



Demagogisch Zembla wekt onterecht twijfel aan toelaatbaarheid glyfosaat

Posted on 17 mei 2025 by Arnout Jaspers

Zembla, het documentaire programma van BNN/Vara, heeft een reputatie hoog te houden van actiejournalistiek, waarbij er nauwelijks een ondergrens is aan de demagogie en manipulatie die de redactie toelaatbaar acht om de gewenste boodschap uit te dragen. Zelfs menige collega-journalist wil niets meer met ze te maken hebben.

Een van de dingen waar *Zembla* campagne tegen voert, is glyfosaat (het werkzame bestanddeel van Round-Up). Dit onkruidbestrijdingsmiddel wordt in de landbouw veel gebruikt om al het onkruid op de akkers te verdelgen, zodat daarna met een schone lei begonnen kan worden voor de volgende oogst (glyfosaat wordt in de bodem snel afgebroken, zodat het zaai gewas er geen last van heeft).

Glyfosaat opnieuw voor tien jaar toegelaten

Zoals alle bestrijdingsmiddelen in de EU, moet ook glyfosaat periodiek opnieuw beoordeeld worden op veiligheid voor mens en natuur. Na uitvoerig onderzoek door

vier EU-landen, waaronder het CTGB in Nederland, is glyfosaat in november 2023 opnieuw voor tien jaar toegelaten. Daarmee sluit de EU zich aan bij de VS, Australië, Japan en nog veel meer landen.

In een wereld waar respect voor feiten en rationele argumenten leidend is, zou het glyfosaat-verhaal nu klaar zijn. Maar daarvoor moet je niet bij de milieubeweging en hun spreekbuizen binnen de NPO zijn. Daar heerst onverkort het dogma: glyfosaat is des duivels, [dus het moet en zal verboden worden](#).

Zembla doet nu alsof ze de *smoking gun* gevonden hebben die bewijst dat glyfosaat nooit had mogen worden goedgekeurd. Op hun website [kopten](#) ze afgelopen week: 'Ctgb erkent: decennialang niet streng genoeg bij beoordelen kankerrisico's pesticiden'. Dat is volgens vertrouwd Zembla-recept: schaamteloos generaliseren, selectief citeren en uit z'n verband rukken. Wat er feitelijk aan de hand is: in februari hield het CTGB een expert-bijeenkomst over een- of tweezijdig toetsen van de resultaten van dierproeven, mede omdat *Zembla* dit technische detail naar boven haalde in hun eerdere uitzending. Hoewel tweezijdig toetsen nu de standaard is, zeiden de experts bij de CTGB-bijeenkomst dat eenzijdig toetsen bij het onderzoek naar mogelijke carcinogeniteit van een stof de voorkeur verdient.

Het is op zich een interessant discussiepunt: wat is de juiste statistische toets om data over proefdieronderzoek te analyseren? Als je wilt weten of een stof kankerverwekkend is, geef je de ene groep proefdieren (vaak ratten of muizen) een dosis glyfosaat mee in hun voer, en een ongeveer even grote controlegroep krijgt schoon voer. En dan registreer je gedurende een vooraf vastgelegde periode of een van de groepen vaker kanker krijgt, en welk type kanker dat dan is.

Een-zijdig en twee-zijdig toetsen

Zulk onderzoek levert nooit op, dat alle honderd glyfosaat-ratten binnen twee jaar kanker krijgen, tegen geen enkele van de honderd controle-muizen. Dan had je geen statistische toets nodig. Je zult altijd iets vinden (fictief voorbeeld) in de trant van: in de glyfosaat-groep kregen 14 muizen lymfeklierkanker, en in de controlegroep 11 muizen. Bewijst dit dat glyfosaat lymfeklierkanker veroorzaakt?

Je moet dan een statistische toets doen, met als uitkomst dat dit resultaat wel of niet 'significant' is. Die toets kent twee varianten: je kunt een-zijdig of twee-zijdig toetsen. Verderop wordt tot in detail, met getallenvoorbeelden, uitgelegd hoe dit

werkt. Maar waar *Zembla* nu op inzoomt, is dat bij zulk proefdieronderzoek doorgaans twee-zijdig getoetst wordt, terwijl je bij een een-zijdige toets wat sneller zou concluderen dat een resultaat significant is.

Statistische ongecijferdheid

Dan is het in principe mogelijk, dat als je de data van bovenstaand experiment twee-zijdig test, het oordeel luidt: niet significant (dus geen bewijs voor carcinogeniteit van glyfosaat), terwijl dat oordeel bij eenzijdig testen kan zijn: wel significant.

Zembla maakt daar van, dat je met tweezijdig toetsen twee keer zoveel ratten met kanker nodig hebt om een resultaat als significant te kwalificeren dan als je eenzijdig toetst (dus 28 tegen 11 in dit voorbeeld, als 14 tegen 11 net significant zou zijn bij eenzijdig toetsen). Dit is gewoon nonsens, statistische ongecijferdheid, zo werkt die toets niet.

Essentieel is, dat er niet één juiste methode van toetsen is, met de implicatie dat alle andere fout zijn. Hoe en wat je toetst, en wanneer je iets 'statistisch significant' noemt, hangt van allerlei keuzes af die te maken hebben met de opzet van het onderzoek, en zelfs met de opvattingen van de onderzoeker. Dat betekent zeker niet, dat je elke uitkomst uit een experiment kunt krijgen die je wilt; als het resultaat overduidelijk is (zeg, 55 tegen 11), dan zal elke toets concluderen dat hier echt iets aan de hand is. De interpretatiekwestie speelt pas bij de twijfelgevallen die net wel of net niet 'significant' zijn.

