

Statistiek voor tandartsen (I)

A.H.B. Schuurs, tandarts^{*)}

H.J. Duivendoorn, methodoloog^{**)}

M.A. van 't Hof, statisticus^{***)}

Uit de ^{*)}vakgroep Cariologie en Endodontologie van het Academisch Centrum Tandheelkunde Amsterdam (ACTA), de ^{**)}vakgroep Medische Psychologie en Psychotherapie, Faculteit der Geneeskunde van de Erasmus Universiteit te Rotterdam en de ^{***)}Mathematisch Statistische Afdeling, Faculteit der Geneeskunde en Tandheelkunde van de Katholieke Universiteit te Nijmegen.

Trefwoord: Statistiek

Datum van acceptatie: 17 mei 1990.

Adres: Dr. A.H.B. Schuurs, Louwesweg 1, 1066 EA Amsterdam.

Samenvatting

De tandheelkundige literatuur is begrijpelijker en waardevoller als de lezer over kennis inzake de statistiek en methodologie beschikt. Een aantal basisbegrippen en eenvoudige statistische technieken worden in een serie artikelen, waarvan hier het eerste als inleiding, toegelicht. In deze bijdrage zijn enkele valkuilen aangeduid en wordt iets gezegd over 'uitval' en over systematische vertekening van de werkelijkheid (I) als een verkeerd samengestelde, dat wil zeggen een selecte groep, wordt onderzocht.

SCHUURS AHB, DUIVENVOORDEN HJ, VAN 'T HOF MA. Statistiek voor tandartsen. Ned Tijdschr Tandheelkd 1990; 97: 335-8.

1 INLEIDING

Het begrijpen van onderzoeksartikelen vereist vaak kennis van methodologie en statistiek die vele lezers van (tandheelkundige) tijdschriften vermoedelijk ontberen.¹ Omdat veel van ons handelen op statistiek is gebaseerd, is inzicht in de gebruikte statistiek gewenst, ook indien in een ontoegankelijk jargon beschreven. Daarom zal in een reeks artikelen een aantal statistische technieken uiteengezet worden. Daarbij komen ook methodologische aspecten aan bod.

2 BESCHRIJVEN EN VERKLAREN

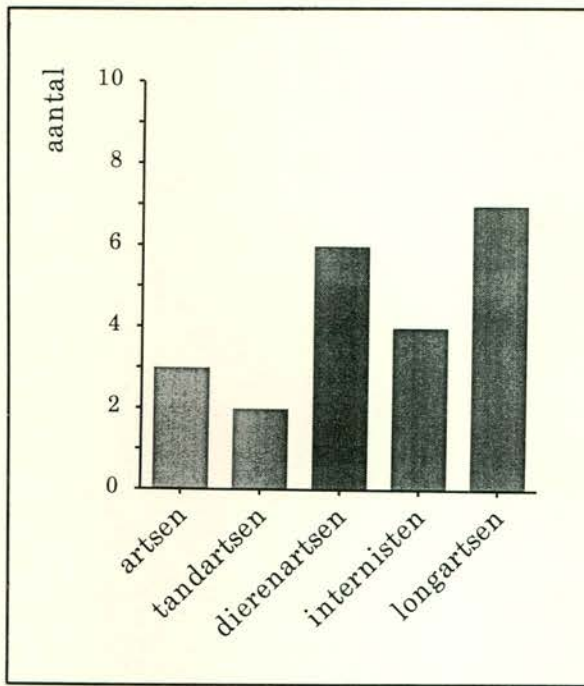
2.1 Beschrijvende statistiek

De in cijfers uitgedrukte gegevens over patiënten worden met de beschrijvende statistiek geordend en overzichtelijk samengevat.^{2,3} Daardoor ontstaat een gemakkelijk te inspecteren beeld van de onderzochte groep, onder andere over hoe de meetwaarden gespreid liggen en of er 'uitschieters' (extreme waarden) zijn. Een opsomming van alle verzamelde gegevens in een lijst (staat) geeft weinig zicht op een eventueel daarin aanwezig patroon. De resultaten van een onderzoek worden overzichtelijker weergegeven door ze bijvoor-

