



Waarschijnlijk baseer jij veel beslissingen op statistisch geschatte kansen. Als komend weekend de kans op mooi weer groot is, besluit je misschien wel om naar het strand te gaan. Wanneer je een vakantie boekt, doe je er wellicht een annuleringsverzekering bij. En de keuze voor of tegen vaccinatie kan ook te maken hebben met wat jij onder een kleine kans op bijwerkingen verstaat.

SANNE WILLEMS

Waarschijnlijk, wellicht, misschien, grote kans. Dit zijn allemaal manieren om een kans in woorden uit te leggen. Ook professionals doen dat, vooral in gesprekken. Je arts vertelt bijvoorbeeld dat je een grote kans hebt op genezing, maar dat mensen zoals jij vaak last hebben van

bijwerkingen van het medicijn.

Maar wat bedoelt de arts dan precies met deze woorden 'grote kans' en 'vaak'? Welke kans heeft deze arts in gedachten? En vooral: komt deze kans overeen met hoe jij hem interpreteert? En zou een andere arts dezelfde

woorden kiezen voor die kansen? Om inzicht te krijgen in de verschillen onderzocht ik samen met mijn collega's Ionica Smeets en Casper Albers de variatie in de interpretatie van kans- en frequentiewoorden.

Waarom in woorden?

De toekomst is onzeker. Als we toch voorspellingen maken, bijvoorbeeld over de kans op regen, dan doen we dat met statistische modellen. Deze modellen gebruiken gegevens uit het verleden, maar die bieden natuurlijk geen garantie voor de toekomst. Een voorspelling blijft altijd een schatting en die is nooit exact.

Deze onzekerheid is een reden om terughoudend te zijn in het geven van exacte getallen. Dat zou namelijk de verwachting kunnen scheppen dat de kans precies bekend is, in plaats van geschat. Vooral in gesprekken gebruiken we daarom vaak kanswoorden (onwaarschijnlijk, grote kans, misschien) of frequentiewoorden (vaak, nooit, meestal). Deze uitspraken geven indirect al de statistische onzekerheid aan.

Maar juist door die indirecte onzekerheid kan er verwarring ontstaan. Want wat als je arts met grote kans 90 procent kans op genezing bedoelt, en jij denkt dat de kans 'maar' 60 procent is? Zo'n verschil in interpretatie kan ervoor zorgen dat jij sterk twijfelt over

een behandeling die eigenlijk een erg grote kans van slagen heeft.

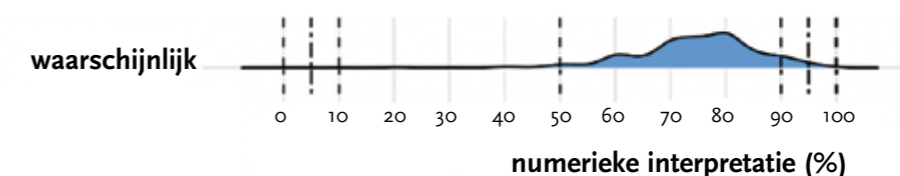
Grote variatie in interpretatie

Stel je leest de zin 'Waarschijnlijk wint het team de wedstrijd', hoe groot denk jij dat dan de kans is dat het team de wedstrijd wint? In ons onderzoek legden wij dat soort vragen voor aan onze deelnemers, elke zin met een ander kans- of frequentiewoord.

In figuur 1 zie je de resultaten voor waarschijnlijk. Op de horizontale as staan alle percentages van 0 tot 100 procent en de hoogte van de grafiek geeft aan hoeveel deelnemers een bepaald percentage kozen. Voorbeeld: de grafiek is het hoogst bij 80 procent, dit betekent dat dat het meest gekozen percentage is voor de interpretatie van waarschijnlijk.

De grafiek is erg breed. Hij begint bij ongeveer 45 procent en heeft een lange staart die doorloopt tot 100 procent. De interpretatie van het woord waarschijnlijk is dus erg gevarieerd.

Tussen 60 procent en 90 procent is de grafiek vrij hoog. Dus als jij waarschijnlijk gebruikt om een kans van 90 procent aan te duiden, is het niet ondenkbaar dat iemand anders dit interpreteert als een kans van nog geen 70 procent. Dat kan dus veel verwarring opleveren.



Figuur 1. De verdeling (dichtheidfuncties) van de interpretatie van kanswoord 'waarschijnlijk'; dit is een soort uitgesmeerd histogram, hierdoor vloeit de grafiek uit buiten de grenzen 0 en 100 procent

@Sanne: Dat de grafiek buiten de grenzen vloeit lijkt in tegenspraak met de lopende tekst.

Figuur 2 geeft de resultaten van alle onderzochte kans- en frequentiewoorden weer. Bij de meeste uitspraken zie je hetzelfde verschijnsel: de grafieken zijn heel breed. Er zitten dus grote verschillen in de interpretatie van veel van deze woorden.

