

De mountainplot: statistiek met een grillige piek

H.K. de WOLF¹, W.H. WOERTMAN² en M. de METZ¹

Inleiding

Een methodevergelijking is één van de evaluaties die wordt uitgevoerd bij de evaluatie van een nieuwe klinisch-chemische bepaling. Regressieanalyse maakt de gemiddelde afwijking tussen twee methoden inzichtelijk en is tevens van belang om individuele en groepsgewijze afwijkingen in kaart te brengen. Immers, individuele afwijkingen rond de klinische afkapgrenzen kunnen van doorslaggevend belang zijn in de klinische evaluatie van de test. Daarnaast geldt dat structurele, groepsgewijze afwijkingen binnen de vergelijkingsstudie informatie geven over bias door interferentie, batch-to-batchverschillen of non-lineariteit. Aanwijzingen hierover zijn cruciaal voor de evaluatie van de nieuwe methode. Voor de visuele inspectie van dergelijke afwijkingen propageren de NCCLS-richtlijnen het gebruik van de bland-altmanplot (1), een verschilplot waarbij de verschillen in de vorm van een scatterplot over het meetgebied zijn weergegeven. Een andere verschilplot die binnen de klinische chemie wordt gehanteerd is de relatief onbekende mountainplot, de omgeklapte versie van de empirische cumulatieve verdelingsplot (2). Gebruik van deze plot werd destijds gepropageerd om snel een indruk te krijgen van de belangrijkste statistische parameters van een dataset (3). In de mountainplot wordt voor elk verschil x (X-as) het percentage van de waargenomen verschillen dat onder x ligt (Y-as) weergegeven, waarbij ter hoogte van de mediaan een tweede neerwaartse schaal wordt gebruikt. Dit levert een piekvormige plot op waarbij de aandacht wordt getrokken naar de piek (detectie bias en afwijkingen in de normaalverdeling) en de staart (detectie grote individuele verschillen) van de grafiek. Aan de hand van een aantal simulaties maken we de kracht en zwakte van de beide verschilplots inzichtelijk in het weergeven van afwijkingen die kunnen optreden binnen een methodevergelijking.

Methoden

Drie denkbeeldige, eenheidsloze methodevergelijkingen zijn gesimuleerd ($n=500$). In de eerste studie, die als uitgangspunt dient voor de twee andere simulaties, is een realistische dataset gesimuleerd met een hoge mate van correlatie ($r = 0,999$). De X-variabele van de eerste dataset is gesimuleerd uit een uniforme verdeling van 10 tot 500. De Y-variabele is geconstrueerd door eerst

$Y=X$ te stellen, en door vervolgens aan Y ruis toe te voegen met behulp van herschaalde, normaalverdeelde fouten, waarbij voor grotere waarden van X de SD van de fout groter was (doch relatief gezien kleiner). De tweede dataset is gemaakt door de handmatige toevoeging van vijftig meetpunten aan de eerste dataset met een groepsgewijze afwijking van +5 (onderworpen aan een normaalverdeelde fout) in het meetgebied 150-500. De derde dataset maakt weer gebruik van de X-variabele van de eerste dataset, maar de Y-variabele is geconstrueerd via $Y = 1,08 X - 20$ en door aan deze Y-waarden een normaalverdeelde fout toe te voegen (waarbij de SD over het hele meetgebied gelijk is).

Van alle datasets is een bland-altmanplot en een mountainplot gemaakt van de absolute verschillen. De mountainplot wordt gemaakt door de verschillen tussen de beide methoden naar rangorde van grootte te schikken en te converteren naar percentielen (percentiel = $\text{rang} \cdot 100 / (N+1)$). Tenslotte worden alle percentielen groter dan 50% getransformeerd volgens 'nieuwe percentiel = 100-oude percentiel'. In de beide plots zijn de 'limits of agreement' opgenomen, het interval waartussen 95% van de totale analytische fout zich bevindt (bland-altmanplot: gemiddelde afwijking $\pm 1,96$ SD, mountainplot: gebied tussen de 2,5^e en 97,5^e percentiel). De manier waarop de dataset in de beide verschilplots worden weergegeven is met elkaar vergeleken.

Resultaten

Figuur 1A en 1B geven de bland-altman- en mountainplot weer van de eerste dataset, een dataset met een hoge mate van correlatie en normaal verdeelde methodenverschillen. De bland-altmanplot (figuur 1A) van de absolute verschillen toont een scatterplot rond $Y = 0$ met een uitwaaiende vorm. Het interval van de totaal-analytische fout (-11,2 tot +11,3) is weergegeven in de figuur middels streepjeslijnen. Omgerekend betreft dit een relatieve fout van -5,0 tot +4,8% (percentage ten opzichte van de gemiddelde waarde van X en Y), die het hoogst is in het lage meetgebied en stabiliseert in het hogere meetgebied (data niet weergegeven). De mountainplot van de absolute verschillen (figuur 1B) toont een symmetrische curve met een scherpe piek bij $X = 0$ (median fout). Het interval van de absolute totaal-analytische fout [-12,1 tot +11,4] is weergegeven in de figuur en komt overeen met een relatieve fout van -5,7 tot +6,0%. De bland-altmanplot vertoont een regelmatig uitziende scatterplot zonder uitbijters. Het patroon van methodeverschillen past bij een toevallig verdeelde fout en men kan hieruit opmaken dat er geen