beeld te rangschikken naar grootte (tabel I) en gecompriemd te presenteren door hun *frequentieverdeling* op te geven: vermeld wordt dan niet dat Simon één vulling heeft, Jan, Klaas en Hugo twee en Piet en Wilfred drie vullingen, enzovoorts, maar dat bij de onderzochten één vulling eenmaal, twee vullingen driemaal en drie vullingen tweemaal voorkwamen (tabel I). Met een dergelijke weergave gaat echter enige informatie verloren: in de eerste kolom van tabel I zijn de scores van individuen wel en in de kolom met de frequentieverdeling niet meer te zien. Bij de *relatieve* frequentieverdeling zijn de absolute getallen vervangen door percentages (fracties), verkregen door het aantal individuen/waarnemingen

Tabel I. In kolom 1 en 2 staat per individu, gerangschikt naar volgorde van gebitscontrole, het aantal vullingen vermeld. In kolom 3 en 4 zijn de individuen gerangschikt naar oplopend aantal vullingen. In kolom 5 en 6 is het aantal vullingen in 'klassen' ingedeeld en wordt vermeld hoeveel individuen in elke klasse aanwezig zijn. In kolom 7 is de absolute frequentie vervangen door de relatieve (door deling met N) en in kolom 8 is de cumulatieve relatieve frequentie vermeld.

Kolom 1 Naam	2 Aantal vullingen	3 Naam	4 Aantal vullingen	5 Aantal vullingen	6 Frequentie	7 Relatieve frequentie	8 Cumulatieve relatieve frequentie
1. Jan	2	Simon	1	1	1×	0.1	0.1
2. Piet	3	Jan	2				
3. Bob	8	Klaas	2	2	3×	0.3	0.1 + 0.3 = 0.4
4. Simon	1	Hugo	2				
5. Rob	6	Piet	3				
6. Klaas	2	Wilfred	3	3	2×	0.2	0.4 + 0.2 = 0.6
7. Albert	6	Albert	6				
8. Hugo	2	Rob	6	6	2×	0.2	0.6 + 0.2 = 0.8
9. Wilfred	3	Bob	8				
10. Henk	8	Henk	8	8	2×	0.2	0.8 + 0.2 = 1.0
						—+ 1.0	



Afb. 1. Staafdiagram. Aantallen deelnemers aan een cursus per categorie.

per klasse te delen door het totale aantal (tabel I, voorlaatste kolom) Door telkens de relatieve frequentie van een klasse bij de voorgaande op te tellen ontstaat de cumulatieve relatieve frequentieverdeling (laatste kolom tabel I).

De data laten zich ook compact weer- geven met een maat voor het niveau (bij- voorbeeld het gemiddelde) en een maat voor de variabiliteit in de gegevens; hierop wordt in een vervolgartikel nader inge- gaan.

Beschrijvende statistiek is zinvol daar waar het gaat om schattingen van een be- paalde, duidelijk gedefinieerde populatie ('bevolking') te verkrijgen.

Grafische voorstellingen van de data ge- ven veelal dezelfde informatie als tabellen, maar dan gevisualiseerd. Veel gebruikt zijn staafdiagrammen (afb. 1) en frequentie- polygonen (afb. 2).

