarerministeren at nedsætte en arbejdsgruppe vedr. børneleukæmi og pesticider. En af arbejdsgruppens opgaver er at komme med fagligt bidrag til nærværende notat, udarbejdet af Miljø- og Fødevarerministeriet (MFVM), med følgende opdrag:

"Beskrivelse af muligheder for at bruge befolkningsundersøgelser i vurdering af pesticider og hvad det betyder for vurdering af risikoen for gravide - herunder begrænsninger ved denne type undersøgelser og behov for yderligere undersøgelser."

Notatet vil danne baggrund for, hvordan resultater af fremtidige befolkningsundersøgelser omhandlende pesticider og sygdomsudvikling skal håndteres, dvs. hvordan styrker og svagheder ved undersøgelserne skal vurderes, og hvordan resultaterne kan bruges ved risikovurdering af pesticider.

Der er, i tillæg til dette notat, udarbejdet en kort opsummering af arbejdsgruppens anbefalinger for brug af epidemiologiske undersøgelser i risikovurdering af pesticider. Opsummeringen indeholder et beslutningstræ, der visuelt sammenfatter de overordnede anbefalinger, der er beskrevet i nærværende notat.

2. Risikovurdering af pesticider ift. menneskers sundhed

Hvis et pesticid ønskes markedsført på det europæiske marked, skal det godkendes iht. proceduren givet i Europa-Parlamentets og Rådets forordning Nr. 1107/2009² (pesticidforordningen). De regulatoriske myndigheder udfører en formel risikovurdering ift. mennesker og miljø for hvert pesticid. Der stilles flest krav til selve pesticidaktivstoffet ift. undersøgelse af korttids- og langtidseffekter for mennesker, og væsentlig færre til det

¹ Oplæggets titel kan oversættes til: "Beboelsesmæssig nærhed til landbrugets ukrudtsmidler under graviditeten og børneleukæmi i den danske nationale fødselskohorte".

² <https://eur-lex.europa.eu/legal-content/DA/TXT/PDF/?uri=CELEX:02009R1107-20170828>

markedsførte pesticidprodukt, hvor primært akutte effekter undersøges. Risikovurderingen har endvidere fokus på særlige hjælpestoffer, der gør et aktivstof mere virksomt over for skadevoldere (såkaldte ”synergister”) eller beskytter de sprøjtede planter mod pesticidets giftige virkning (såkaldte ”safeners”). For de øvrige hjælpestoffer i det færdige pesticidprodukt foretages der ikke en særskilt risikovurdering. Det skyldes, at hjælpestoffer generelt antages at være mindre problematiske end selve de aktive stoffer, og derfor vil være dækket ind af risikovurderingen af aktivstoffernes langtidseffekter, mens de akutte effekter af det endelige pesticidprodukt altid bliver undersøgt. Desuden vil mange hjælpestoffer høre under andre kemikalielovgivninger såsom EU’s generelle kemikalieforskrift REACH³ (Europa-Parlamentets og Rådets forordning (EF) nr. 1907/2006)⁴. Der bliver altid foretaget en søgning på toksikologisk viden om alle stoffer i det færdige pesticidprodukt, og denne viden bliver inddraget i risikovurderingsprocessen.

Risikovurdering af pesticider ift. menneskers sundhed er baseret på en lang række af påkrævede toksikologiske forsøg, der skal udføres iht. internationalt accepterede og standardiserede forsøgsprotokoller fra OECD⁵, og som har til formål at identificere og karakterisere eventuelle farer. Ud fra dyreforsøg identificeres aktivstoffets kritiske effekter og på baggrund af disse etableres såkaldte sundhedsmæssige ”referenceværdier” for stoffet. De såkaldte referenceværdier er mål for en mængde af et specifikt pesticid, som et menneske vurderes at kunne udsættes for uden helbredsrisiko. Referenceværdierne beskytter alle befolkningsgrupper, inklusive gravide, det ufødte barn og børn. Referenceværdierne fastsættes ud fra dyreforsøg, hvor der identificeres en dosis, hvor der ikke kan påvises skadelige effekter, og denne dosis reduceres så yderligere ved brug af usikkerhedsfaktorer. Usikkerhedsfaktorerne dækker bl.a. over forskelle mellem dyr og mennesker (her anvendes ofte en usikkerhedsfaktor på 10) samt over forskelle mellem mennesker imellem (her anvendes tilsvarende oftest en usikkerhedsfaktor på 10). Yderligere usikkerhedsfaktorer kan anvendes ud fra en ekspertvurdering. Der sættes referenceværdier for eksponering for et aktivstof i forbindelse med markudbringning af et pesticidprodukt, hvor eksponeringen primært sker ved hudkontakt og indånding⁶, såvel som for indtagelse af rester af pesticidet via kosten⁷.

³ Registration, Evaluation, Authorisation and restriction of CHemicals (registrering, vurdering og godkendelse af samt begrænsninger for kemikalier. EU’s generelle kemikalieforskrift)

⁴ <https://eur-lex.europa.eu/legal-content/DA/TXT/PDF/?uri=CELEX:02006R1907-20140410&from=EN>

⁵ The Organisation for Economic Co-operation and Development (Organisationen for Økonomisk Samarbejde og Udvikling)

⁶ Referenceværdierne AOEL (Acceptable Operator Exposure Level) og AAOEL (Akut AOEL) angiver en grænse for acceptabel eksponering af sprøjteføreren ved gentagen anvendelse af pesticidet hhv. acceptabel eksponering af sprøjteføreren inden for en dag, hvor sidstnævnte kun fastsættes for akuttoksiske stoffer.

⁷ Referenceværdierne ADI (Acceptabelt Dagligt Indtag) og ARfD (Akut ReferenceDosis) angiver det højeste beregnede indtag af restkoncentrationer af et stof i kosten, man kan have hver dag hele livet uden sundhedsmæssig risiko hhv. den mængde af et stof, som kan indtages i en periode på 24 timer eller mindre uden sundhedsmæssig risiko.

For de enkelte pesticidprodukter skal der, ud over undersøgelse af akut giftighed samt evne til at fremkalde irritation og allergi, også bestemmes optagelse gennem huden (dermal absorption). Den dermale absorption angiver, hvor stor en procentdel af aktivstoffet, der optages i kroppen ved hudkontakt med det specifikke pesticidprodukt. Menneskers eksponering for et givent aktivstof (primært ved hudkontakt og indånding) kan estimeres vha. forskellige beregningsmodeller ud fra den produktspecifikke dermale absorption og de specifikke anvendelser af produktet. Risikoen for mennesker ved eksponering for et pesticid vurderes ved at sammenholde den etablerede referenceværdi for aktivstoffet med den estimerede eksponering ved en given anvendelse.

Der foretages også en risikovurdering ift. at mennesker udsættes for pesticider eller deres nedbrydningsprodukter via indtag af mad og drikke, som kan indeholde rester af pesticider. For hvert aktivstof fastsættes en referenceværdi, som beskrevet ovenfor, der fastsætter den mængde af stoffet, som et menneske vurderes at kunne indtage hver dag livet igennem uden sundhedsmæssig risiko; for aktivstoffer, der kan være akut farlige, fastsættes desuden en grænseværdi, der angiver den mængde stof, som et menneske vurderes at kunne indtage over kort tid (1 dag) uden sundhedsmæssig risiko. Ud fra disse referenceværdier og de specifikke anvendelser af det enkelte pesticidprodukt bliver der fastsat en maksimal grænseværdi for restkoncentrationer af aktivstoffet, der må kunne findes i en behandlet afgrøde; her tages højde for, at det samme pesticid/aktivstof kan være anvendt i en lang række forskellige afgrøder. Hvis disse maksimale grænseværdier overholdes, er de restkoncentrationer, der findes i fødevarer, vurderet at være sundhedsmæssigt acceptable.