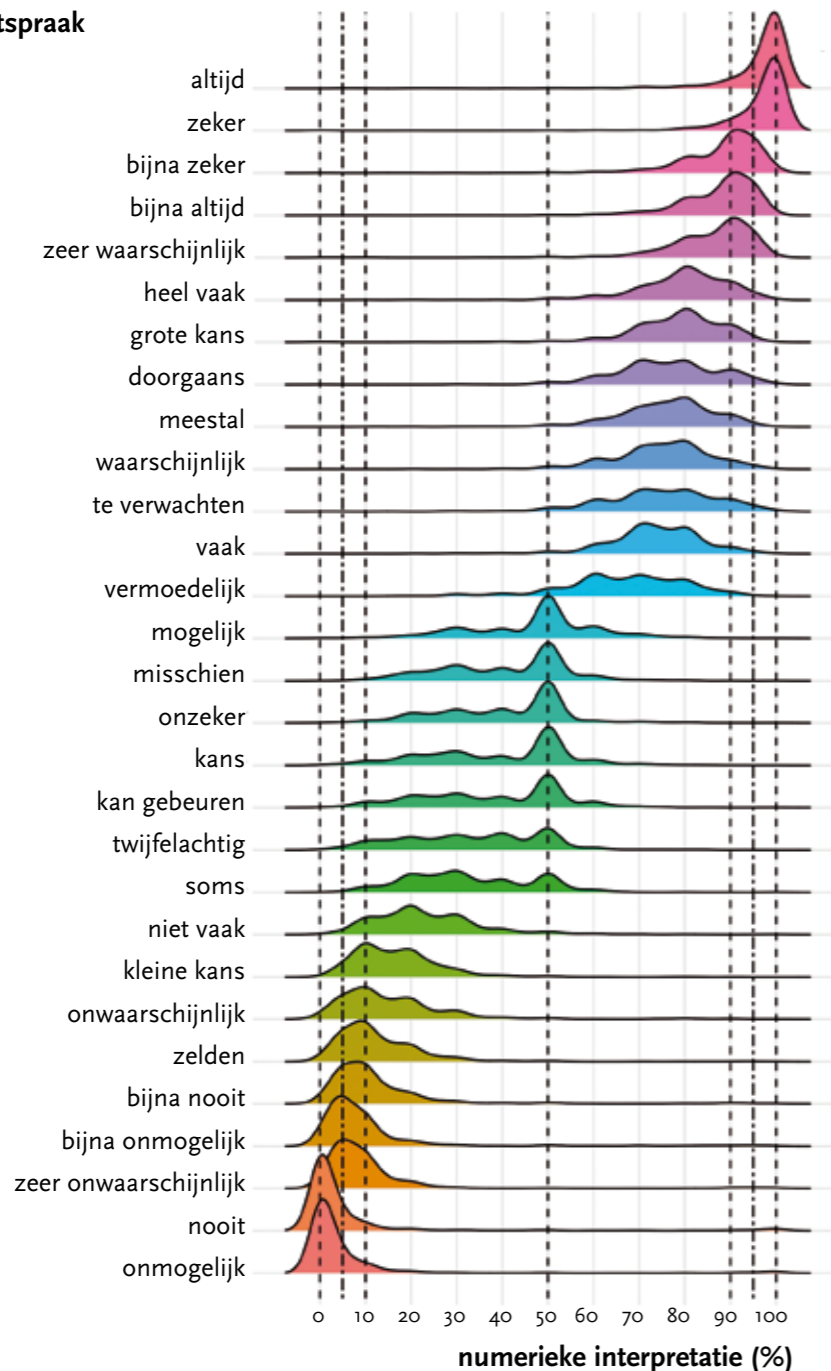
Alleen bij de extreme uitspraken, zoals altijd, zeker, nooit en onmogelijk, heeft iedereen bijna dezelfde interpretatie. De grafieken zijn minder breed en er zijn hoge pieken rond 100 en 0 procent. Dit is niet onverwacht; bij deze woorden is er weinig ruimte voor eigen interpreta-

tie. Wat me wel verbaast is dat sommige deelnemers een onmogelijke gebeurtenis toch nog een kans van 10 procent gaven. Het onmogelijke is dus toch mogelijk?

Helpt ervaring?

Hoe zit dat bij de experts? Je zou misschien verwachten dat statistici het onderling wel eens zijn over de interpretatie van verbale kansuitspraken. Daarom vroegen wij

Kansuitspraak



Figuur 2. De verdeling (dichtheidsfuncties) van de interpretatie van kans- en frequentie woorden; dit zijn een soort uitgesmeerde histogrammen, hierdoor vloeien de grafieken uit buiten de grenzen 0 en 100 procent

deelnemers of zij regelmatig (wekelijks of maandelijks) statistische analyses doen. Zo konden we de antwoorden van de twee groepen vergelijken en nagaan of ervaring met statistiek voor overeenstemming in de interpretatie zorgt.

In figuur 3 vergelijken we de resultaten van de twee groepen voor vijf kansuitspraken. De rode grafieken geven de antwoorden van de statistici weer, en die van de andere deelnemers zijn blauw. De paarse gebieden laten de overlap tussen de twee groepen zien. Er is veel overlap tussen de twee groepen, het grootste gedeelte van de grafiek is namelijk paars. En statistici zijn het onderling ook niet eens zijn over de interpretatie van de verschillende uitspraken, want ook de rode grafieken zijn erg breed.