2.2 Verklarende statistiek

Naast de beschrijvende statistiek kent men de inferentiële (verklarende) statistiek, ge- bezigd om te toetsen, te schatten en te beslissen. Bij *toetsen* wordt de juistheid van een hypothese (veronderstelling) getoetst aan onderzoeksgegevens. Een hypothese kan gegrond zijn op een inval of op algeme- ne opvattingen, die men statistisch wil be- vestigen, maar kan ook een onderdeel van een nauwkeurig geformuleerde theorie be- treffen, die gebaseerd is op eerder onder- zoek of ziektebeschrijvingen. Toetsing be- helst het gebruik van statistische methoden om de juistheid van hypothesen na te gaan. Toevalsfluctuaties (Eng.: random errors) in de bevindingen bij een steekproef (een

deel van de populatie) zijn hierbij van be- lang. In een vervolgartikel wordt op 'hypo- these' en 'toetsen' dieper ingegaan.

3 STATISTIEK LIEGT?

Verkeerd of misleidend gebruik van de sta- tistiek kan schade aanrichten. Een hoofd- inspecteur van de politie te Rotterdam merkte op dat in zijn stad '.....slechts een kwart' van alle winkeldiefstallen door 'junks' werd bedreven en driekwart door 'respectabele' burgers (Volkskrant 31-1-1989). Deze opmerking suggereert dat het met de criminaliteit van verslaafden wel meevalt. Echter, in Rotterdam wonen min- stens 100 maal zoveel respectabele burgers dan 'hard drugs' gebruikers, een gegeven dat de uitspraak van de inspecteur een an- dere dimensie geeft.

Een ander voorbeeld: tegenwoordig wordt wel gezegd dat cariës een overwon- nen probleem is. Als de dmft-index voor 1000 kinderen 0,5 is, kunnen in extremo 950 kinderen cariësvrije gebitten hebben en 50 kinderen elk 20 dmf-elementen bezitten. Voor een bepaalde groep is cariës niet overwonnen.

Statistiek 'liegt' alleen door al dan niet bewust, verkeerd gebruik en door foutieve inferenties. Inferenties zijn gevolgtrekkings- afleidingen op grond van waarnemings- gen. Bewust of onbewust kan men onder- zoeksgegevens manipuleren. Tegen onbe- wuste manipulatie kan men zich wapenen. Zo is het goed dat meer dan één persoon waarnemingen doet, maar dan is het wel nodig de bevindingen van de onderzoekers met elkaar te vergelijken; ook hierop wordt in een latere bijdrage ingegaan.

4 VALKUILEN

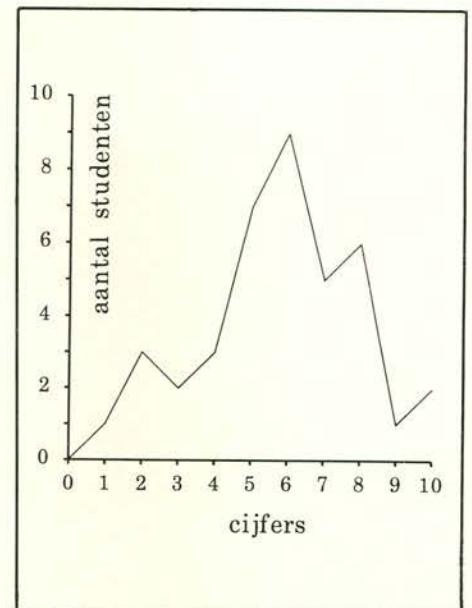
Een proces van onderzoek bergt tal van valkuilen in zich, hier geïllustreerd met een voorbeeld (4.1) en enkele opmerkingen over 'uitval' (het verdwijnen van personen tijdens onderzoek) en het gebruik van ge- gevens uit archieven (4.2 en 4.3).

4.1 Verkeerde oplossing van een vraag

Het lijkt op het eerste gezicht redelijk te verwachten dat handvaardige studenten la- ter als tandarts meer tevreden patiënten zullen hebben dan weinig handvaardige. De achterliggende gedachte is dat hand- vaardige ex-studenten betere tandheelkun- dige restauraties afleveren dan de weinig handvaardige. Om dit te onderzoeken zou men, enige jaren na afstuderen, enkele pa- tiënten van twee handvaardige en twee wei- nig handvaardige ex-studenten kunnen vra- gen naar de mate van hun tevredenheid en de kwaliteit van hun restauraties kunnen beoordelen.