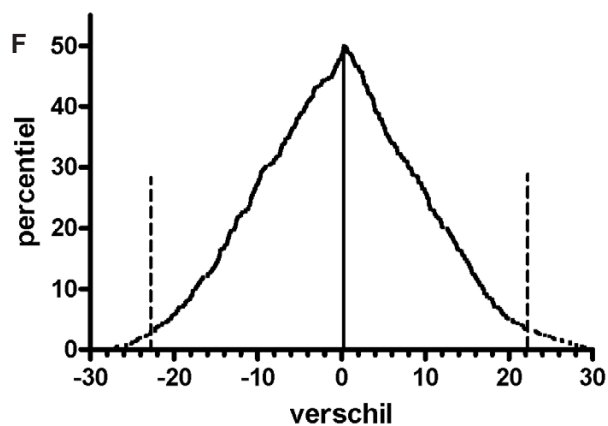
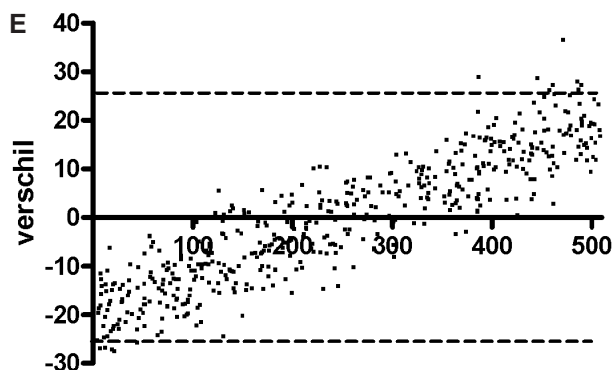
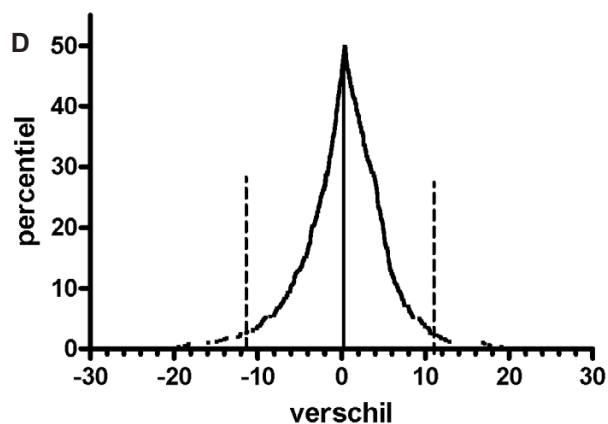
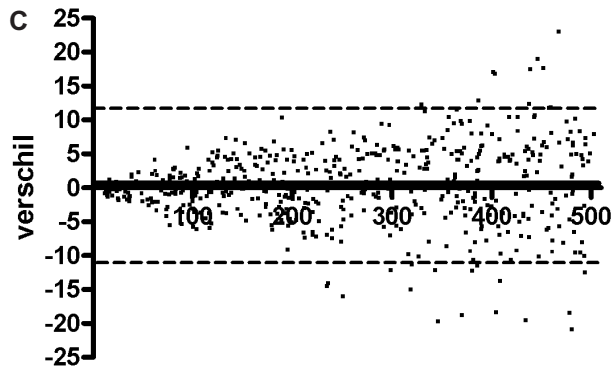
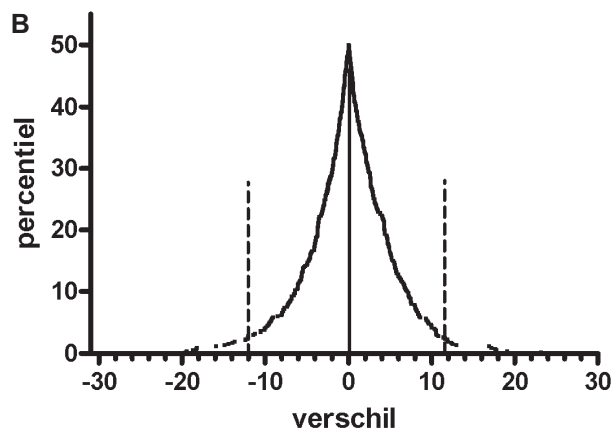
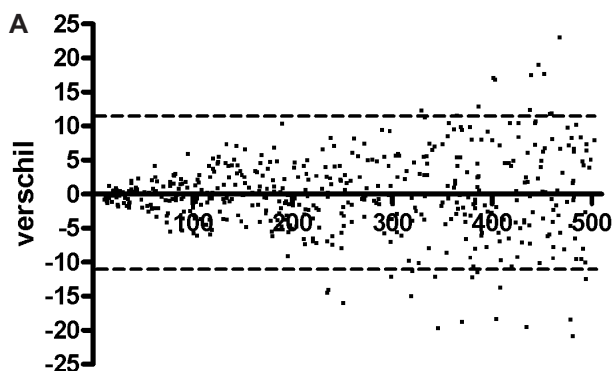
Canisius-Wilhelmina Ziekenhuis¹ en UMC St Radboud², Nijmegen

E-mail: h.d.wolf@cwz.nl

sprake lijkt te zijn van structurele fouten. Inherent aan de wijze van weergave, kan op basis van de mountainplot geen uitspraak gedaan worden over individuele afwijkingen binnen de dataset. De piek en staarten van de plot zijn symmetrisch. Hiermee wordt de, op basis van de bland-altmanplot verkregen, indruk over reguliere, random verdeelde verschillen bevestigd en worden groepsgewijze afwijkingen onwaarschijnlijk geacht.

