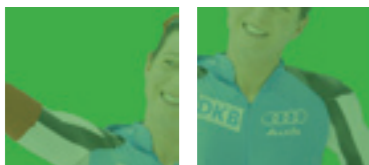


■ BANG



WAARMEE WERDEN WE DEZE MAAND BANG GEMAAKT? NWT CONTROLEERT DE MEEST ANGSTAANJAGENDE VERHALEN OP WETENSCHAPPELIJK GEHALTE.

Bloedschande

Dopingspaspoort als fabeltje

sportarbitrage (CAS) aan, nadat zij door de internationale schaatsbond (ISU) twee jaar was geschorst – wat op haar leeftijd in feite het einde van haar carrière betekent.

Pechstein is een lakmoesproef voor de bewijskracht van het zogeheten bloedpaspoort. Zo'n paspoort bevat de resultaten van alle bloedtests die vanaf het begin van de topsportcarrière zijn afgenomen. Bij de test meet men een stuk of vijf natuurlijke parameters, zoals het gehalte aan rode bloedlichaampjes.

Pechstein onderging sinds 2000 bijna honderd keer tests, waarbij nooit een verboden stof werd aangetroffen. Toch is ze nu veroordeeld, omdat één parameter verdachte schommelingen vertoonde rond het toernooi in Calgary (november 2007) en rond het wereldkampioenschap in Hamar (februari 2009). Die parameter is het percentage reticulocyten, of nieuwe rode bloedlichaampjes. Dit was abnormaal hoog vlak voor beide toernooien, en abnormaal laag een week erna. Volgens de ISU, gevolgd door het CAS in zijn vonnis van 25 november 2009, zijn dit directe aanwijzingen voor het voorafgaand gebruik van EPO, de verboden stof die de aanmaak van rode bloedlichaampjes stimuleert. EPO zelf is al na een paar dagen niet meer aantoonbaar.

In Nederland heeft chemometrisch consultant Klaas Faber zich Pechsteins zaak aangetrokken. Hij schreef een advies voor haar advocaten en publiceerde artikelen over de zaak in juridische bladen. Volgens Faber zijn de conclusies die de ISU uit Pechsteins bloedwaarden trekt een typisch geval van 'martel de data totdat die bekennen'.

Het grote strijdpunt is: welke bloedwaarden zijn nu precies 'abnormaal'? Als je een grote groep topsporters van beide geslachten, verschillende rassen en verschillende leeftijden test, is de natuurlijke variatie enorm. De variatie per individu in de loop der tijd is, zo veronderstellen de dopingautoriteiten, veel kleiner. De eerste testwaarden in het bloedpaspoort dienen daarom juist om grenzen te schatten voor de natuurlijke variatie van de sporter in kwestie. Hoe meer tests, hoe nauwkeuriger die schatting, en des te smaller de marge waarbinnen de bloedwaarden nog als normaal gelden.

Grofvuil

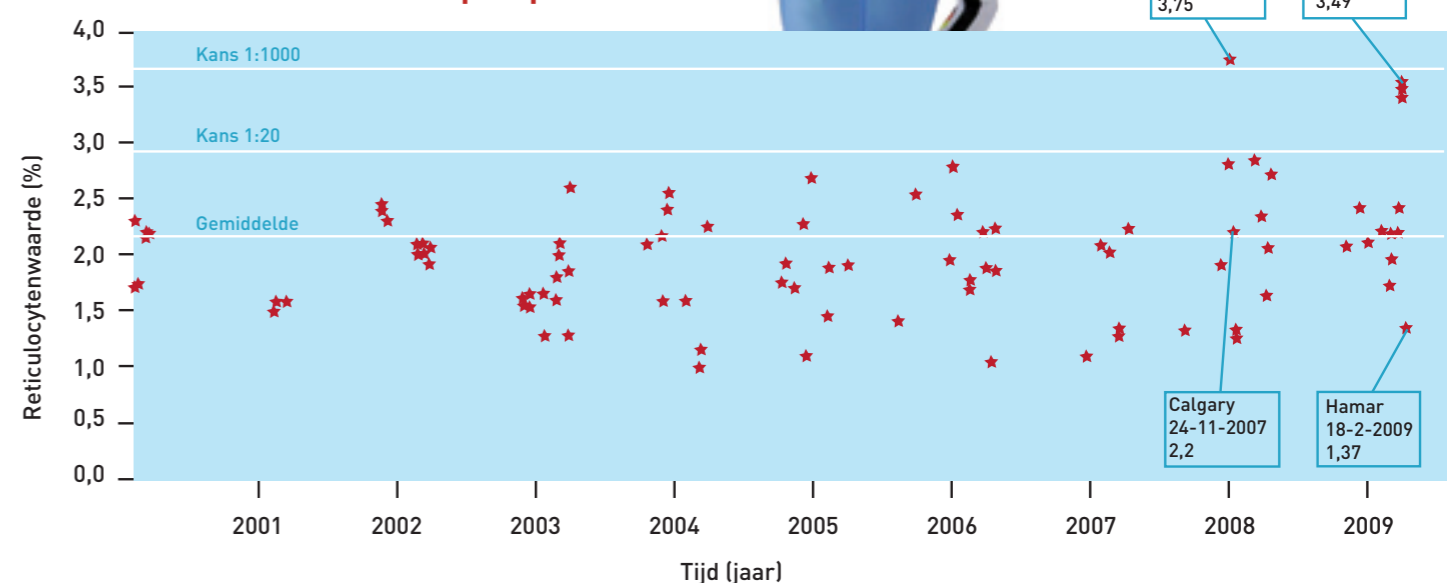
In zekere zin trekt de sporter dus de strop steeds strakker om de eigen nek. En dat is ook de bedoeling van de bedenker van het bloedpaspoort, Pierre-Edouard Sottas. Hij deed in 2007 verslag van een meerjarig experiment met tientallen dopingvrije atleten, waarvan sommigen op zeker moment een gecontroleerde hoeveelheid EPO of een placebo kregen. Sottas' conclusie: het bloedpaspoort detecteert veel meer dopinggebruikers dan andere tests, die algemene grenzen voor

abnormale bloedwaarden hanteren.