Laten we aannemen dat het onderzoek geen relatie tussen handvaardigheid en te- vredenheid oplevert. Het zal de lezer bij nader inzien duidelijk zijn dat de gekozen werkwijze niet zo vanzelfsprekend is. En- kele bezwaren tegen dit onderzoek zijn:

1. Tevredenheid wordt wellicht meer door de wijze van omgaan met de patiënten bepaald dan door de kwaliteit van de vullingen.
 2. Het aantal onderzochte praktijken en patiënten is wel erg klein. Over het aan- tal dat onderzocht had moeten worden, handelt een volgend artikel.
- Een voldoende grote én geschikte steek- proef is nodig om te voorkomen dat er



Afb. 2. De data van tabel 1 in een polygoon uitgebeeld.

sprake zal zijn van systematische vertekening (Eng.: bias) in de gegevens (Eng.: data).

4.2 'Bias'

Napijn na een endodontische behandeling kan spontaan verdwijnen (natuurlijk beloop), maar door de patiënt/tandarts worden toegeschreven aan voorgeschreven medicatie. Dit geldt ook voor een groep patiënten. Om de effectiviteit van een pijnstillertje te onderzoeken dient men over een tweede groep patiënten, die geen medicatie krijgt, te beschikken. Om bevooroordeelde waarnemingen te voorkomen is de waarnemer niet op de hoogte van de behandelingswijze. Dit alles geschiedt om bias te voorkomen.

Bias in de onderzoeksgegevens kan in enkele hoofdgroepen worden ingedeeld; zij zijn het gevolg van selectie, informatie- en confounding bias. Dit alles komt neer op verschillende vormen van onvergelykbaarheid.⁴ Ter illustratie worden enkele voorbeelden gegeven.

4.2.1 Uitval

Tot een onderzoeksverslag behoren gegevens over het onderzochte 'materiaal', het aantal proefdieren, de onderzoekspersonen, enzovoorts. Tijdens het onderzoek kunnen proefdieren sterven, mede als gevolg van het experiment. Bij enquêtes kunnen mensen vanwege een veelheid van redenen niet meedoen, door ziekte, overlijden, verhuizingen, wegeringen. Het aantal uitvallers kan hoog oplopen, met name bij een postaal onderzoek en een na-onderzoek.

Door uitval kan het beeld, verkregen in het onderzoek, ernstig van de werkelijkheid afwijken. Wat bij voorbeeld te denken van de conclusie van een fictief onderzoek dat 75% (tabel II) van volledige prothesen na tien jaar nog voldoet? Als we naar de data in tabel II kijken, lijken ten minste vijf verschillende conclusies mogelijk.

1. In het *slechtste* geval moeten *alle niet te beoordelen* prothesen als 'mislukt' worden beschouwd, met onder andere als argumentatie dat het niet-reageren op de oproep door ontevredenheid over de als slecht beoordeelde prothese werd veroorzaakt.

Dus geslaagd is
 $(300/1000) * 100 = 30\%$.

2. In een *gunstiger* geval mag – gebaseerd op het aantal van 300 'voldoende' prothesen op de 400 prothesen beoordeelde – ook 3/4 van de niet-beoordeelde prothesen als 'geslaagd' worden geklasseerd. Dan blijft het totale percentage 'geslaagd' dus 75%.

3. In het *gunstigste* geval kan men alle

Tabel II. Beoordeling van prothesen na tien jaar.

	n	%
Oorspronkelijke aantal geplaatste prothesen	1000	100
Aantal na 10 jaar beoordeeld	400	40
Aantal na 10 jaar als 'goed' beoordeeld	300	75*
<i>Redenen voor niet-beoordeling</i>		
Patiënt kan niet komen (ziekte, e.d.)	70	7
Niet beoordeelbaar wegens overlijden	40	4
Geen reactie op controle-oproep (verhuizingen, ontevreden over prothese?)	490	49

* Hier hadden ook andere percentages kunnen staan – zie tekst.

uitval, weinig realistisch, 'geslaagd' noemen.

