

De klinische toepasbaarheid van de Dysphonia Severity Index (DSI)

Marieke Hakkesteegt

Afdeling KNO - Gehoor- en Spraakcentrum, Erasmus MC, Rotterdam

Samenvatting

Voor het evalueren van behandeling van stemstoornissen zijn onder andere objectieve meetinstrumenten nodig. De Dysphonia Severity Index (DSI) is een dergelijke objectieve maat. De benodigde parameters zijn de hoogste haalbare frequentie, de minimale intensiteit, de maximale fonatieduur en het percentage jitter. Het scorebereik ligt ongeveer tussen de -5 en +5.

De DSI heeft een relatie met de perceptuele beoordeling van stemkwaliteit; hoe slechter de stemkwaliteit hoe lager de score op de DSI. Er blijkt geen significant verschil tussen mannen en vrouwen op de DSI. Wel is er een effect van leeftijd: met het ouder worden, wordt de DSI van mensen zonder stemklachten lager. Dit effect bestaat voor zowel mannen als vrouwen.

Bij herhaalde metingen van de DSI blijkt dat er een verschil van minimaal 2,5 moet zijn om van een significante verandering te kunnen spreken. De uitkomst van de meting is niet afhankelijk van de onderzoeker die de meting uitvoert.

De DSI kan worden gebruikt bij het evalueren van interventie, zoals stemverbeterende chirurgie of stemtherapie, bij groepen patiënten. Op individueel niveau lijken kleine veranderingen minder goed te evalueren, omdat deze vaak binnen de bandbreedte van de normale variatie vallen.

Summary

For evaluation of effects of intervention for voice disorders objective tools are needed. The Dysphonia Severity Index (DSI) is such an objective tool. The parameters of the DSI are the highest possible frequency, the lowest intensity, the maximum phonation time and the percentage jitter. The scores on the DSI are approximately in a range of -5 till +5. The DSI is related to the perceptual evaluation of voice quality; the worse the voice quality, the lower the score on the DSI. There are no significant differences between males and females. In a population without voice disorders, age appeared to have a significant influence on the DSI: with advancing age, the DSI is getting lower, in both males and females. In repeated measures of the DSI (without intervention in between), the difference between measures has to be at least 2.5 to be a significant change. The results of measures

are not dependent of the investigator.

The DSI is applicable for evaluation of intervention for voice disorders, such as phonosurgery or voice therapy in groups of patients. On an individual level it seems that small changes are not really well detected with the DSI, since those changes are often within the normal variations.

Inleiding

In alle (para)medische beroepsgroepen is tegenwoordig grote belangstelling voor 'evidence based' handelen. Zo ook in de logopedie in het algemeen en, in het kader van dit artikel, voor stemstoornissen in het bijzonder. Om het effect van een bepaalde behandeling te kunnen beoordelen, is het noodzakelijk over uitkomstmaten te beschikken waarmee dit resultaat gemeten kan worden. Dergelijke uitkomstmaten moeten aan diverse eisen voldoen.

Allereerst is het van belang te weten of je meet wat je denkt te meten. Bij het beoordelen van stemkwaliteit gaat het om wat we kunnen horen aan de stem, de perceptuele beoordeling. Een objectieve maat voor stemkwaliteit zou dus gerelateerd moeten zijn aan deze perceptuele beoordeling. Ook moet bekend zijn wat de meetfout of spreiding in de meting is. Bij veel metingen is het zo dat een herhaalde meting niet exact dezelfde uitkomst heeft als een (kort) daarvoor uitgevoerde meting. Zeker als er, zoals bij metingen van de stem, een mondelinge instructie wordt gegeven door de onderzoeker, en er van de proefpersoon een prestatie wordt verwacht. Het weten van deze spreiding is belangrijk om bij het evalueren van bijvoorbeeld stemtherapie te kunnen beoordelen of een *verschil* tussen de meting voor en na de behandeling ook daadwerkelijk op een *verandering* duidt.