2.1. Risikovurderingsprocessen

Godkendelsen af et aktivstof i EU er bl.a. baseret på en vurdering af risiko for mennesker ved eksponering for pesticidet. Risikovurderingsprocessen ift. menneskers sundhed er typisk baseret på dyreforsøg og kan inddeles i fire trin:

1. Fareidentifikation - Undersøgelse af stoffets skadelige effekter på helbredet.
2. Farekarakterisering - Undersøgelse af sammenhængen mellem stigende dosis/eksponering og skadelig(e) effekt(er) for at kunne fastlægge "ingen effekt"-niveauet, hvor der ikke kan påvises en skadelig påvirkning; dette kaldes også en "dosis-respons" vurdering. Ud fra "ingen effekt"-niveauet og ved brug af usikkerhedsfaktorer bestemmes referenceværdier for stoffet.
3. Eksponeringsvurdering - Den potentielle eksponering af mennesker ved en given anvendelse af et pesticid bliver estimeret ud fra beregningsmodeller. Vurderingen sker for fire grupper: sprøjteførere, arbejdere i landbruget, naboer til marker og forbigående. Eksponeringen via fødevarer vurderes særskilt.
4. Risikovurdering - Den potentielle eksponering sammenholdes med en stofs specifik referenceværdi for at vurdere, om nogle af de eksponerede grupper udsættes for en øget risiko.

2.2. Myndighedskrav til data ifm. risikovurdering af pesticider

Der stilles en lang række datakrav til aktivstoffer i pesticider og til pesticidprodukter, som fremsat i hhv. Kommissionens forordning Nr. 283/2013⁸ og 284/2013⁹. Den krævede datapakke for aktivstoffer ift. menneskers sundhed er primært baseret på dyreforsøg, men også "reagensglasforsøg" og computer-baserede forudsigelser i visse tilfælde. Datapakken, der kræves for pesticidprodukter, er mere begrænset end for aktivstoffet, og handler primært om undersøgelse af produktets akutte giftighed samt irriterende og allergifremkaldende egenskaber. Data kan baseres på dyreforsøg og/eller "reagensglasforsøg", og behov for klassificering af et pesticidprodukt for akutte effekter kan også afklares ved beregning, hvis den/de akutte effekter er undersøgt for de enkelte stoffer i produktet jf. reglerne i forordning (EF) nr. 1272/2008 om klassificering, mærkning og emballering af kemiske stoffer og blandinger (CLP-forordningen¹⁰).

Data fra laboratorieforsøg (dyreforsøg og undertiden "reagensglasforsøg") er den primære kilde til identifikation og karakterisering af aktivstoffets farlighed (jf. punkt 1 og 2 under afsnit 2.1). Forsøgene kræves udført efter internationalt accepterede og standardiserede forsøgsprotokoller (OECD test guidelines) og udført iht. god laboratoriepraksis (GLP). Forsøgsprotokollerne er meget detaljerede ift. forsøgsdesign og hvilke undersøgelser/målinger (de såkaldte "endpoints"), der skal foretages for det enkelte forsøg. Der stilles således strenge krav til forsøgenes udførelse og afrapportering, og kun hvis disse er opfyldt, vil forsøgene kunne tages med i betragtning. Ved myndighedsvurdering af de enkelte forsøg opereres med tre kategorier: acceptable, supplerende/understøttende og ikke-acceptable forsøg, og det er kun forsøg i de første to kategorier, der medtages i risikovurderingen.

For et aktivstof stilles der overordnet krav om undersøgelser af stoffets "skæbne" i kroppen (optagelse, fordeling, omsætning og udskillelse), stoffets akutte effekter inkl. irriterende og allergifremkaldende egenskaber, effekter ved gentagen eksponering over kortere tid (28 og 90 dage), effekter på arveanlæggene, effekter ved gentagen eksponering over længere tid (1,5 og 2 år) og kræftfremkaldende egenskaber, effekter på forplantningsevnen og fosterskadende effekter, hormonforstyrrende egenskaber, effekter på nervesystemet samt effekter af eventuelle nedbrydningsprodukter. Der stilles endvidere krav om anvendelse af humane data, hovedsageligt beskrevet under forordningens afsnit 5.9 "Medicinske data". Disse data kommer f.eks. fra rapporter udarbejdet i forbindelse med utilsigtet arbejdsmæssig eksponering eller forsætlig selv-forgiftning og fra eventuelle helbreds-/monitoreringsundersøgelser af personalet på fabrikker, der fremstiller pesticidet. Data kan også stamme fra den videnskabelige, kollegial-evaluerede alment tilgængelige litteratur, herunder publicerede befolkningsundersøgelser. Specifikt står der i datakravene, at "Der skal fremlægges relevante epidemiologiske undersøgelser, i det omfang de foreligger." Pesticidforordningen kræver, at relevant litteratur, der er publiceret inden for de seneste 10

⁸ <https://eur-lex.europa.eu/legal-content/DA/TXT/PDF/?uri=CELEX:32013R0283&from=EN>

⁹ <https://eur-lex.europa.eu/legal-content/DA/TXT/PDF/?uri=CELEX:32013R0284&from=EN>

¹⁰ Classification, Labelling and Packaging - <https://echa.europa.eu/da/regulations/clp/legislation>

år før indsendelse af ansøgningen om (forny) godkendelse af et aktivstof, skal indgå i risikovurderingen af pesticider jf. artikel 8 stk. 5 (ældre litteratur må også gerne indgå, såfremt det kan retfærdiggøres og i øvrigt er identificeret og udvalgt i overensstemmelse med EFSA's¹¹ retningslinjer på området¹²). Befolkningsundersøgelser og andre humane data vil være relevante særligt ved vurdering af, om et aktivstof skal have godkendelsen fornyet, da der typisk ikke findes publiceret litteratur om et nyt aktivstof ved den første godkendelse (med mindre det allerede er godkendt uden for EU).

Som nævnt har dyreforsøg hidtil været centrale for risikovurderingen. Imidlertid står det klart, at mere komplekse sygdomsbilleder hos mennesker ikke altid er dækket af de dyreforsøg, der stilles krav om, som f.eks. børneleukæmi og visse neurologiske sygdomme. På baggrund af en litteraturgennemgang af en lang række epidemiologiske undersøgelser af pesticiders mulige helbredseffekter (Ntzani 2013), blev EFSA's panel vedr. plantebeskyttelsesmidler og deres restkoncentrationer (PPR Panelet) bedt om at undersøge den sandsynlige medvirken af pesticideksponering som en risikofaktor for bl.a. børneleukæmi. PPR Panelets arbejde i denne forbindelse sandsynliggør¹³, at en miljømæssig påvirkning af moderen under graviditeten i visse tilfælde kan være en risikofaktor for udvikling af blodkræft i børnenes første leveår (EFSA PPR Panel 2017b), og dette fanges ikke af de klassiske dyreforsøg med eksponering i fostertilstanden. Endvidere blev det påpeget at generelt for børneleukæmier er de hidtil påkrævede standardiserede dyreforsøg ikke tilstrækkelige. En tilsvarende problemstilling kunne gælde for andre kræftformer og andre sygdomme¹⁴.

I de gældende datakrav for pesticid-aktivstoffer er der to typer af obligatoriske forsøg, hvor drægtige dyr eksponeres for et stof: et to-generationers reproduktionstoksisk forsøg med særligt fokus på forplantningssystemet samt ungerens vækst og udvikling (OECD test nr. 416¹⁵) og et udviklingsforsøg med fokus særligt på effekter i drægtige dyr og fostre (OECD test nr. 414¹⁶). I to-generationersforsøget udsættes gnavere kontinuerligt gennem to generationer for et stof, dvs. også under drægtighedsperioden, for særligt at undersøge for eventuelle effekter på forplantningsevnen, og i denne forsøgsprotokol stilles der ikke krav om blodanalyse (hverken af afkommet eller de voksne dyr). I de såkaldte udviklingsforsøg bliver moderdyrene eksponeret for et stof under (dele af) graviditeten, men her udtages ungerne ved kejsersnit umiddelbart inden fødslen, da fokus er på fosterskader, og ungerens langsigtede helbredstilstand undersøges derfor ikke. Der findes endnu en forsøgsprotokol

¹¹ European Food Safety Authority (den europæiske fødevarerikkerhedsautoritet)

¹² <https://efsa.onlinelibrary.wiley.com/doi/epdf/10.2903/j.efsa.2011.2092>

¹³ Panelet anvendte en konceptuel ramme med begivenhedskæder, der leder fra en initierende molekylær hændelse til et skadeligt udfald for organismen, kendt som "adverse outcome pathway" (AOP)

