



# Een liniaal langs littekens in de literatuur

Bram Hengeveld \*

Iedereen kan zich voorstellen dat twee timmerlieden die met twee verschillende centimeters meten, samen scheve huizen zullen bouwen. Daarom hebben we duidelijke afspraken over wat een centimeter is. De rolmaat en de duimstok van timmerlieden zijn gemaakt op basis van die afspraken en leveren in diverse omstandigheden dezelfde resultaten: het meetinstrument is betrouwbaar en valide. Maar waar een lengte tegenwoordig relatief eenvoudig te meten is, stelt een litteken ons voor meer complexe vragen. En die vragen worden op verschillende manieren beantwoord. In de medische wereld zijn daarom allerhande meetinstrumenten te vinden voor littekens. Want hoe meten we eigenlijk een litteken? En verschillen patiënten en medici van mening over wat daarin van belang is? Carrière et al. zochten het uit in een interessante (en open access beschikbare) systematische literatuurstudie (1).

## De review

Het doel van de studie was het maken van een overzicht van gebruikte meetinstrumenten en gemeten karakteristieken van littekenkwaliteit in gepubliceerde onderzoeken. Deze karakteristieken betroffen visuele (bijv. kleur), tactiele (bijv. soepelheid) en sensorische (bijv. pijn) eigenschappen. Bij het maken van het overzicht werd er onderscheid gemaakt tussen vier groepen instrumenten: schalen die alleen door patiënten of medici worden ingevuld, of door medici en patiënten. Tevens werd gezocht naar apparaten (measurement devices) die een aspect van littekenkwaliteit meten.

Het doen van een meta-analyse van de klinimetrische eigenschappen van de diverse schalen en instrumenten was geen onderdeel van de review. Er wordt dus geen antwoord gegeven op de vraag hoe goed de gevonden instrumenten presteren. De resultaten van de review vormen de basis van een Delphi-onderzoek om te bepalen hoe littekenkwaliteit gedefinieerd moet worden. In een Delphi-onderzoek wordt op een methodische manier gevraagd naar kennis en meningen van experts met als doel consensus te bereiken over een onderwerp. In dit geval een internationale groep littekenexperts en mensen met littekens.

De databases van PubMed en EMBASE werden op 31 oktober 2018 doorzocht naar Engelstalige publicaties waarin tenminste één eigenschap van littekens werd gemeten. Hierbij was het van belang dat er ook beschreven werd hoe die eigenschap werd gemeten. Publicaties die dit niet deden, werden uitgesloten. Dierstudies werden ook uitgesloten.

In totaal werden 440 complete artikelen geïncludeerd. Daarin werden in totaal 909 instrumenten gebruikt: 375 keer een clinician reported, 236 patient reported, 27 patient

## Systematische review en meta-analyse

De begrippen systematische review en meta-analyse worden soms door elkaar gebruikt. Het zijn echter twee verschillende begrippen. In een systematische review wordt op een duidelijk beschreven en grondige manier gezocht naar publicaties over een specifieke onderzoeksvraag. Hierbij wordt zoveel mogelijk alle relevante en beschikbare literatuur doorzocht. Meta-analyse is een statistische techniek die vervolgens op de gevonden data kan worden toegepast. Meta-analyse stelt ons in staat om diverse eigenschappen van onderzoeken (zoals de onderzoeksgrootte) mee te wegen in een getalsmatige samenvatting van alle beschikbare data. Wie meer wil weten over de methodiek van systematische review en meta-analyse kan zich wenden tot het gratis Cochrane Handbook of Systematic Reviews <https://training.cochrane.org/handbook/current>.