De stem verandert in de loop van het leven, maar er zijn natuurlijk ook verschillen tussen de gemiddelde mannen- en vrouwenstem. Bovendien zijn er grote verschillen tussen stemmen: we kunnen een persoon vaak herkennen door alleen naar de stem te luisteren. Om te kunnen beoordelen wat je aan een objectieve meting hebt, is het van belang te weten hoe de uitkomsten zijn bij proefpersonen zonder stemklachten, van diverse leeftijden en zowel van mannen als vrouwen. Verder moet de meting in staat zijn een verandering te detecteren: dus niet alleen grote verschillen in stemkwaliteit tussen twee fragmenten van dezelfde stem moeten kunnen worden gemeten, maar liefst ook de wat kleinere verschillen.

Er bestaat (wereldwijd) geen consensus over de te gebruiken metingen voor stemkwaliteit. Wel is bekend dat multi-parametrische maten meer geschikt zijn dan een enkelvoudige. Van de diverse maten die worden gebruikt, is nog maar weinig betreffende bovenstaande kenmerken bekend. Daarom heeft ons onderzoek zich gericht op een al bestaande, multi-parametrische maat, namelijk de door Wuyts et al. ontwikkelde Dysphonia Severity Index (DSI) (Wuyts et al., 2000). De DSI is een index, met de volgende parameters: de hoogst haalbare frequentie (F_{0-high}), de minimale intensiteit (I_{low}), de maximale fonatieduur (MPT) en jitter. De DSI wordt als volgt berekend: $DSI = 0,13 \times MPT + 0,0053 \times F_{0-high} - 0,26 \times I_{low} - 1,18 \times \text{jitter} (\%) + 12,4$. De DSI is zodanig geconstrueerd dan een stem die op de GRBAS schaal van Hirano (Hirano, 1981) wordt beoordeeld als G0 (normale stem) overeenkomt met

een DSI van +5, terwijl een ernstig dysfone stem (G3) overeenkomt met een DSI van -5. Maar ook scores hoger dan +5 en lager dan -5 zijn in principe mogelijk.

Het voordeel van de DSI is dat hij in de dagelijkse praktijk tamelijk snel en eenvoudig is te meten. De DSI werd door ons nader onder de loep genomen. Onder andere werd onderzocht:

- De relatie tussen de perceptuele beoordeling en de DSI.
- Het effect van geslacht en leeftijd op de DSI.
- De inter-beoordelaarsvariabiliteit en test-hertest betrouwbaarheid van de DSI.
- De toepasbaarheid van DSI (en VHI) bij het meten van effect van interventie bij stemstoornissen.

In de volgende paragrafen zullen de opzet en resultaten van deze onderzoeken kort worden besproken.

De relatie tussen de perceptuele beoordeling en de DSI

Het uitgangspunt bij de ontwikkeling van de DSI was dat het een objectieve meting van stemkwaliteit moest zijn. Om te onderzoeken of een nieuwe meetmethode meet wat je wilt meten, zou deze moeten worden vergeleken met de 'gouden standaard'. Echter, een echte gouden standaard om stemkwaliteit te meten bestaat niet. Meestal wordt de perceptuele beoordeling als zodanig beschouwd (Kreiman, Gerratt, Kempster, Erman, & Berke, 1993; Ma & Yiu, 2006), ondanks de matige betrouwbaarheid hiervan (De Bodt, Wuyts, Van de Heyning, & Croux, 1997; Dejonckere et al., 1998; Kreiman & Gerratt, 1998; Webb et al., 2004). Bij het ontwikkelen van de DSI is uitgegaan van de perceptuele beoordeling, gescoord op de G van de GRBAS schaal van Hirano (Hirano, 1981).

Bij een groep patiënten werd onderzocht hoe de relatie tussen de DSI en de score op de G is. Daarnaast hebben we een groep personen zonder stemklachten, de controlegroep, vergeleken met de groep patiënten met stemklachten, om te onderzoeken in hoeverre de DSI in staat is onderscheid te maken tussen mensen met en zonder stemproblemen.