¹⁴ Jf. hypotesen om føtal oprindelse af sundhed og sygdom (DOHaD: Developmental Origins of Health and Disease)

¹⁵ https://www.oecd-ilibrary.org/environment/test-no-416-two-generation-reproduction-toxicity_9789264070868-en

¹⁶ https://www.oecd-ilibrary.org/environment/test-no-414-prenatal-development-toxicity-study_9789264070820-en

med eksponering i fostertilstanden, et såkaldt udvidet en-generations reproduktionstoksisk forsøg, hvor bl.a. forplantningsevne og fosterudvikling undersøges, og her er blodanalyser påkrævet (OECD test nr. 443¹⁷). I de gældende datakrav er denne nyere forsøgsprotokol imidlertid ikke obligatorisk, men kan dog blive krævet i visse tilfælde. Endvidere skal det bemærkes, at protokollen ikke er designet til at undersøge kræft, og det er tvivlsomt, om den er følsom nok. For de mere komplekse sygdomsbilleder hos mennesker findes der endvidere ikke altid en dyremodel, der fuldstændigt afspejler de væsentligste elementer i sygdomsudviklingen og sygdommen i sig selv; dette gør sig f.eks. gældende ift. børneleukæmi. For en række andre helbredseffekter mangler der tilsvarende en fuldt dækkende dyremodel, såsom effekter på hjerte-kar-systemet, det metaboliske system og effekter på hjernen i et livstidsperspektiv (risiko for udvikling af demens); her udføres f.eks. undersøgelser af væv fra de relevante organer, men de eksisterende forsøgsprotokoller kan ikke fuldt identificere alle tænkelige helbredseffekter.

3. Introduktion til epidemiologiske undersøgelser

Epidemiologiske undersøgelser, en særlig type ”befolkningsundersøgelser”, beskæftiger sig typisk med udvalgte påvirkninger i en given population, der kan have indflydelse på sygdomme mv. Undersøgelserne kan være beskrivende eller analytiske. I den beskrivende epidemiologi undersøges hyppigheden af sygdomme over tid eller på et givet tidspunkt. Den analytiske epidemiologi har til formål at belyse, om udvalgte faktorer øger risikoen for bestemte sygdomme, f.eks. om et pesticid øger risikoen for leukæmi hos børn. I analytiske studier indgår normalt en eller flere referencegrupper, der ideelt ikke er udsat for f.eks. det pesticid, der ønskes belyst i forhold til sammenhæng med sygdom. Epidemiologiske undersøgelser foretages i menneskers vanlige omgivelser, dvs. ikke i et kontrolleret miljø som ved laboratorieundersøgelser. Dette er både en styrke, fordi menneskers virkelige livsbetingelser indgår, og en udfordring, fordi menneskers liv omfatter komplekse påvirkninger. Derfor kan der i epidemiologiske undersøgelser være indbygget systematiske fejl (bias), herunder informationsbias, der refererer til ukorrekt bestemmelse af eksponering og/eller sygdomsudfald. Mange kroniske sygdomme er multifaktorielle, dvs. at de kan have flere forskelligartede påvirkninger som årsag. Det kan derfor virke forstyrrende på undersøgelsens resultater og tolkning, hvis der ikke tages højde herfor, når data analyseres. Sådanne mulige forstyrrende faktorer (såkaldte ”konfoundere”¹⁸) skal derfor så vidt muligt måles og medtages i analyserne. Det svageste led i epidemiologiske studier er ofte valide oplysninger om eksponeringen (sammensætning, omfang og tidspunkt), og i mindre grad helbredseffekterne.

Epidemiologiske undersøgelser inddeles typisk på basis af design i følgende fire typer: populationsbaserede korrelationsundersøgelser, tværsnitsundersøgelser, case-kontrol samt

¹⁷ https://www.oecd-ilibrary.org/environment/test-no-443-extended-one-generation-reproductive-toxicity-study_9789264185371-en

¹⁸ Konfounding betyder sammenblanding og opstår, når sammenhængen mellem den undersøgte eksponering og sundhedsudfaldet påvirkes af en tredje faktor (konfounderen), uden at denne faktor indgår i årsagskæden fra eksponering til udfald.

kohorteundersøgelser (prospektive og retrospektive). I faktaboksen nedenfor er der en kort beskrivelse af de enkelte typer. Undersøgelserne kan baseres på selvrapporterede oplysninger eller objektive oplysninger fra f.eks. registre, både hvad angår eksponering og helbredseffekter. Der er fordele og ulemper ved hver type design og dataindsamlingsmetode, og tolkning af resultater afhænger af problemstillingen, herunder type af påvirkning og sygdom. Kvaliteten af eksponeringsestimatet og muligheden for at vurdere sammenhæng mellem eksponering og udfald er afgørende for kvaliteten af undersøgelsen.

Som mål for sammenhængen mellem en eksponering og en helbredseffekt anvendes ofte et af to associationsmål: "relative risk" (RR) eller "odds ratio" (OR). Den relative risiko, RR, angiver forholdet mellem sandsynligheden for en helbredseffekt i en eksponeret gruppe og sandsynligheden for udfaldet i en sammenlignelig ikke-eksponeret gruppe. "Odds ratio" angiver chancen for at et udfald indtræder ved en given eksponering sammenholdt med chancen for at udfaldet indtræder i fraværet af eksponeringen¹⁹. Forskellige niveauer af eksponering kan sammenlignes for at undersøge, om større eksponering giver øgede effekter (undersøgelse af et muligt "dosis-respons" forhold).

I meta-analyser og poolede analyser kan resultater eller data fra flere uafhængige undersøgelser af samme problemstilling søges kombineret. Typisk er tolkning usikker for problemstillinger, hvor der kun findes få epidemiologiske undersøgelser. Endvidere er det normalt underbyggende for tolkning af resultater fra epidemiologiske undersøgelser, hvis disse støttes af velgennemførte dyreforsøg og mekanistiske undersøgelser.

¹⁹ For kontinuerte helbredseffekter, f.eks. sædkvalitet eller IQ, anvendes effektestimater, der angiver, hvor meget en given ændring i eksponeringen påvirker associationen.

Fakta-boks: Fire typer af epidemiologiske undersøgelser

Populationsbaserede korrelationsstudier (eng.: ecological studies) undersøger sammenhængen mellem mulige risikofaktorer og helbredseffekter i geografiske og/eller tidsmæssige afgrænsede populationer. Risikofaktorer og helbredseffekter udtrykkes som gennemsnitlige forekomster, som for hver geografiske eller temporære subpopulation sammenlignes statistisk. Korrelationsstudier er hypotesedannende og har ofte indikeret årsagssammenhænge før andre epidemiologiske, kliniske eller eksperimentelle studier (f.eks. kost og cancer). Resultaterne bør dog altid valideres af andre studier eller metoder, herunder efter Bradford-Hill kriterierne for årsagssammenhæng i biologiske systemer.

Tværsnitsstudier (eng.: cross-sectional studies) undersøger forekomsten af helbredseffekter, adfærd og/eller risikofaktorer hos individer i en population på undersøgelsestidspunktet. Man kan undersøge sammenhænge, f.eks. sammenhængen mellem en eksponering og et sygdomsudfald, hvis flere forskellige faktorer inkluderes i samme studie. Tværsnitsstudier kan ikke anvendes til at undersøge årsager til helbredseffekter, kun eventuelle sammenhænge, da man ikke ved om en given eksponering indtraf før eller efter effektens/sygdommens opståen. Resultaterne bør derfor valideres af andre studier eller metoder, herunder efter Bradford-Hill kriterierne for årsagssammenhæng i biologiske systemer.

Case-kontrol studier bruges primært til at undersøge hvilke faktorer, der øger risikoen for en sygdom. De udføres ved at sammenligne deltagere med sygdommen (cases) med deltagere uden sygdommen (kontrolpersoner). Derved kan man analysere for, om der er risikofaktorer, som forekommer hyppigere hos cases end hos kontrolpersoner. Da informationer om f.eks. eksponering indsamles retrospektivt, dvs. efter cases er blevet syge, risikerer man, at eksponering (sgraden) klassificeres forkert. Det kan f.eks. ske, hvis nogle deltagere har svært ved at huske væsentlige eksponeringer (recall bias) eller eksponeringsdata er baseret på "proxy data", altså data, der er indsamlet med et andet formål for øje. Studiedesignet er ret følsomt over for fejlkilder og ikke velegnet, hvis eksponeringsfrekvensen er lav.