Dus:

geslaagd is $\frac{(330 + 530)}{1000} * 100 = 83\%$.

De conclusie dat 75% nog voldoet, is slechts een van de mogelijke. Hoewel in de tabel correct de redenen van de uitval worden vermeld, was een zuiverder beeld ontstaan door de vermelding dat het percentage van geslaagde prothesen tussen 30% en 83% ligt. Door de laatste twee percentages te middelen wordt $(30+83) = 60,5\%$ gevonden.

In een vervolgonderzoek wordt vijf procent uitval door sommigen als nog aanvaardbaar beschouwd. Anderen noemen een uitval van 10% tijdens het onderzoek naar klinisch beloop van een ziekte zorgelijk en bij een uitval van 20% zouden de resultaten zelfs niet meer de moeite van het lezen waard zijn.⁵

4.2.2 Retrospectief onderzoek en 'Berkson fallacy'

In een retrospectief onderzoek werkt men met in het verleden verzamelde gegevens. Archiefgegevens werden vaak met een ander doel dan het doen van onderzoek bewaard. Een nadeel van het gebruik van dergelijke gegevens kan zijn dat sommige voor het onderzoek noodzakelijke data niet beschikbaar zijn (informatie-'bias').

Een tweede nadeel is dat de patiënten van ziekenhuizen en faculteit tandheelkunde een zekere 'bias' kunnen tonen, bijvoorbeeld doordat daar patiënten komen op verwijzing (selectie-bias), met als gevolg dat de populatie van verwezen patiënten

kan verschillen van de populatie van niet-verwezen patiënten. Er is sprake van onvergelykbaarheid door selectie. Ook door het toelatingsbeleid ('admissibility bias' en 'admission bias' = Berkson fallacy) en verwijzingsbeleid (complicaties) kunnen onjuiste conclusies worden getrokken. Dit geldt temeer voor conclusies gebaseerd op *verhoudingen van aantallen* patiënten met *combinaties* van aandoeningen. Berkson fallacy genoemd.^{4 6 7} Zo zou in een orthodontische praktijk een relatie kunnen worden gevonden tussen agenesie en gebits-traumata, domweg omdat de kans in de orthodontische praktijk te worden ingeschreven groter is als beide afwijkingen aanwezig zijn. Echter, onder de bevolking zal die relatie afwezig zijn.

4.2.3 Confounding bias

Bij de vaststelling van de levensduur van kroon- en brugwerk speelt niet alleen de vaardigheid van de tandarts, maar ook het feit of hij al dan niet bereid is welhaast 'hopeloze' gebitten te behandelen een rol. Alleen al hierdoor kan een in wezen kundige tandarts bij vergelijking met werk van collegae worden gediskwalificeerd. Maar eigenlijk mogen niet alle tandartsen in dezen als een homogene groep worden behandeld. De gebitsituatie van de patiënten is hier namelijk een versturende (confounding) factor. Dit probleem is op te lossen in experimenteel onderzoek (clinical trial).

5 SLOT

Dit alles betekent niet dat alle onderzoek in klinieken en ziekenhuizen vertekeningen

SUMMARY

STATISTICS FOR DENTISTS

Key word: Statistics

Knowledge of statistics and methodology enhances the understanding of the dental literature. Most dentists are, however, unfamiliar with these subjects. Therefore, in a series of articles a number of statistical methods will be described. The present article deals with some pitfalls. Attention is given to, among others, non-response due to drop outs and several forms of bias.