Faber heeft een scala aan bezwaren tegen Sottas' methode, die erop neerkomen dat Sottas de kans op 'vals-positieven' (onschuldigen met een abnormale testuitslag) flink onderschat. Bovendien verwijt hij Sottas een

“Het bloedpaspoort zoals het nu wordt gebruikt, moet bij het grofvuil.”

Pechsteins bloedpaspoort



de ISU dat ze zich in het geval van Pechstein niet eens aan hun eigen protocol houden. Volgens dit protocol moet de kans dat een uitschieter in de bloedwaarden op toeval berust, kleiner zijn dan 1:1000. Volgens Fabers berekening voldoet de uitschieter op 6 februari 2009 in Hamar niet aan dit criterium: deze is niet extreem genoeg (zie grafiek hierboven).

Faber is geen onafhankelijke derde, omdat hij Pechsteins advocaten van betaald advies heeft voorzien. Toch ontkent ook Faber niet dat de grote schommelingen in Pechsteins bloedwaarden rond Calgary en Hamar reden zijn voor verdenking. Zijn punt is echter dat die schommelingen op zich geen bewijs kunnen vormen voor een veroordeling. Faber: "Het bloedpaspoort is wel bruikbaar om mensen te volgen, en als daardoor verdenking rijst, kun je extra gaan testen op verboden stoffen. Maar Pechsteins veroordeling berust alleen maar op statistiek. Er zijn geen getuigen, geen extra tests, niks. Het bloedpaspoort zoals het nu wordt gebruikt, moet bij het grofvuil."

Richard Gill, hoogleraar statistiek aan de Universiteit Leiden en niet bij de zaak betrokken, komt na bestudering van het materiaal tot vergelijkbare conclusies. Gill verbaast zich er vooral over dat Sottas geen expliciete berekening geeft voor de kans op een vals-positieve uitslag. In zijn getuigeverklaring verwijst Sottas slechts naar 'normale' varia-

ties in de populatie topsporters als geheel en concludeert dan: "In de praktijk heb ik nooit zulke sterke verhogingen gezien bij een gezonde topsporter die geen EPO gebruikte."

Gill: "Sottas treedt niet op als een statisticus die de methode gebruikt die hij zelf heeft ontwikkeld en propageert. Nee, in het vonnis van de CAS spreekt de medicus die zegt: ik zie iets raars in die bloedwaarden, hier moet wel doping gebruikt zijn. Hij noemt de mogelijkheid niet eens dat de schommelingen toeval zijn."

Ook als die berekening wel was gedaan en er kwam een zeer kleine kans op toeval uit, bijvoorbeeld 1:10.000, dan is het zeer de vraag hoe significant die is. Dat betekent namelijk bepaald niet dat de kans op dopinggebruik dan 9999:10.000 is. Immers, van honderden sporters zijn in totaal tienduizenden testresultaten beschikbaar. Als een dopingjager in al die gegevens gaat zoeken naar 'iets raars', zul je gegarandeerd een paar zeer zeldzame uitschieters vinden; zo werkt toeval nu eenmaal. Als je die uitschieter eerst gebruikt om een vermeende dopingzondaar te identificeren en vervolgens gebruikt als bewijs voor dopinggebruik door diezelfde persoon, bega je de beruchte *prosecutor's fallacy*, de aanklagersdwaling.

Dogma

Gill signaleert een nog fundamenteeler probleem. Volgens hem zijn de berekeningen

allemaal gebaseerd op de zogeheten Gaussverdeling. Dit veronderstelt dat allerlei meetwaarden die een natuurlijke variatie vertonen, dat op een standaardmanier doen: bijna alle waarden in een kluitje rond het gemiddelde, terwijl uitschieters heel snel zeldzamer worden naarmate ze verder afwijken.

Volgens Gill is dit niet meer dan een 19e-eeuws dogma, waarvan bewezen is dat het met name in medische kwesties vaak niet klopt. Het probleem is dat de juiste kansverdeling in veel gevallen niet goed experimenteel is bepaald. Wel is volgens Gill duidelijk dat natuurlijke uitschieters veel minder zeldzaam zijn dan volgens de Gauss-verdeling, zodat de kansen op 'vals-positieven' nu schromelijk worden onderschat.

Het advies van Gill klinkt vertrouwd, hoewel ontmoedigend: er is meer onderzoek nodig. "Zet bijvoorbeeld een statisticus als ik en een farmacoloog bij elkaar, neem twee promovendi in dienst en onderzoek een groot aantal sporters gedurende lange tijd. Pas dan kun je betrouwbare uitspraken doen op grond van zo'n bloedpaspoort."

Gill ziet dat echter niet zo snel gebeuren, omdat volgens hem de dopingjagers totaal andere prioriteiten hebben. "Ze willen topsporters afschrikken van het gebruik van doping. Blijkbaar zijn ze bereid daarvoor een zeker percentage onschuldige sporters op te offeren. En dat zijn er meer dan ze nu denken." ●