Kohortestudier følger en gruppe af deltagere (en kohorte) over tid for at undersøge sammenhængen mellem en eller flere risikofaktorer (f.eks. en eksponering) og en helbredseffekt. Kohortestudier er prospektive eller retrospektive, altså fremadskuende eller tilbageskuende ved kortlægning af eksponeringen. Begge typer vil forsøge at bestemme eksponeringen før helbredseffekten indtræffer, men det kan som regel gøres mere nøjagtigt i prospektive studier. Kohortestudier er mindre følsomme over for fejlkilder end case-kontrol studier, men der kan være problemer med systematiske fejl i forbindelse med udvælgelse af deltagere samt frafald af deltagere over tid. Kohortestudier er generelt længere tid om at levere resultater end case-kontrol studier, hvilket især gør sig gældende ved sjældne helbredseffekter.

4. Befolkningsundersøgelser af pesticiders mulige helbredseffekter

I 2013 publicerede EFSA en videnskabelig rapport med en systematisk gennemgang af publicerede befolkningsundersøgelser mellem 2006 og 2012, der viser sammenhæng mellem eksponering for pesticider (også erhvervsmæssig) og helbredseffekter ("Literature review on epidemiological studies linking exposure to pesticides and health effects"²⁰, Ntzani 2013). Helbredseffekterne blev grupperet i 23 hovedkategorier, og i meta-analyser blev der fundet statistisk signifikante sammenhænge mellem pesticideksponering og følgende helbredsudfald: lever-, bryst- og mavekræft, amyotrofisk lateral sklerose, astma, type II diabetes, Parkinsons sygdom og børneleukæmi. Trods et stort antal tilgængelige forskningsartikler og analyser (>6,000) var kvaliteten af undersøgelserne generelt lav og med mange bias i en grad, så artiklens forfattere ikke fandt tilstrækkelig evidens for en årsagssammenhæng for hovedparten af helbredseffekterne. Årsagen skyldes et antal begrænsninger ved befolkningsundersøgelserne (f.eks. lille antal deltagere, selvrapporteret eksponering, retrospektivt design) og en stor forskelligartethed i data. Forfatterne bemærkede særligt, at eksponeringen for pesticider var meget bredt defineret (f.eks. blot "pesticid-eksponering", kategorier som "ukrudtsmidler"/"insektmidler" eller anden gruppering af pesticider med diverse kemiske og toksikologiske egenskaber) i befolkningsundersøgelserne, hvilket begrænsede værdien af de udførte meta-analyser. Brede pesticiddefinitioner nedsætter muligheden for at identificere sammenhænge mellem eksponering for individuelle pesticider eller relaterede grupper af pesticider med effekter på helbredet og dermed at kunne reagere regulatorisk.

I en anden nylig oversigtsartikel (Schüz 2016) er der fundet nogen evidens for, at forældres erhvervsmæssige eksponering for pesticider er en risikofaktor for udvikling af leukæmi hos deres børn, mens der er mindre evidens for en sammenhæng mellem børneleukæmi og pesticideksponering af børn efter fødslen.

Som en opfølgning på det systematiske review blev EFSA's panel vedr. plantebeskyttelsesmidler og deres restkoncentrationer (PPR Panelet) bedt om at udarbejde en "videnskabelig udtalelse" med en vurdering af de metodologiske begrænsninger ved befolkningsundersøgelser af helbredseffekter ved eksponering for pesticider (EFSA PPR Panel 2017a), se afsnit 5.

5. Begrænsninger ved befolkningsundersøgelser af pesticiders mulige helbredseffekter

Der blev identificeret flere metodologiske begrænsninger ud fra de mange befolkningsundersøgelser omhandlende pesticiders mulige helbredseffekter (Ntzani 2013, PPR panel 2017a), som også inkluderer mange af de begrænsninger, der gælder generelt for epidemiologiske undersøgelser. De væsentligste begrænsninger er:

²⁰ Rapportens titel kan oversættes til: "Litteraturgennemgang af epidemiologiske undersøgelser, der kæder eksponering for pesticider sammen med helbredsmæssige virkninger".

- Mangel på prospektive undersøgelser og hyppig brug af forsøgsdesign, der er følsomme over for bias (såsom case-kontrol undersøgelser og tværsnitsundersøgelser).
- Særlig kritisk er en manglende detaljering i vurdering af eksponering for pesticider. Information om eksponering for et specifikt pesticid mangler ofte; tilsvarende er samtidig eksponering for andre pesticider eller andre faktorer (f.eks. i arbejdsmiljøet), ikke (tilstrækkelig) estimeret.
- Mangler i vurderingen af effekter på helbredet (brede udfaldsdefinitioner, selvrapporterede udfald eller surrogatudfald)
- Mangelfuld analyse og afrapportering (tolkning af effekt-estimer, kontrol for "konfunderere" og multipel hypotese-testning, hvor der ikke er justeret for test af adskillige hypoteser)
- Selektiv afrapportering, publikationsbias mv.

En uddybning af nogle af de nævnte begrænsninger samt yderligere observationspunkter følger her:

Begrænsninger i undersøgelsesdesign

Prospektive undersøgelser har størst validitet, men er ikke altid en realistisk mulighed. Det gælder især for sjældne udfald som cancer og fødselsdefekter. For sygdomme med lang latensperiode (dvs. at der går lang tid fra eksponering til sygdommen konstateres), og hvor udvikling af sygdom er forårsaget af længere tids eksponering, vil måling af eksponering til et enkelt tidspunkt ofte ikke reflektere langtidseksponeringen. Dette er især væsentligt for ikke-persistente/letnedbrydelige pesticider, hvis niveau i biologiske prøver, der danner baggrund for eksponeringsvurderingen, ikke er konstant. Undersøgelser, der viser en sammenhæng mellem en eksponering baseret på en enkelt måling i urin og et langtidsudfald, skal derfor tolkes med stor forsigtighed. Dersom der er frosne urinprøver (biobank) indhentet på det relevante eksponeringstidspunkt (f.eks. under graviditet) vil målinger dog kunne give et godt estimat, især hvis eksponeringen antages at have været kontinuerlig i den relevante periode.

Fostre er særligt sårbare, men derudover udgør børn også en særlig sårbar gruppe, idet deres energiomsætning, fysiologi, kost og eksponeringsmønstre afviger fra voksne. Det betyder, at de typisk er udsat for højere eksponeringer og samtidig er de mere følsomme for kemikaliers skadelige effekter. Der kan også være et tidsmæssigt vindue, hvor børn er særligt følsomme for skadevirkninger. Sådanne vinduer er dog typisk ikke kendt og vil oftest afhænge af den sygdomsfremkaldende mekanisme. Forskel i følsomhed mellem køn bør også overvejes ifm. påvirkning af forplantningsevnen og hormonforstyrrende effekter.

Relevans af genetiske variationer i befolkningen

Genetiske variationer vurderes normalt at være af mindre betydning for epidemiologiske undersøgelser. Det er dog værd at nævne, at naturligt forekommende genetiske variationer kan øge eller nedsætte nedbrydning af kemiske stoffer i kroppen, kan prædisponere for en given sygdom, og de kan være særligt udbredt i visse områder. Den genetiske sammensætning i en population kan derfor påvirke om en given sygdom udvikles, og risikoestimerterne for udfaldet vil tilsvarende blive påvirket, hvis ikke der (kan) tages højde

for genetiske variationer i analyserne. Det kan betyde, at undersøgelser udført i forskellige dele af verden kan finde forskellige sammenhænge/udfald.