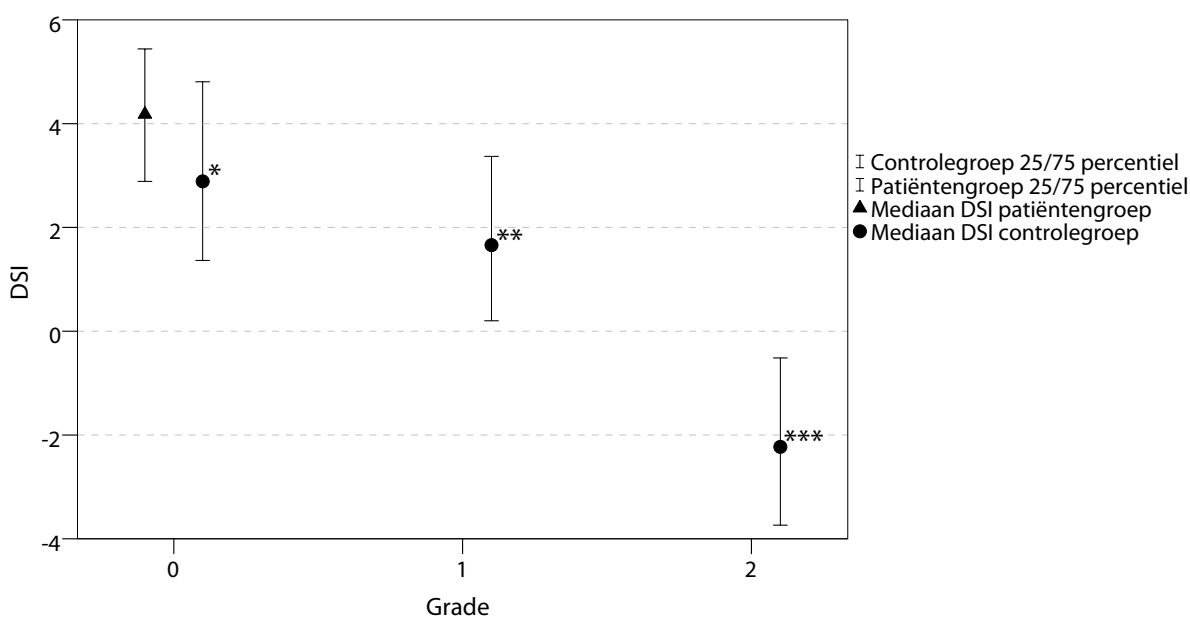
De proefpersonen

Patiënten: N= 294 (196 vrouwen, 98 mannen, leeftijd 14-87 jaar, gemiddeld 44). Dit waren zowel patiënten met benigne organische als niet-organische stemstoornissen. Controlegroep: N=118 (69 vrouwen, 49 mannen, leeftijd 20-79, gemiddeld 44).

Resultaten

De resultaten zijn elders uitgebreid beschreven (Hakkesteeft, Brocaar, Wieringa, & Feenstra, 2008). Hier volgt een beknopte versie. In figuur 1 is zichtbaar dat de personen in de controlegroep een significant betere DSI hebben dan de patiënten die een perceptuele beoordeling