In figuur 1C-D worden de bland-altman- en de mountainplot met elkaar vergeleken op basis van de tweede dataset die is ontstaan door aan de eerste dataset handmatig een afwijkende serie meetpunten toe te voegen. Het betreft een serie van 50 meetpunten die een gemiddelde systematische bias van +5 hebben. De bland-

altmanplot van deze aangepaste dataset (figuur 1C) vertoont een scatterplot rond $Y = +0,5$ die bijna identiek aan de plot van de eerste dataset (figuur 1A). De groep afwijkende meetpunten zijn nauwelijks zichtbaar en zonder voorinformatie kan de gebruiker hier gemakkelijk aan voorbijgaan. De mountainplot van de tweede dataset (figuur 1D) toont een scherpe piek bij $X = +0,3$. In vergelijking met de mountainplot van de eerste dataset (figuur 1B) vertoont de plot van een duidelijke toename in asymmetrie in de rechterpoot van de curve ter hoogte van $X = 4$ tot 6. Dit duidt op een niet-normaal verdeelde dataset. Op deze manier wordt de gebruiker van de mountainplot wel alert gemaakt op groepsgewijze afwijkingen binnen de betreffende dataset. De gevoeligheid waarmee de mountainplot



Figuur 1. Bland-altman-(A, C en E) en mountainplots (B, D, F) van drie gesimuleerde datasets. Ononderbroken lijn: gemiddelde/mediane afwijking. Onderbroken lijn: interval van de totaal-analytische fout.

systematische bias detecteert is het grootst bij grote afwijkingen en is onafhankelijk van de mate waarin de meetserie is uitgespreid over het meetgebied (data niet weergegeven). Voor de bland-altmanplot geldt precies de omgekeerde situatie. Door het eendimensionale karakter van de mountainplot is de afwijkende groep als zodanig niet herkenbaar. Doordat men de ligging binnen het meetgebied niet kan herleiden, kan de gebruiker de relevantie van de geobserveerde bias niet inschatten.

In figuur 1E-F worden de bland-altman- en de mountainplot met elkaar vergeleken op basis van de derde dataset die is ontstaan door de eerste dataset aan een nieuwe bias ($Y = 1,08X-20$) en een herschaalde fout te onderwerpen. De bland-altmanplot van deze dataset levert een ellipsvormig scatterplot op rond $Y=-0,1$ zonder uitbijters (figuur 1E). De oriëntatie van de puntenwolk duidt onmiskenbaar op een proportionele bias. De mountainplot van de derde dataset toont een curve met een brede doch symmetrische piek bij $X=+0,3$ (figuur 1F). De breedte van de piek en het bijbehorende grote interval van de totaal analytische fout (-23,2 tot +24,0) dient de gebruiker alert te maken op de mogelijkheid van een proportionele bias. De breedte van de piek is gecorreleerd aan de mate van proportionele bias (data niet weergegeven). Op deze manier kan de situatie ontstaan dat de aanwezigheid van een kleine proportionele bias, op basis van de symmetrische vorm van de mountainplot, onopgemerkt blijft.

Conclusie

De mountainplot is een elegante methode van weergave van de distributie van methodeverschillen. Visuele inspectie van de plot kan kleine afwijkingen ten opzichte van de normaliteit van de dataset aan het licht brengen, afwijkingen die bij gebruik van de bland-altmanplot over het hoofd kunnen worden gezien. Echter, de gevoeligheid waarmee de plot dergelijke afwijkingen visualiseert is afhankelijk van de soort en de grootte van de afwijkingen. Op die manier kunnen situaties ontstaan waarbij zelfs notoir niet-normaalverdeelde datasets onopgemerkt blijven. Bovendien blijft de gebruiker via de mountainplot in het ongewisse over de relevantie van de geobserveerde afwijkingen, iets wat door de extra dimensie van de bland-altmanplot wél zichtbaar wordt. Sterkten en zwakten van de mountainplot maken dat deze binnen een methodevergelijking complementair aan de Bland-Altmanplot dient te worden gehanteerd.

Referenties

1. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986; 307-310.
2. Krouwer JS, Monti KL. A simple, graphical method to evaluate laboratory assays. *Eur J Clin Chem Biochem* 1995; 33: 525-527.
3. Monti KL. Folded empirical distribution function curves-mountain plots. *The Amer Stat* 1995; 49: 342-346.