Udfordringer ved eksponeringsvurdering

De væsentligste begrænsninger ved befolkningsundersøgelser af pesticider skyldes usikkerhed i eksponeringsestimatet. De fleste af de aktivstoffer, der er godkendt i dag, har en kort halveringstid, og de anvendes typisk i flere forskellige pesticidprodukter, hvis anvendelse varierer med afgrødetype og sæson. En præcis vurdering skal altså i mange tilfælde kunne opfange en periodisk langtidseksponering af ikke-persistente stoffer såvel som være i stand til at kvantificere eksponeringen for enkelte pesticider. Bestemmelse af den interne eksponering²¹ via måling af urin-metabolitter (nedbrydningsprodukter) fungerer bedst, når metabolitterne er unikke for enkelte eller grupper af pesticider, og pesticiderne har samme virkningsmekanisme. Her kan urinmålinger afspejle den reelle samlede eksponering fra forskellige kilder (fødevarer, omgivelser, biocidanvendelse osv.). Hvis metabolitterne omvendt er fælles for en større gruppe af helt forskellige pesticider, er vurderingen vanskelig. Det skyldes dels, at potensen af de afledte pesticider kan variere meget og dels, at de pågældende pesticider kan komme fra andre kilder, f.eks. fødevarer, uden relevans for vurderingen i det pågældende studie. Målingen af ”gruppe-metabolitter” kan dog kombineres med andre eksponeringsoplysninger (spørgeskema, pesticidrester i fødevarer osv.), hvorved eksponeringen for enkeltstoffer i gruppen kan kvantificeres. Herved kan der også korrigeres for forskelle i potens. Der må dog forventes en væsentlig grad af usikkerhed forbundet med et eksponeringsestimat udledt fra et større antal ”gruppe-metabolitter”. SAPEA²², der yder videnskabelig rådgivning ifm. politiske beslutningsprocesser, har nyligt udgivet en rapport omhandlende forbedringer til den europæiske proces for godkendelse af pesticider, herunder epidemiologiske undersøgelser. I rapporten konkluderes det, at epidemiologiske undersøgelser skal etablere en sammenhæng mellem helbredseffekter og et specifikt aktivstof eller en gruppe af stoffer, der formodes at have samme toksiske virkemåde, for at kunne bidrage mere direkte til regulatoriske beslutninger (SAPEA 2018).

6. Anvendelse af befolkningsundersøgelser ved risikovurdering af pesticider

6.1. Integrering af befolkningsundersøgelser i risikovurderingsprocessen

Velgennemførte epidemiologiske undersøgelser kan medvirke til at pege på potentielle sammenhænge mellem pesticideksponering og helbredseffekter og er dermed relevante for fareidentifikation (trin 1 i risikovurderingsprocessen, afsnit 2.1). Særligt kan undersøgelserne bidrage med vigtig information om effekter i mennesker, der ikke nødvendigvis fanges af de eksperimentelle undersøgelser, som der kræves af myndighederne. Befolkningsundersøgelser kan også bidrage med information om potentielt øget risiko for særligt sårbare populationer, særligt følsomme perioder i livet og kønsspecifikke effekter.

²¹ Eksponering, hvor der er taget højde for optagelse af stoffet/stofferne i kroppen

²² Science Advice for Policy by European Academies

Hidtil har befolkningsundersøgelser af pesticider generelt ikke kunnet bidrage til vurdering af sammenhænge mellem graden af eksponering og effekt ("dosis-respons"-vurdering), og derfor heller ikke til fastsættelse af de tidligere omtalte referenceværdier (trin 2 i risikovurderingsprocessen, afsnit 2.1). Dette skyldes, at eksponeringsniveauet typisk ikke er kendt eller er upræcist bestemt.

Eksponeringsvurderingen (trin 3 i risikovurderingsprocessen, afsnit 2.1) har som nævnt ovenfor generelt været af lav kvalitet i befolkningsundersøgelser, da eksponeringen typisk er svær at estimere over længere tidsperioder. Det forventes dog, at biomonitoreringsundersøgelser og kvantitativ eksponering for pesticider vil ses fremadrettet, da den teknologiske udvikling vil gøre det muligt ift. pris, testmuligheder og logistik.

Humane data fra forskellige kilder kan dog bruges til at underbygge eller udbygge konklusionerne fra den samlede toksikologiske datapakke for et pesticid, f.eks. i forhold til målorganer og reversibilitet af effekter. Ved fastsættelse af referenceværdier til brug ved risikokarakterisering (trin 4 i risikovurderingsprocessen, afsnit 2.1) er det værd at bemærke, at data fra befolkningsundersøgelser alene kan bruges til at øge sikkerhedsniveauet, da det er specificeret i den gældende forordning 1107/2009 artikel 4(6), at data indsamlet fra mennesker ikke må benyttes til at sænke sikkerhedsmarginen etableret på basis af dyreforsøg.

Overordnet set kan humane data anvendes på flere stadier i risikovurderingsprocessen, men resultater fra enkeltstående eller få epidemiologiske undersøgelser kan normalt ikke anvendes i en regulatorisk kontekst. Hvis resultater fra epidemiologiske undersøgelser skal anvendes ved risikovurdering eller regulatorisk, skal pålideligheden (reabiliteten, fra engelsk "reliability") af undersøgelserne overordnet være høj. Hvordan pålideligheden kan vurderes fremgår af afsnit 6.2.

6.2. Vurdering af pålideligheden ("reabiliteten") af befolkningsundersøgelser

Den overordnede vurdering af pålideligheden af befolkningsundersøgelser bør bero på en ekspertvurdering af den samlede evidens fra undersøgelse til undersøgelse. Det er ikke muligt at opstille rigide og utvetydige kriterier, men der kan dog opstilles en (ikke-udtømmende) række af kvalitetsparametre, der skal vurderes for hver undersøgelse. I Tabel 1 er syv sådanne hovedparametre opsummeret. I tabellen er også angivet en vægtning af kvaliteten af de enkelte parametre i tre kategorier (høj, middel, lav). For hver undersøgelse kan en evidensvægtning af alle parametre give en overordnet vurdering af undersøgelsens pålidelighed. For uddybning af parametrene og inspiration til yderligere relevante faktorer henvises til EFSA PPR Panel 2017a. Analogt til at den overordnede kvalitetsvurdering af laboratorieforsøg med pesticider inddeles i en af tre kategorier som acceptabel, supplerende/understøttende eller ikke-acceptabel, kan befolkningsundersøgelser vurderes på tilsvarende vis ud fra kvalitet: (a) høj pålidelighed (b) middel pålidelighed eller (c) lav pålidelighed. Det vil så kun være undersøgelser af overordnet høj eller middel kvalitet (a og b), der kan indgå i den regulatoriske risikovurdering; undersøgelser af middel kvalitet (b) vil kun kunne betragtes som supplerende. De listede kvalitetsparametre skal betragtes som

retningsgivende, og kategoriseringen af de enkelte parametre kræver nøje ekspertvurdering og skal ses i forhold til hver enkelt befolkningsundersøgelse. Særlig vil kvalitetsvurdering af udfaldet og til dels også eksponeringen afhænge af den specifikke undersøgelse, og det er her ikke muligt at opsætte entydige kvalitetsparametre. Endelig skal det understreges, at en eller flere befolkningsundersøgelser af et pesticid indgår i en samlet risikovurdering af pesticidet; humane data skal integreres med den øvrige toksikologiske datapakke i en samlet vurdering af evidensbyrden (weight of evidence, WoE) (se EFSA Scientific Committee 2017). Dette omtales videre i næste afsnit.

Tabel 1. Kvalitetsparametre til brug ved vægtning af kvalitet af befolkningsundersøgelser (EFSA PPR Panel 2017a)

Parameter	Høj	Middel	Lav
Undersøgelsens design og udførelse	Prospektive kohorteundersøgelser. Præ-specificeret hypotese (stof- og udfaldsspecifik).	Case-kontrol undersøgelser. Retrospektive og prospektive kohorteundersøgelser, hvor designet ikke tilstrækkeligt dækker relevante befolkningsgrupper, forventede fordelinger af eksponering eller udfald.	Tværsnitsundersøgelser, populationsbaserede korrelationsundersøgelser. Case-kontrolundersøgelser, hvor designet ikke tilstrækkeligt dækker relevante befolkningsgrupper, forventede fordelinger af eksponering eller udfald.
Population	Tilfældig udvælgelse af individer. Forsøgspopulationen er stor nok til at sikre tilstrækkelig statistisk styrke. Veldefinerede populationskarakteristika (følsomme undergrupper er inkluderet).	Ikke-tilfældig udvælgelse af individer, f.eks. baseret på frivillighed til at deltage. Undersøgelsens statistiske styrke tvivlsom, ikke detaljeret begrundet. Populationskarakteristika ikke tilstrækkeligt defineret.	Ingen detaljeret information om udvælgelse af undersøgelsespopulationen. Ringedefinition af populationskarakteristika.