Bij tweezijdig toetsen is de vraag: wijkt het aantal glyfosaat-ratten met kanker significant af van het aantal controle-ratten met kanker? Bij eenzijdig toetsen is de vraag: is het aantal glyfosaat-ratten met kanker significant *hoger* dan het aantal controle-ratten met kanker? Het verschil is, dat tweezijdig toetsen de mogelijkheid open laat, dat de glyfosaat-ratten *minder* kanker krijgen dan de controle-ratten. Dat komt inderdaad voor bij zulke experimenten: door stom toeval, of omdat de toegediende stof werkelijk een beschermend effect heeft.

Dierproeven zijn slechts een ingrediënt in

onderzoek

Als je die mogelijkheid uitsluit, mag je al bij een iets kleiner overschot aan tumoren bij de glyfosaat-ratten concluderen dat dit 'significant' is. Maar wat is 'significant'? Het is een conventie uit de pre-historie van de statistiek, dat de grens voor een 'significant' resultaat ligt bij een kans van 5% dat dit resultaat toch puur door toeval tot stand gekomen is. Dus of je nu een- of tweezijdig test, er is een kans van 5% dat het 'significante' resultaat niet meer dan een toevallige fluctuatie is.

Dat is een zeer lage drempel om zulke verstrekkende conclusies te trekken: je wilt immers 99,999% zeker weten dat glyfosaat niet kankerverwekkend is bij mensen. Daarom zijn zulke dierproeven niet meer dan een van de vele ingrediënten in een compleet pakket aan onderzoek, op grond waarvan het CTGB en vergelijkbare instanties in nog drie EU-landen geconcludeerd hebben dat glyfosaat geen kanker veroorzaakt bij mensen.

Zulke dierproeven worden nu meestal door de onderzoekers tweezijdig getoetst. Toetsen verandert uiteraard niets aan de oorspronkelijke data, dus het CTGB zou deze data met de eenzijdige toets kunnen heranalyseren. Zolang die heranalyse niet gedaan is, is het een slag in de lucht om te beweren dat eenzijdig toetsen bij deze dierproeven wel zou hebben aangetoond dat glyfosaat kankerverwekkend is bij ratten.

Drie experimenten

Ter illustratie hoe subtiel het verschil tussen een- en tweezijdig toetsen is, heb ik Grok drie experimenten laten doorrekenen om te onderzoeken of linkshandigen een hoger IQ hebben dan rechtshandigen. Bij het eerste experiment scoorden 36 linkshandige proefpersonen een gemiddeld IQ van 104 (rechtshandigen hebben een gemiddelde van 100, met standaarddeviatie 15). Zowel de een- als de tweezijdige toets zeggen hier: dit verschil is niet significant (de kans dat er bij toeval twee of drie extreme slimmeriken onder die 36 zitten, is te groot).

Tweede experiment: 36 linkshandigen scoorden een gemiddeld IQ van 105. Nu zegt de tweezijdige toets nog steeds dat dit niet significant is, maar volgens de eenzijdige toets is dit wel significant.

Derde experiment: 36 linkshandigen scoorden een gemiddeld IQ van 106. Dit

resultaat is zowel volgens de een- als de tweezijdige toets significant.

We zien hier dus, hoe 2 IQ-punten in het gemiddelde bij zowel de een- als de tweezijdige test het verschil maakt tussen wel of niet significant, maar dat de eenzijdige test wat eerder omslaat in zijn oordeel.

Terugvertaald naar de glyfosaat-dierproeven: er is geen sprake van, dat het gangbare tweezijdige toetsen experimenten onder de mat veegt waarin proefdieren overduidelijk vaker kanker kregen. Het gaat om subtiele verschillen. Zembla stelt dat de eenzijdige toets 'strenger' zou zijn, maar strikt genomen is die toets juist slapper: de lat om een resultaat significant te noemen ligt lager dan bij de tweezijdige test.

Obsessie

Stel, je onderzoekt of een nieuw medicijn werkt tegen kanker, en je doet een vergelijkbare dierproef. Dan kun je een- of tweezijdig toetsen of dat medicijn werkt (dat wil zeggen: significant minder kanker bij de dieren die het medicijn kregen, vergeleken met de dieren die een placebo kregen).

Dan kan het best zijn, dat de eenzijdige toets zegt: het medicijn werkt, terwijl de tweezijdige toets zegt: dat is nog niet bewezen. Als we moeten besluiten of we dit medicijn gaan toelaten op de markt en voorschrijven aan patiënten, zou u dan meer vertrouwen hebben in de eenzijdige of de tweezijdige toets?

Er is geen antwoord dat in alle gevallen het beste, of het enige wetenschappelijk verantwoorde antwoord is. Hoe dan ook is de obsessie van *Zembla* met dit zeer technische aspect van de statistiek een storm in een glas water, vanwege al het andersoortige onderzoek dat ook gedaan is om de veiligheid van glyfosaat te onderbouwen.

Wynia's Week verschijnt drie keer per week, **156 keer per jaar**, met even onafhankelijke als broodnodige artikelen en columns, video's en podcasts. U maakt dat samen met de andere donateurs mogelijk. Doet u weer mee? Kijk [**HIER**](#).
Hartelijk dank!