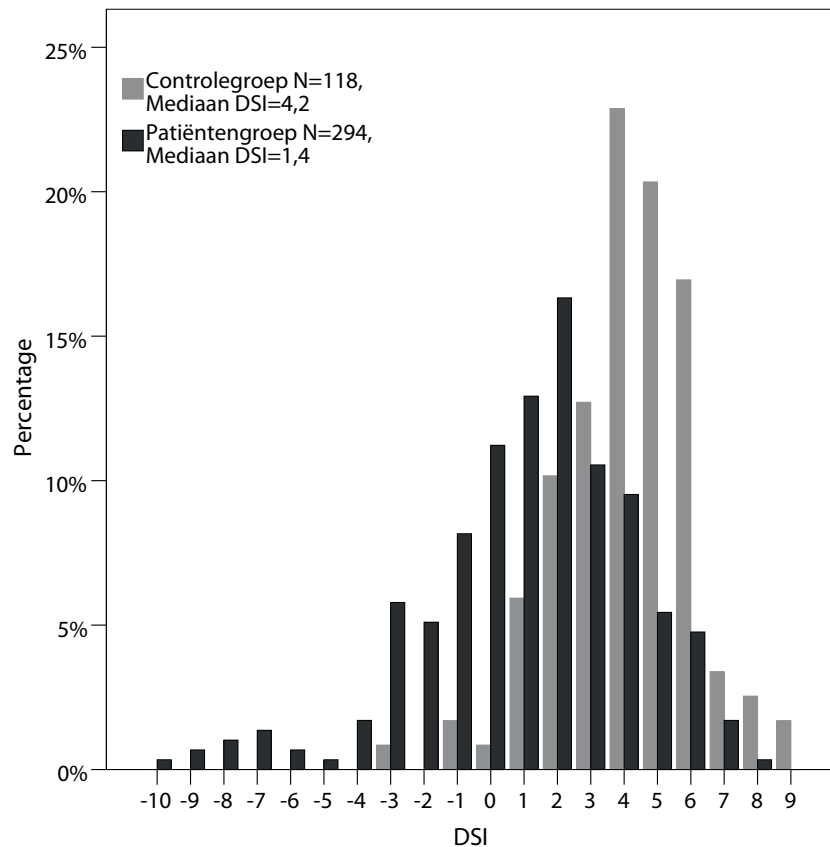
van G0 kregen. Ook zijn de verschillen tussen de overige perceptueel beoordeelde ernstcategorieën significant. In figuur 2 zijn de resultaten van de DSI weergegeven voor zowel de controles als de patiënten. Er is sprake van een overlap tussen beide groepen. Om te onderzoeken of de DSI patiënten en mensen zonder stemklachten van elkaar kan onderscheiden, werd bij verschillende afkappunten zowel de sensitiviteit (correct positieven) als de specificiteit (correct negatieven) bepaald. Deze blijken bij een DSI van 3,0 optimaal. Dus een DSI van minder dan 3 duidt op een stemstoornis.



Figuur 1: Mediaan van de DSI met 25ste en 75ste percentiel per G-score voor de patiëntengroep en de controlegroep. * = Verschil met controlegroep is significant ($p < 0,05$); ** = Verschil met patiënten met G0 is significant ($p < 0,05$); *** = Verschil met patiënten met G0 and G1 is significant ($p < 0,05$).

Bespreking

Patiënten die met een klacht van de stem komen, hebben niet altijd een in de spreekstem hoorbare afwijking. Soms komen mensen met klachten die alleen na (forse) stembelasting optreden, of juist bij hoge eisen die aan de stem worden gesteld (zingen, luid spreken). Dit verklaart dat er ook in de patiëntengroep wel G0 scores zijn. De scores van deze patiënten op de DSI zijn echter significant lager dan die van de controlegroep. De DSI meet niet alleen de kwaliteit van de stem, maar ook juist de functie van de stem, wat bij de groep met niet hoorbare afwijkingen in de spreekstem blijkbaar een geschikte wijze van evalueren is. Het blijkt dat hoe groter de hoorbare afwijking van een stem is (hogere score op G), hoe slechter de DSI is. Met de objectieve DSI wordt de ernst van de afwijking van de stem dus overeenkomend



Figuur 2: Verdeling van de DSI in de patiëntengroep en de controlegroep.

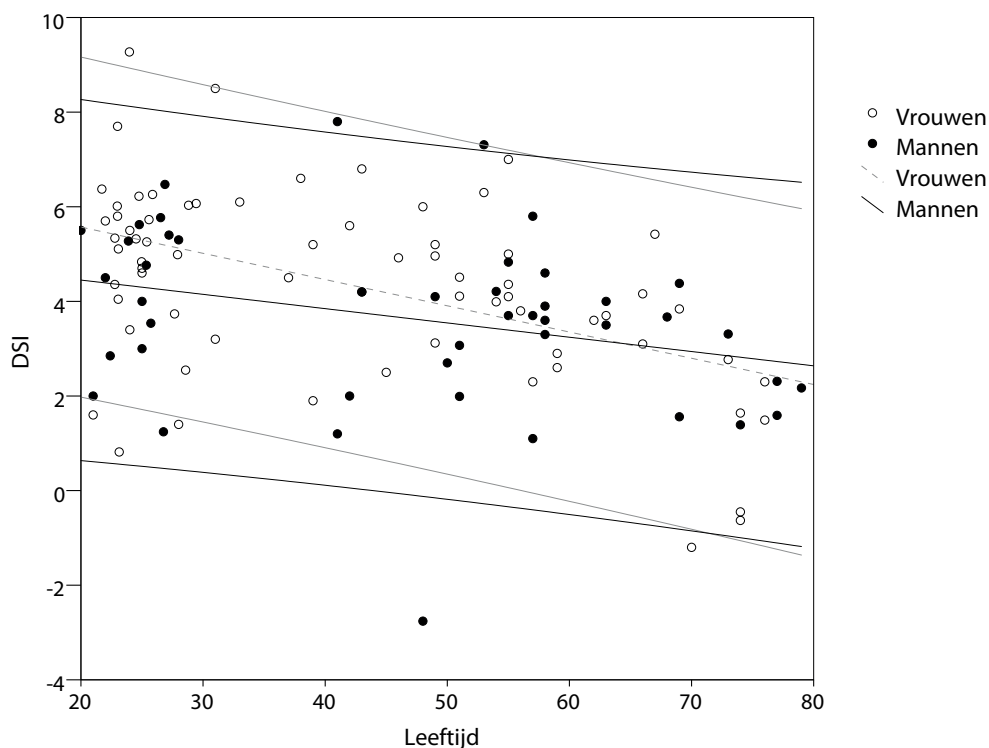
met de subjectieve beoordeling gemeten. Toch zijn er ook patiënten die een DSI hebben die zou duiden op een goede stem ($DSI > 3$). Mogelijk gaat dit om mensen die vooral klachten hebben met betrekking tot de (duur)belasting van de stem.