Meer ervaring met kansen zorgt dus niet voor een eenduidige interpretatie. Misschien is dit niet zo vreemd. Als ik met mijn collega's de uitkomsten van modellen bespreek, dan doen we dat eigenlijk altijd in getallen. We gebruiken onderling dus zelden kanswoorden en bespreken daarom ook nooit de interpretatie daarvan.

Hoe dan wel?

Niet iedereen interpreteert kanswoorden op dezelfde manier. Dit kan tot verwarring leiden. Hoe lossen we dat op? Een voor de hand liggende oplossing lijkt om de verbale kansuitspraken te verbannen en alleen nog met numerieke percentages te werken. Maar helaas is het niet zo mak-

kelijk. Veel mensen vinden percentages namelijk lastig te begrijpen. Het gebruik van percentages leidt dus óók tot misverstanden.

Wetenschappers aan verschillende universiteiten doen onderzoek naar de duidelijkste weergave van kansen. Zo bekijken ze bijvoorbeeld de combinatie van woorden en kansen en het effect van verschillende weergaven van percentages, bijvoorbeeld '25 procent' versus '25 van de 100'. Ook onderzoeken ze welke factoren van invloed zijn, zoals gecijferdheid.

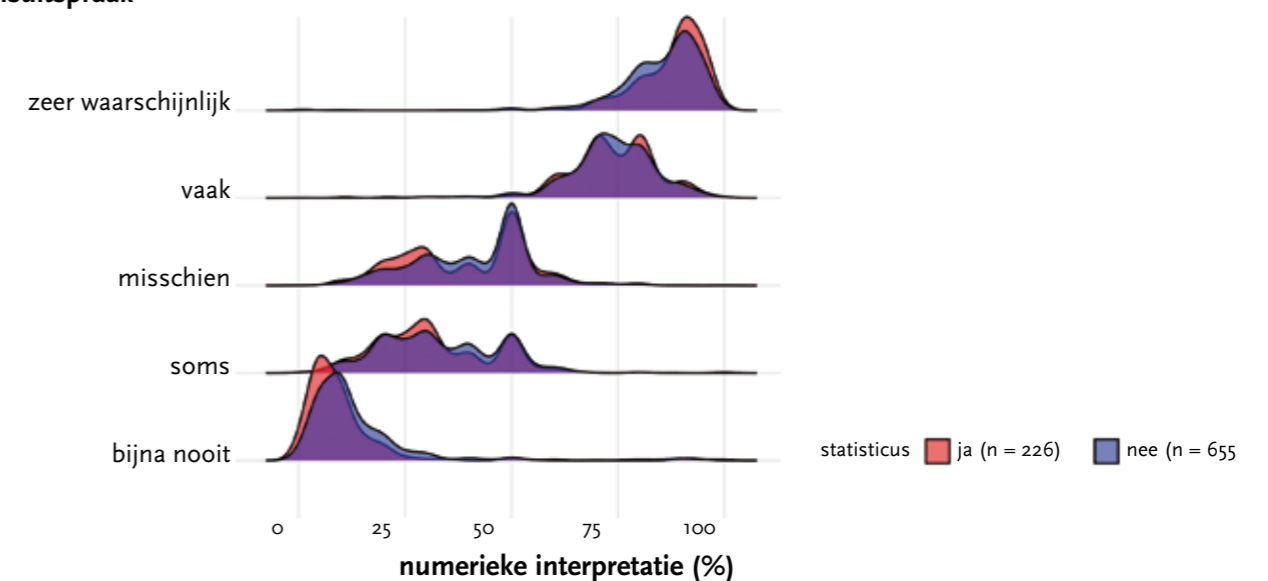
Dé oplossing is nog niet gevonden. Het is wel duidelijk dat woorden alleen niet genoeg zijn. Gebruikt jouw arts toch een keer een kanswoord? Vraag dan om verduidelijking.

VERANTWOORDING

Dit artikel is eerder verschenen op het nieuwe online blog van de VVSOR (blog.vvsor.nl). Het is gebaseerd op de resultaten van een onderzoek dat de auteur heeft uitgevoerd in samenwerking met Prof.dr. Ionica Smeets en Prof.dr. Casper Albers. De volledige resultaten zijn gepubliceerd in het wetenschappelijke artikel van S. Willems, C. Albers & I. Smeets. (2020). Variability in the interpretation of probability phrases used in Dutch news articles; A risk for miscommunication. *JCOM: Journal of Science Communication*, 19(2). https://jcom.sissa.it/archive/19/02/JCOM_1902_2020_A03

SANNE WILLEMS is universitair docent bij de sectie Methodologie en Statistiek van Universiteit Leiden. Verder is zij voorzitter van de VVSOR sectie Statistics Communication. E-mail: s.j.willems@fsw.leidenuniv.nl

Kansuitspraak



Figuur 3. De verdeling (dichtheidsfuncties) van de interpretatie van kans- en frequentie woorden; dit zijn een soort uitgesmeerde histogrammen. Let op, hierdoor vloeien de grafieken uit buiten de grenzen 0 en 100 procent