Eksponeringsvurdering	Nøjagtig og præcis kvantitativ eksponeringsvurdering (human biomonitoring (HBM) ²³ eller ydre eksponering ²⁴) ved brug af validerede metoder. Valideret spørgeskema ²⁵ og/eller interview for stof-specifik eksponering besvaret af deltagere.	Ikke-valideret proxymål eller biomarkør i en specificeret matrix og ydre eksponering ¹⁶ . Spørgeskema og/eller interview for stof-specifik eksponering besvaret af deltagere eller stedfortrædere.	Ringe proxymål. Spørgeskema og/eller interview af lav kvalitet; information indsamlet for grupper af stoffer. Ingen indsamling af stof-specifikke eksponeringsdata; generel vurdering af pesticidbrug/-eksponering som "engang/aldrig" eller "hyppig/sjælden".
Udfaldsvurdering	Valid og troværdig udfaldsvurdering. Standardiseret og valideret i undersøgelsespopulationen. Lægejournal eller diagnose bekræftet, hvor relevant.	Standardiseret udfald, ikke valideret i populationen, eller screenings-værktøj anvendt; lægejournal ikke bekræftet, hvor relevant.	Ikke-standardiseret og ikke-valideret helbredseffekt. Utilstrækkelige eller selv-rapporterede udfald (Kvalitet af selvrapportering dog udfaldsafhængig).
"Konfounder"-kontrol	Passende kontrol for vigtige "konfoundere" og effektmodifikatorer ²⁶ relevante for det videnskabelige spørgsmål.	Moderat kontrol for "konfoundere" og effektmodifikatorer. Ikke alle variabler relevante for det videnskabelige spørgsmål er taget i betragtning.	Ingen kontrol for potentielle "konfoundere" og effektmodifikatorer i design- og analysefaserne af undersøgelsen. Overjustering er tilsvarende problematisk.
Statistisk analyse	Passende til undersøgelsens design, understøttet af et tilstrækkeligt antal observationer, maksimal udnyttelse af data, og objektiv rapportering.	Acceptable metoder, analytiske valg medførende informationstab, og uklar rapportering. Analyse er udført "post hoc", men det er tydeligt	Kun deskriptiv statistik eller bivariat analyse er lavet. Statistiske sammenligninger af grupper er ikke udført eller uklart beskrevet.

²³ En metode til at vurdere menneskers eksponering for kemikalier ved at måle disse kemikalier, deres nedbrydningsprodukter eller reaktionsprodukter i væv eller prøver som blod, urin eller hår fra mennesker

²⁴ Potentielt eksponering af hud og åndedrætssystemet, der ikke tager højde for optagelse i kroppen

²⁵ Et spørgeskema, der er udviklet til og valideret i den befolkningsgruppe, som skemaet henvender sig til. Validering foretages i en repræsentativ gruppe og skal vise en tilstrækkelig pålidelighed af spørgeskemaet.

²⁶ Faktorer, der påvirker *effekten* af eksponeringen, altså at samme eksponering har forskellig effekt mellem forskellige grupper.

	Statistiske metoder til kontrol af ”konfoundere” er anvendt, og justerede og ikke-justerede estimater er præsenteret. Undergruppe- og interaktionsanalyser er udført.	angivet.	Mangler i analysen (f.eks. justering for multipel testning).
Rapportering	Centrale elementer under ”Materialer og Metoder” og ”Resultater” er rapporteret med tilstrækkelige detaljer, og der er adgang til programmeringskoder og evt. data fra de gennemførte analyser. Antal individer i hver fase af undersøgelsen er rapporteret. En plausibel biologisk mekanisme for en observeret sammenhæng er givet.	Nogle elementer under ”Materialer og Metoder” eller ”Resultater” er ikke rapporteret tilstrækkeligt detaljeret. Fortolkning af resultater er moderat adresseret.	Mangler i rapportering (f.eks. fortolkning af effektestimater, kontrol for ”konfoundere”). Selektiv rapportering. Ingen eller ringe information om andre relevante faktorer, som kan påvirke forholdet mellem eksponering og helbred. Ikke-begrundede konklusioner.

6.3. Relevansvurdering af befolkningsundersøgelser

I dette afsnit beskrives kort en tilgang til at kombinere og summere resultater fra forskellige befolkningsundersøgelser omhandlende en sammenhæng mellem pesticideksponering og sygdom hos mennesker. Tilgangen til vægtning af befolkningsundersøgelser kan tage udgangspunkt i principper som de modificerede Bradford Hill-kriterier. De omfatter en gruppe af kriterier, der samlet kan belyse en potentiel årsagssammenhæng mellem en given eksponering og en observeret effekt (styrke, konsistens, specificitet, temporalitet, biologisk gradient, plausibilitet, overensstemmelse, forsøg, rækkefølge af nøglebegivenheder og analogi). I Tabel 2 er der listet en række overvejelser, der er væsentlige i forhold til de nævnte centrale kriterier for evidensvægtningen.

Relevansen af de enkelte parametre vil afhænge af den konkrete problemstilling og undersøgelse. Overvejelser om betydningen af multikausalitet bør indgå, da multikausalitet f.eks. kan have stor betydning for konsistens og specificitet. Det er også vigtigt at understrege, at der er tale om en vægtning af kriterierne (jf. EFSA Scientific Committee 2017), at biologiske systemer er komplekse, og at en mulig kausal sammenhæng ikke kan udelukkes, hvis alle kriterierne ikke er opfyldt.

Tabel 2. Kriterier til brug ved vægtning af befolkningsundersøgelser for evidens-integrering (baseret på modificerede Bradford Hill-kriterier) (EFSA PPR Panel 2017a)

Kategori	Overvejelser
Sammenhængens styrke og signifikans	Jo stærkere associationen er, des større er sandsynligheden for, at der er tale om en kausal association. Styrken af en sammenhæng (dvs. associationsmålets størrelse og den statistiske signifikans) har dog kun værdi, hvis de tilgrundliggende data og metoder, herunder inddragelse af andre kontekstuelle faktorer og sammenligning med evidensbyrden (WoE) i litteraturen, er af tilstrækkelig høj kvalitet, som gennemgået ovenfor. Statistisk signifikans er ikke det samme som biologisk betydning. Lavere p-værdier angiver ikke nødvendigvis, at effekten er større eller vigtigere, og højere p-værdier angiver ikke nødvendigvis mindre betydelig effekt. Størrelsen af p-værdien afhænger af stikprøvestørrelsen og målepræcisionen (Wasserstein & Lazar 2016). Sammenhængens styrke vurderes derfor også ud fra konfidensintervaller.
Sammenhængens konsistens	Sammenhænge bør generelt være konsistente på tværs af flere uafhængige undersøgelser, særlig de, der er udført med forskellige forsøgsdesigns og i forskellige populationer under forskellige forhold. Dog kan f.eks. multikausalitet forklare en umiddelbar inkonsistens. Tilsvarende bør der også være en høj grad af konsistens af resultater på tværs af alle bevistyper (epidemiologi, dyreforsøg, ”reagensglasforsøg” mv). Mulige årsager til tilsyneladende inkonsistens skal vurderes omhyggeligt og kan bruges til at identificere nye forskningsområder.
Specificitet	Det oprindelige evidenskriterie, hvor forbindelsen af et specifikt udfald med en eksponering kan give et stærkt argument for årsagssammenhæng, er videreudviklet og kan have nye og interessante implikationer i en dataintegrationskontekst. Ved hjælp af dataintegration er det muligt at en molekylær virkningsmekanisme kan identificeres med en præcist defineret sammenhæng mellem ét stof og de forskellige udfald associeret med komplekse eksponeringer. En tilsyneladende mangel på specificitet kan således hjælpe med til at identificere de(t) enkelte stof(fer), der faktisk er associeret med sygdom (Fedak 2015).
Temporalitet (tidsrelation)	Årsagen kommer tidsmæssigt før udfaldet. Bevis for en tidsmæssig rækkefølge mellem eksponering for et stof og fremkomsten af effekten inden for en passende tidsramme udgør et af de bedste argumenter for en årsagssammenhæng. Derfor er undersøgelsesdesigns, som sikrer en tidsmæssig progression mellem de to mål, de mest overbevisende til at vise årsagssammenhæng.
Biologisk gradient (”dosis-respons”)	Øgede effekter associeret med større eksponering eller varighed af eksponering indikerer kraftigt en årsagssammenhæng. En observeret ”dosis-respons” sammenhæng kan dog skyldes confounding, når eksponeringen sam-varierer med den virkelige årsagsfaktor. Fraværet af en sådan association udelukker heller ikke en årsagssammenhæng.
Biologisk plausibilitet	Data, der er forklaret og støttet af sandsynlige biologiske mekanismer baseret på eksperimentel evidens, styrker sandsynligheden for, at en sammenhæng er kausal. Manglende mekanistiske data skal imidlertid ikke ses som et bevis imod en årsagssammenhæng.