& clinician reported en 271 keer een apparaat. Er werden 31 verschillende (versies van) clinician reported schalen gevonden. De POSAS2 (<https://www.posas.org/>), de Vancouver Scar Scale (VSS) en de aangepaste VSS door Baryza & Baryza werden hierbij het vaakst gebruikt. Bij de patiëntschalen werden 39 verschillende (versies van) meetschalen gevonden. De patiëntversie van de POSAS en de Burn Specific Health Scale-brief werden het vaakst gebruikt. De twee aanpassingen van de VSS door Nedelec en Oliveira waren de meest toegepaste van in totaal negen verschillende gecombineerde schalen. Tenslotte werden er maar liefst vijftig verschillende apparaten gebruikt in de gevonden literatuur. De Cutometer werd met 32 keer het vaakst gebruikt, gevolgd door de Deraspectometer met 25 maal.

In tabel 1 staat de rangschikking van de vijf meest gebruikte eigenschappen van littekenkwaliteit van de vier soorten instrumenten.

Carrière et al. merken op dat er een aanzienlijke variatie bestaat in de wijze waarop omgegaan wordt met het meten van littekenkwaliteit; van consensus is geen sprake. Verreweg de meeste onderzoeken gebruikten de meetinstrumenten om patiënten klinisch te evalueren (88%). In 31 onderzoeken werd een instrument zelf getest op betrouwbaarheid en/of validiteit. Van de onderzoeken betrof de grootste groep (40%) patiënten met chirurgische littekens, gevolgd door littekens van brandwonden (36%). Tien procent betrof onderzoek onder patiënten met keloid littekens.

### Bespreking

De review van Carrière et al. is keurig volgens de PRISMA-richtlijnen (2) beschreven. Mogelijke belangenconflicten en de financiering door de Nederlandse Brandwondenstichting worden duidelijk vermeld. Enigszins opvallend is het feit dat van de hits in Pubmed en EMBASE in totaal 28% uiteindelijk ook bruikbaar was voor inclusie. Dat is, in vergelijking met veel andere systematische reviews, behoorlijk hoog. De keuze voor de twee databases en het weglaten van bijvoorbeeld de Cochrane CENTRAL database van de Cochrane wound en Skin groups wordt niet onderbouwd. Pubmed en EMBASE zijn weliswaar twee van de belangrijkste databases, maar het kan zijn dat er publicaties (en instrumenten) zijn gemist. Een voorzichtig antwoord op die vraag kan gevonden worden als we kijken naar vergelijkbare reviews zoals die door Deng (3), Lee (4), Nguyen (5), Junker (6), Bae (7), Jones (8), Verhaegen (9), Fearmonti (10) en Durani (11). De review van Carrière et al. is (veelal) recenter en geeft een meer omvattend beeld dan deze reviews, waarvan sommige zich op één soort eigenschap of soort instrument richten of (zeer) summier hun methodieken beschrijven. De Matching Assessment of Scars and Photographs (MAPS) (12), die door Nguyen et al. (5) wordt genoemd, is echter niet terug te vinden in de hier besproken review. De MAPS is gebaseerd op de Seattle Scar Scale, waarvan een aantal aanpassingen zijn opgenomen in de review van Carrière et al. Er lijkt echter geen lijst te zijn van alle geïncludeerde publicaties, dus het is helaas