LITERATUUR

- ¹SCHEUTZ F, ANDERSEN B, WULFF HR. What do dentists know about statistics? Scand J Dent Res 1988; 96: 281-7.
- ²WILLEMSSEN EW. Understanding statistical reasoning. San Francisco: W.H. Freeman, 1974; ch. 1.
- ³ARMITAGE P, BERRY G. Statistical methods in medical research. Oxford: Blackwell Scientific Publications, 1987; ch. 1.
- ⁴SACKET DL. Bias in analytic research. J Chron Dis 1979; 32: 51-63.
- ⁵TIJSSEN JGP. Lezen tussen de regels van gerapporteerd geneesmiddelenonderzoek. Geneesmiddelenbulletin 1987; 21: 73-83.
- ⁶TUGWELL PX. Hoe moeten medische tijdschriften worden gelezen? III. Het beoordelen van het klinische beloop en de prognose van een ziekte. Ned Tijdschr Geneesk 1983; 127: 2367-71.
- ⁷DE JONGE H, RUMKE CL, VAN STRIK R. Medische statistiek. Afd. Medische statistiek, Vrije Universiteit te Amsterdam.

geeft, wel dat de onderzoeker bedacht moet zijn op de vraag over welke populatie hij een uitspraak wil doen. Men moet bedacht zijn op een goede opzet van het onderzoek, op de bepaling van een adequate omvang van de te onderzoeken groep, op systematische fouten indien het patiënten betreft die verwezen zijn of als men met bijvoorbeeld archiefkaarten werkt. In de tandheelkunde faculteiten zijn 'clinical trials', het evalueren van klinisch werk en het bezien of bepaalde afwijkingen gerelateerd zijn (zolang het geen verwijzingsproblemen betreft) zeer goed uit te voeren.

Bias is vaak een methodologisch oplosbaar probleem. De statistiek speurt naar deze systematische fouten, om tot 'unbiased' schattingen te komen.

Uit de historie

De eerste orthodontische patiënt

Na een voordracht door Ch.F.L. Nord, gehouden voor de 'Vereeniging van Nederlandse Tandartsen', mei 1927, verzucht één der aanwezigen tijdens de discussie:

'De allereerste orthodontische patiënt die ik heb behandeld heeft mij doen schrikken, zoals in mijn latere 18-jaren praktijk nooit een geval mij heeft doen schrikken. Nog langen tijd daarna kreeg ik elken keer, wanneer ik de deur voorbijging van het huis waar de patiënt woont, hartkloppingen. In 1908 kreeg ik n.l. in behandeling een meisje van een jaar of 10, 11 met een heel nauw gehemelte. Ik heb toen een plaat gemaakt

precies op de wijze zoals de heer Nord gedemonstreerd heeft. Ik had haar gezegd dat zij er niet aan mocht komen, maar, het kind was heel ijverig, zij heeft gezien hoe ik de schroef een slag omdraaide en toen zij thuis kwam heeft zij dat ook eens gedaan en meermalen gedaan. Op een avond, ongeveer een uur of elf, werd ik opgebeld of ik onmiddellijk komen wilde, want het kind had een geweldig dik gezicht. Ik ging er met een taxi heen en inderdaad, het kind had zoo'n gezicht. Ik betastte de bovenkaak en het bleek mij, dat beide helften bewegen konden. Nooit ben ik zoo geschrokken als toen. Ik heb het ding onmiddellijk uitgeno-

men. Het is terechtgekomen zonder malheur, ik heb twee of drie jaar geleden die patiënt nog gezien. Ik heb verscheidene Misserfolge er mede gehad en ik ben persoonlijk buitengewoon huiverig om deze methode toe te passen. Dat neemt niet weg dat ik het nog wel eens weer zal probeeren, maar, of het nu inderdaad iets is wat wij in de gewone praktijk kunnen toepassen, zoals de heer Nord zeide, ohne Gefahr und ohne Risiko, dat weet ik nog niet. Daarvoor zou ik meer ervaring moeten hebben.'

Bron: Tijdschr Tandheelkd 1928; 35: 147.