**Statistische significantie zegt niks over een effect of de grootte van een effect. Statistische significantie zoals deze op vandaag wordt gebruikt gaat ervan uit dat als een effect groter is dan een bepaalde grenswaarde het effect te groot is om toe te schrijven aan kans. Er moet dan meer aan de hand zijn. Statistische significantie wil niet zeggen dat het effect causaal is. Daar zitten een hele hoop haken en ogen aan. En De Snoo maakt zich schuldig aan exact hetzelfde als waar hij de pesticiden-industrie van beschuldigd, namelijk rekenen met voorbedachten rade. Als je namelijk de tweezijdige toets vervangt door de eenzijdige toets, dan verhoog je ook de kans op vals-positieven. Dat is waarom altijd tweezijdig wordt getoetst, of eenzijdig met dezelfde grenswaarde als voor de tweezijdige toets. Die grenswaarde is trouwens een KEUZE en ik kan laten zien hoe die keuze van invloed is op de uitkomst.**

**Verder zit de kracht van het vaststellen of een middel causale gevolgen heeft (zoals het ontstaan van kanker) zit hem niet in de statistiek, of de gebruikte statistische toets, maar in het experiment zelf: de randomisatie procedure, de grootte van het aantal dieren. Dat muizen of ratten kanker krijgen betekent nog niet dat wij kanker krijgen. Daarom zijn er ook studies gedaan bij mensen. Maar daar wordt niks over gezegd. Alles wordt gegooid op de keuze van de statistiek. En ik wil laten zien dat dit inderdaad van invloed is, maar aan beide kanten snijdt.**

**De keuze voor het tweezijdig toetsen komt voor uit het idee wat we geen idee hebben welke kant het effect op kan gaan. De Snoo heeft een punt dat je niet verwacht dat een pesticide kankerberschermend is. Je wilt dus één kant op toetsen. Maar die toets heeft een grotere kans op vals-positieven en dat wil je niet. De reden voor tweezijdig toetsen, ook al gaat het effect maar één kant op, is dus om het aantal vals-positieven klein te houden, maar ook om de nulhypothese tegengewicht te bieden. De nulhypothese is namelijk al problematisch genoeg. In de frequentistische statistiek wordt vaak als volgt geredeneerd: als een verschil tussen groepen niet statistisch significant is, dan is er GEEN verschil tussen de groepen. Dat is een onjuiste aanname. De hele manier van frequentistisch toetsen is daarmee zwaar belemmerend. Er wordt vaak gebruikt gemaakt van aannames, zoals normaliteit, die je niet kunt waarmaken. Het frequentistisch toetsen, tweezijdig of eenzijdig, is op zichzelf al limiterend voor beide partijen.**

**Als ik de tweezijdige toets vervang door de eenzijdige toets met eenzelfde criterium als voor de tweezijdige toets, dan verandert er niks aan de bevindingen. In plaats van te focussen op het vervangen van de standaard tweezijdige toets door de standaard eenzijdige toets (waarbij de bevindingen WEL zullen veranderen), is het veel belangrijker om te laten zien hoe de beschuldiging van de een (aannames voor het eigen gewin van de pesticiden industrie) ook teruggekaatst naar de ander (aannames voor het eigen gewin van de milieubeweging). Eigenlijk zeggen de bezwaarmakers hier dat we de grens moeten verlagen om iets causaal kankerverwekkend te noemen. Maar of iets causaal is laat zich nooit statistisch bepalen. En de oproep om een andere toets te gebruiken riekt ook naar een keuze die komt vanuit een sterk geloof. Dat is geen wetenschap.**

**Ik snap de verdunning-uitspraak, maar dat heeft niks te maken met de grootte van het effect, maar met de grens van het statistisch toelaatbare. Met andere woorden: het criterium waaronder we een verschil wel of niet statistisch significant noemen. De Snoo wil graag dat de ene standaard wordt vervangen door de andere standaard in de wetenschap dat de andere standaard hem zal helpen. Althans, daar lijkt het op, maar wat ik wil laten zien is wat er gebeurd als je onder verschillende assumpties naar dezelfde data gaat kijken. Ik wil laten zien dat zijn manier van kijken en argumenteren ook feilbaar is. Ideologisch, niet wetenschappelijk.**

**Ik zou graag willen laten zien wat er van de uitkomsten overblijft als ik verschillende statistische toetsen gebruik. Ik zou willen laten zien dat tekortkomingen die zij opnoemen ook voor hen geldend zijn. Ik wil laten zien wat er gebeurt als ik aannames verander. Ik wil laten zien onder welke voorwaarden iets wel of niet statistisch significant wordt genoemd.** **Ik wil laten zien dat statistisch significant niks te maken heeft met causaal. Ik wil het Bayesiaans aanvliegen waarin ik statistische significantie laat vieren en kijk naar de grootte van een effect onder verschillende aannames op basis van de data. Ook wil ik kijken naar het met menselijk onderzoek en de relatie tussen pesticiden en kanker? Welke statistiek wordt daar gebruikt. Maar het allerliefst wil ik korte metten maken met zinnen zoals deze:**