### Betrouwbaarheid en validiteit

Het nut van een meetinstrument is van een aantal zaken afhankelijk. Zo moet het natuurlijk makkelijk toepasbaar zijn (denk aan de rolbandmaat die in je broekzak past), maar we moeten er ook betrouwbaar mee kunnen meten en het instrument moet valide zijn. Deze twee termen zijn paraplubegrippen die globaal als volgt uit te leggen zijn: betrouwbaarheid geeft aan hoe goed we met een instrument onderscheid kunnen maken tussen individuen (of bijvoorbeeld littekens) op basis van specifieke kenmerken. Betrouwbaarheid is niet zozeer een kenmerk van een instrument, maar van de meting. De betrouwbaarheid moet zoveel mogelijk onafhankelijk zijn van zaken als tijd, omstandigheden of de persoon die de test uitvoert. Maar een betrouwbare meting is nog steeds niet valide als deze niet ook daadwerkelijk meet wat we willen meten. Met een rolbandmaat kan je misschien betrouwbaar de lengte en de breedte van een litteken bepalen, maar het is geen valide instrument om iets te zeggen over littekenkwaliteit. Dat begrip omvat veel meer facetten. Sommige daarvan zijn minder direct te meten dan de lengte en de breedte. We noemen verzamelbegrippen als littekenkwaliteit een construct. De validiteit van een meetinstrument voor littekenkwaliteit geeft dus aan of we ook daadwerkelijk het construct meten. Het bepalen van betrouwbaarheid en validiteit van (multi-item) meetschalen is een wetenschap op zich, maar onmisbaar voor een goed meetinstrument. Voor degenen die hier meer over willen weten, wordt het boek van Streiner en Norman (14) van harte aanbevolen.

niet te verifiëren of dit de MAPS betreft. De MAPS beschrijft tevens de locatie van het litteken als karakteristiek; deze eigenschap is niet terug te vinden in de supplementen bij de publicatie van Carrière et al. Navraag bij de auteurs leert dat de publicatie van Masters (12) - en daarmee de MAPS - inderdaad niet terug is gekomen in de hits van hun zoekstrategie. De complete lijst met referenties van hun geïncludeerde artikelen is nogmaals door de auteurs bij het Journal Plastic and Reconstructive Surgery - Global Open aangeboden.

**Tabel 1. Rangorde van meest gebruikte karakteristieken. Vertaald uit Carrière et al (1)**

Type instrument	Eerste	Tweede	Derde	Vierde	Vijfde
Clinici	Dikte	Pigmentatie	Vasculariteit	Plooibaarheid	Onregelmatigheid van oppervlak
Patiënten	Pijn	Jeuk	Kleur	Dikte	Plooibaarheid
Combinatie	Vasculariteit	Dikte	Jeuk	Plooibaarheid	Pijn
Apparaten	Vasculariteit	Plooibaarheid	Dikte	Pigmentatie	Vochtgehalte

Ook opvallend is de relatief lage frequentie waarin functionele beperkingen ten gevolge van littekens zijn gemeten. Deze karakteristiek komt alleen voor in de patient reported scale en komt daar in totaal slechts zes keer voor, ondanks dat dit meermaals wordt gezien als een essentieel onderdeel van een goed meetinstrument in de hierboven genoemde publicaties. Vanzelfsprekend is dit geen tekortkoming in de review van Carrière et al.; het geeft vooral het belang van hun werk aan.

Een bespreking van de klinimetrische eigenschappen van de tests, waar deze voorhanden zou zijn, is wel de grote afwezigheid in de review van Carrière et al. Als we afgaan op deze review valt echter ook de relatief kleine hoeveelheid onderzoeken naar ontwikkeling en klinimetrische eigenschappen van de instrumenten om littekens te beoordelen op. Deze is met respectievelijk 23 en 31 publicaties veel kleiner dan de hoeveelheid gebruikte (ad hoc aangepaste versies van) instrumenten. Ook dit geeft de urgentie aan van het werk van de auteurs. Het gebruik van niet goed onderzochte meetinstrumenten kan namelijk flinke gevolgen hebben. Marshall et al. (13) vonden dat trials over schizofrenie waarin niet-gepubliceerde meetinstrumenten werden gebruikt 37% vaker een positieve uitkomst rapporteerden. Veel publicaties zouden zelfs een omgekeerde conclusie (nl. géén superioriteit van behandeling) hebben als een gepubliceerde meetschaal zou zijn gebruikt.