Logisk sammenhæng	Fortolkningen afbeviser bøg give mening og ikke være i strid med, hvad der vides om biologien af det pågældende udfald under eksponering-til-sygdom paradigmet. Hvis det gør, bør arten tættest på mennesket betragtes som den mere relevante.
Eksperimentel evidens	Resultater fra randomiserede eksperimenter yder stærkere bevis for en kausal sammenhæng, frem for resultater baseret på andre forsøgsdesigns. Alternativt kan en sammenhæng fra en ikke-eksperimentel undersøgelse overvejes som kausal, hvis en randomiseret forebyggelse (eliminering/minimering af eksponeringen) bekræfter fundet.
Rækkefølge af nøglebegivenheder	Der gives en klar beskrivelse af hver nøglebegivenhed (dvs. målbare parametre fra en kombination af ”reagensglasforsøg”, dyreforsøg eller humane datakilder), som ligger til grund for den etablerede ”mode of action” ²⁷ (MoA)/”adverse outcome pathway” ²⁸ (AOP) for en specifik helbredseffekt. En fuldt belyst MoA/AOP er ikke påkrævet for brug af befolkningsundersøgelser i sundhedsrisikovurderingen.
Analogi	Ved stærk evidens for en årsagssammenhæng mellem et specifikt stof og et givent udfald, bør en svagere evidens, for at et lignende (analogt) stof kan forårsage et tilsvarende udfald, accepteres. I dag er der mange værktøjer til at søge efter og finde analoger, inklusiv mønstre i sygdomsprogression, typiske risikofaktorer og konfunderere samt biologisk virkemåder, og analogi bør derfor i dag også bruges til at fremsætte og teste mekanistiske hypoteser.

6.4. Vurdering af den biologiske relevans af en statistisk signifikant sammenhæng

Befolkningsundersøgelser bidrager i stigende omfang til etablering af (årsags)sammenhæng, og et vigtigt element i denne forbindelse er etablering af biologisk sandsynlighed, der dog er afhængig af forståelsen af sygdomsmekanismer. Som opfølgning på litteraturgennemgangen af befolkningsundersøgelser, der kæder humane helbredseffekter sammen med pesticideksponering (Ntzani 2013), blev EFSA’s panel vedr. plantebeskyttelsesmidler og deres restkoncentrationer (PPR Panelet) bedt om at undersøge den sandsynlige medvirken af pesticideksponering som en risikofaktor for udvikling af Parkinsons sygdom og børneleukæmi. Panelet anvendte den konceptuelle ramme med begivenhedskæder, der leder fra en initierende molekylær hændelse til et skadeligt udfald for organismen, kendt som ”adverse outcome pathway” (AOP), til at undersøge den biologiske sandsynlighed for statistiske sammenhænge fundet i befolkningsundersøgelser ved at bruge specifikke symptomer ved sygdommene som et ”skadeligt udfald” (”adverse outcome”) (EFSA PPR Panel 2017a). AOP’er er en form for systematisk review af data, hvor der laves en evidensvægtning baseret på Bradford Hill-kriterier med henblik på regulatorisk anvendelse.

²⁷ (biologisk) virkemåde

²⁸ En konceptuel ramme med begivenhedskæder der leder fra en initierende molekylær hændelse til et skadeligt udfald for organismen

PPR Panelet fik udviklet to kvalitative AOP'er for Parkinsons sygdom, en kvalitativ AOP for leukæmi hos spædbørn ("infant leukemia") samt en formodet AOP for leukæmi hos mindre børn ("childhood leukemia") (EFSA PPR Panel 2017a). AOP'er giver en mekanistisk forståelse af helbredseffekter. Arbejdet med AOP-udvikling varetages af OECD. Det skal understreges, at der ikke stilles krav om mekanisme eller udvikling af en AOP, for at en epidemiologisk undersøgelse kan anvendes i den regulatoriske risikovurdering; men det er selvsagt en styrke, hvis der er en virkningsmæssig underbygning af en epidemiologisk undersøgelse.

7. Forbedring af pesticid-epidemiologiske undersøgelser ift. regulatorisk anvendelse

Eksponeringsvurderingen er ofte et svagt punkt, og bedre metoder til at estimere eksponeringen vil styrke mulighederne for at afdække mulige helbredseffekter. Human biomonitoring (HBM) er et vigtigt redskab, men indebærer en del udfordringer, da pesticider, som er godkendt i dag, har kort biologisk halveringstid og stort set vil være ude af kroppen efter 1-2 døgn. Netop pga. stoffernes hurtige omsætning i kroppen vil HBM typisk kun give en eller flere punktmålinger, da kontinuerlig måling (endnu) ikke er mulig. Ved erhvervsmæssig eksponering, hvor man kender tidspunkt for eksponering og hvilke specifikke pesticider, der er anvendt, er HBM af specifikke pesticider/metabolitter en mulighed. Hvis der skal opnås et pålideligt mål for en samlet langtidseksponering skal eksponeringsvurderingen vha. HBM gentages regelmæssigt over en lang periode (SAPEA 2018). Anbefalingen er derfor, at det kun i visse tilfælde vil være hensigtsmæssigt med HBM, hvis der er tale om punktvis erhvervsmæssige eksponeringer. Ved en mere kontinuerlig eksponering, som f.eks. fra pesticidrester i fødevarer (og sandsynligvis også fra pesticider i husstøv), vil HBM dog kunne være et godt redskab til at vurdere den interne eksponering.

Ud over biomarkører, der reflekterer eksponeringen, så kunne der udvikles nye biomarkører for effekt, der repræsenterer tidlige stadier i sygdomsudviklingen. Det kræver dog, at man har kendskab til sygdomsmekanismen.

I forhold til øget kvalitet af rapportering af epidemiologiske undersøgelser anbefales det at følge de såkaldte STROBE²⁹-retningslinjer (von Elm 2007). Der er udviklet en fælles tjekliste³⁰ for kohorte, case-kontrol og tværsnitsundersøgelser over elementer, der bør inkluderes ved afrapportering af undersøgelserne. Tjeklisten kunne videreudvikles, så den specifikt adresserer undersøgelser vedrørende eksponering for pesticider og helbredseffekter.

Det er desuden væsentligt, at en epidemiologisk undersøgelse er dimensioneret til at kunne detektere ændring i risici for de effekter, som det undersøger. Det kan gøres ved at vurdere studiets statistiske styrke ("power"), f.eks. om studiet har styrke til at påvise 20 % forøgelse i

²⁹ STrengthening the Reporting of OBservational studies in Epidemiology (Styrkelse af rapporteringen af beskrivende epidemiologiske undersøgelser)

³⁰ https://www.strobe-statement.org/fileadmin/Strobe/uploads/checklists/STROBE_checklist_v4_combined.pdf

risikoen for en given effekt. Dette er primært relevant for effekter, der er ”enten-eller”, f.eks. om et barn har leukæmi eller ej, da proceduren er vanskeligere for kontinuerte data (f.eks. sædkvalitet, IQ osv.) (Pejtersen 2015). Tilsvarende vigtig er vurderingen af den statistiske styrke af meta-analyser af flere uafhængige epidemiologiske undersøgelser.