*Het is duidelijk dat consensus op het vlak van het construct littekenkwaliteit nog toekomstmuziek is.*

## Conclusie

Ondanks enkele kritische kanttekeningen is - wat betreft omvang, 'scope' en de wijze waarop methode en bevindingen zijn gerapporteerd - een buiging voor de auteurs op z'n plek. Het is duidelijk dat consensus op het vlak van het construct littekenkwaliteit nog toekomstmuziek is. Hopelijk leidt de Delphi-studie tot vruchtbare grond voor de ontwikkeling van een valide meetinstrument dat voor alle gebruikers, maar zeker ook voor patiënten aanvaardbaar is. Het opstellen en testen daarvan is beslist geen sinecure en de wetenschap erachter een relatief jong gebied binnen de medische wereld (14). Carrière et al. doen de zeer terechte opmerking dat een dergelijk meetinstrument de evidence base voor littekenbehandeling sterk zal verbeteren, maar ook zeker een positieve invloed kan hebben op research waste op dit gebied (15).

Tenslotte rest me nog een woord van dank te richten aan mevrouw Carrière voor haar reactie op dit artikel en haar antwoord op de vragen.

## Literatuur

1. Carrière ME, Kwa KAA, de Haas LEM, et al. **Systematic Review on the Content of Outcome Measurement Instruments on Scar Quality Plastic and Reconstructive Surgery** - Global Open 1 2019.
2. Liberati A, Altman DG, Tetzlaff J, et al. **The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration.** PLoS medicine, 2009;e1000100+.
3. Deng H, Li-Tsang CWP. **Measurement of vascularity in the scar: A systematic review.** Burns, 2019;45:1253-1265.
4. Lee KC, Dretzke J, Grover L, et al. **A systematic review of objective burn scar measurements.** Burns Trauma, 2016;(14) 4.
5. Nguyen TA, Feldstein SI, Shumaker PR, et al. **Review of scar assessment scales.** Semin Cutan Med Surg, 2015;34:28-36.
6. Junker JPE, Philip J, Kiwanuka E, et al. **Assessing quality of healing in skin: review of available methods and devices.** Wound Repair Regen, 2014;22Suppl1:2-10.
7. Bae SH, Bae YC. **Analysis of frequency of use of different scar assessment scales based on the scar condition and treatment method.** Arch Plast Surg, 2014;41:111-15.
8. Jones HG. **Clinimetrics of tristimulus colourimeters in scar assessment: a review of evidence.** J Wound Care, 2012;21:30-35.
9. Verhaegen PDHM, van der Wal MBA., Middelkoop E, et al. **Objective scar assessment tools: a clinimetric appraisal.** Plast Reconstr Surg, 2011;127:1561-1570.
10. Fearmonti R, Bond J, Erdmann D, et al. **A review of scar scales and scar measuring devices.** Eplasty, 2010;10:e43.
11. Durani P, McGrouther DA, Ferguson MWJ. **Current scales for assessing human scarring: a review.** J Plast Reconstr Aesthet Surg, 2009;62:713-720.
12. Masters M, McMahon M & Svens B. **Reliability testing of a new scar assessment tool, Matching Assessment of Scars and Photographs (MAPS).** J Burn Care Rehabil, 2005;26:273-284.
13. Marshall M, Lockwood A, Bradley C, et al. **Unpublished rating scales: a major source of bias in randomised controlled trials of treatments for schizophrenia.** Br J Psychiatry, 2000;176:249-252.
14. Streiner DL, Norman GR. **Health measurement scales: a practical guide to their development and use.** Oxford: Oxford University Press; 2008.
15. Chalmers I, Glasziou PP. **Avoidable waste in the production and reporting of research evidence.** Lancet, 2009;374:86-9.

\* *Bram Hengeveld, wijkverpleegkundige bij Livio (Enschede), protocolschrijver bij Vilans (Utrecht) en zal in maart 2020 zijn MSc in Evidence Based Practice (klinische epidemiologie) van de UvA/AUMC ontvangen. Lid WCS commissie Ulcus cruris en lid van de redactie van WCS Nieuws.*