Relevante forskningsområder til forbedring af pesticid-epidemiologiske undersøgelser kan inkluderes i Miljøstyrelsens Bekæmpelsesmiddelforskningsprogram.

8. Konklusion

For at få et pesticid-aktivstof godkendt kræves en række undersøgelser til afdækning af stoffets toksikologiske profil. Undersøgelserne er primært baseret på dyreforsøg, men der er også krav om, at eksisterende humane data skal inddrages. Sidstnævnte er væsentligt, da mere komplekse sygdomsbilleder hos mennesker ikke altid er dækket af de påkrævede dyreforsøg. Epidemiologiske undersøgelser kan derfor bidrage med vigtig viden om helbredseffekter ved eksponering for pesticider. Pålideligheden af epidemiologiske undersøgelser såvel som af andre data (dyreforsøg og reagensglasforsøg eksempelvis) skal være tilstrækkelig høj, for at de kan indgå i risikovurderingen af pesticider. Begrænsninger ved (pesticid-)epidemiologiske undersøgelser er gennemgået i notatet, hvoraf den væsentligste er usikkerhed i eksponeringsestimater. Der er også opstillet en række kvalitetsparametre (Tabel 1), der vægtes kan bruges ved vurdering af pålideligheden af den enkelte undersøgelse. Desuden kan flere undersøgelser kombineres for bedre at belyse en potentiel årsagssammenhæng mellem en given eksponering og en observeret effekt, hvilket kan ske ud fra vægtning af en række kriterier (Tabel 2). Her er det vigtigt, at de enkelte undersøgelser rent faktisk har statistisk styrke til at detektere ændring i risiko for de helbredseffekter, de undersøger. Befolkningsundersøgelser kan pege på en sammenhæng mellem eksponering for et pesticid og sygdomsudvikling. Det styrker undersøgelsens resultat, hvis der kan etableres en årsagssammenhæng via en biologisk mekanisme, hvilket dog forudsætter forståelse af den underliggende sygdomsmekanisme. En sådan plausibel korrelation er mere overbevisende. Plausible korrelationer, normalt baseret på mange undersøgelser, hvor bias, ”konfounding” og tilfældighed med rimelighed kan udelukkes, bliver betragtet som kausale. Kausale korrelationer, ”årsagssammenhænge”, er mest relevante – også regulatorisk. Der er dog ikke krav om en beskrevet mekanisme, for at en epidemiologisk undersøgelse kan anvendes i den regulatoriske risikovurdering. Evidensbyrden fra humane data skal integreres med den øvrige toksikologiske datapakke i den samlede risikovurdering.

Notatet afsluttes med en række anbefalinger til forbedring af epidemiologiske undersøgelser af pesticiders mulige helbredseffekter, så de kan anvendes bedst muligt i det regulatoriske arbejde.

9. Referencer

EFSA Scientific Committee 2017. Scientific Opinion on the guidance on the use of the weight of evidence approach in scientific assessments. *EFSA Journal* 2017;15(8):4971, 69 pp.

<https://doi.org/10.2903/j.efsa.2017.4971>

EFSA PPR Panel 2017a. Scientific Opinion of the PPR Panel on the follow-up of the findings of the External Scientific Report 'Literature review of epidemiological studies linking exposure to pesticides and health effects'. *EFSA Journal* 2017;15(10):5007

<https://doi.org/10.2903/j.efsa.2017.5007>

EFSA PPR Panel 2017b. Scientific Opinion on the investigation into experimental toxicological properties of plant protection products having a potential link to Parkinson's disease and childhood leukaemia. *EFSA Journal* 2017;15(3):4691, 325 pp.

<https://doi.org/10.2903/j.efsa.2017.4691>

Fedak 2015. Fedak KM, Bernal A, Capshaw ZA, Gross S. Applying the Bradford Hill criteria in the 21st century: how data integration has changed causal inference in molecular epidemiology. *Emerg Themes Epidemiol* 2015;12:14

<https://doi.org/10.1186/s12982-015-0037-4>

von Elm 2007. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: Guidelines for Reporting Observational Studies. *PLoS Med* 4(10): e296.

<https://doi.org/10.1371/journal.pmed.0040296>

Ntzani 2013. Ntzani EE, Chondrogiorgi M, Ntritsos G, Evangelou E, Tzoulaki I. Literature review on epidemiological studies linking exposure to pesticides and health effects. *EFSA supporting publication* 2013:EN-497, 159 pp.

<https://doi.org/10.2903/sp.efsa.2013.EN-497>

Pejtersen 2015. Pejtersen JH, Burr H, Hannerz H, Fishta A, Eller NA. Update on Work-Related Psychosocial Factors and the Development of Ischemic Heart Disease: A Systematic Review. *Cardiol Rev* 2015;23(2):94-8

<https://dx.doi.org/10.1097/CRD.000000000000033>

SAPEA 2018. Improving authorisation processes for plant protection products in Europe: a scientific perspective on the potential risks to human health. Berlin: SAPEA.

<https://doi.org/10.26356/plantprotectionproducts>

Schüz 2016. Schüz J, Erdmann F. Environmental Exposure and Risk of Childhood Leukemia: An Overview. *Arch Med Res.* 2016;47(8):607-614

<https://doi.org/10.1016/j.arcmed.2016.11.017>

Wasserstein & Lazar 2016. The ASA's Statement on p-Values: Context, Process, and Purpose. *The American Statistician*, 70:2, 129 pp.

<https://doi.org/10.1080/00031305.2016.1154108>

10. Ordliste og forkortelser

- AAOEL: Acute Acceptable Operator Exposure Level (acceptabelt akut eksponeringsniveau af sprøjteføreren)
- ADI: Acceptabelt Dagligt Indtag
- AOEL: Acceptable Operator Exposure Level (acceptabelt eksponeringsniveau af sprøjteføreren)
- AOP: Adverse Outcome Pathway; en konceptuel ramme med begivenhedskæder der leder fra en initierende molekylær hændelse til et skadeligt udfald for organismen
- ARfD: Akut ReferenceDosis
- CLP: Classification, Labelling and Packaging (klassificering, mærkning og emballering)
- DOHaD: Developmental Origins of Health and Disease (Føtal oprindelse af sundhed og sygdom)
- EFSA: European Food Safety Authority (Den europæiske fødevarerikkerhedsautoritet)
- GLP: Good Laboratory Practise (god laboratoriepraksis)
- HBM: Human BioMonitoring; en metode til at vurdere menneskers eksponering for kemikalier ved at måle disse kemikalier, deres nedbrydningsprodukter eller reaktionsprodukter i væv eller prøver som blod, urin eller hår fra mennesker
- MoA: Mode of Action; virkemåde
- OECD: The Organisation for Economic Co-operation and Development (Organisationen for Økonomisk Samarbejde og Udvikling)
- OR: Odds Ratio (chancen for at et udfald indtræder ved en given eksponering sammenholdt med chancen for at udfaldet indtræder i fraværet af eksponeringen)
- PPR Panel: Panel on Plant Protection products and their Residues (Panelet for plantebeskyttelsesmidler og deres restkoncentrationer)
- REACH: Registration, Evaluation, Authorisation and restriction of CHemicals (registrering, vurdering og godkendelse af samt begrænsninger for kemikalier. EU's generelle kemikalieforskrift)
- RR: Relative Risk (den relative risiko)

- SAPEA:** Science Advice for Policy by European Academies
(Videnskabelig rådgivning for politikudvikling fra europæiske akademier)
- Synergist:** Et stof eller præparat, der ikke selv er virksomt (eller kun er det i svag grad), men som kan gøre aktivstoffet eller aktivstofferne i et pesticidprodukt mere virksomme
- Safener:** Et stof eller præparat, der tilsættes et pesticidprodukt for at fjerne eller reducere pesticidproduktets giftige virkning på en afgrøde
- STROBE** STrengthening the Reporting of OBservational studies in Epidemiology
(Styrkelse af rapporteringen af beskrivende epidemiologiske undersøgelser)
- WoE:** weight of evidence; evidensbyrde/beviskraft