Het effect van geslacht en leeftijd op de DSI

Het is bekend dat de stem verandert met het ouder worden (Honjo & Isshiki, 1980; Ramig et al., 2001; Sataloff, Rosen, Hawkshaw, & Spiegel, 1997; Teles-Magalhaes, Pegoraro-Krook, & Pegoraro, 2000). Bovendien zijn zeker twee van de parameters van de DSI waarschijnlijk duidelijk verschillend bij mannen en vrouwen, namelijk de hoogste frequentie en de maximale fonatieduur. Dit werd onderzocht bij de hierboven beschreven controle groep van 118 personen zonder stemklachten.

Resultaten

Het verschil tussen de geslachten was op zowel de maximale fonatieduur (MPT) als de hoogste frequentie (F_{0-high}) significant, maar in de uiteindelijke DSI bleek geen significant verschil tussen mannen en vrouwen. De effecten van deze twee parameters heffen elkaar blijkbaar op. Dit komt overeen met de resultaten die Wuyts et al. (Wuyts et al., 2000) hebben gevonden. Wel was er een significant effect van leeftijd op de DSI (figuur 3): met het toenemen van de leeftijd wordt de DSI lager. Dit effect van de leeftijd is voor zowel mannen als vrouwen en is voornamelijk te verklaren uit de hoogste frequentie, die lager wordt bij het toenemen van de leeftijd. Het effect van leeftijd en geslacht op alle afzonderlijke parameters van de DSI is elders beschreven (Hakkesteeft, Brocaar, Wieringa, & Feenstra, 2006).



Figuur 3: Effect van leeftijd op de DSI.

Bespreking

Er blijkt inderdaad een effect van leeftijd op de stem te zijn. Dit betreft het effect bij mensen zonder stemklachten. Dit betekent in de praktijk dat het goed is bij het interpreteren van de DSI de leeftijd in aanmerking te nemen en onderscheid te maken tussen fysiologische en pathologische stemverandering.

De inter-beoordelaarsvariabiliteit en test-hertest betrouwbaarheid van de DSI

Om met behulp van een meting te bepalen of een bepaalde interventie resultaat heeft gehad, is het van belang te weten hoe groot de normale variatie is. Deze kan afhangen van normale variaties in de stemkwaliteit (bijvoorbeeld als gevolg van emoties), normale variaties in het presteren bij de opdrachten, of van de onderzoeker die de metingen verricht (wijze van instructie geven, mate van aanmoediging). De test-hertest betrouwbaarheid van de DSI werd bepaald door bij 30 personen zonder stemklachten op 3 verschillende momenten door 2 verschillende onderzoekers de DSI te meten. Tussen de meetmomenten zat gemiddeld 9 dagen. Het onderzoek werd elders uitgebreid beschreven (Hakkestegt, Wieringa, Brocaar, Mulder, & Feenstra, 2008).

Resultaat

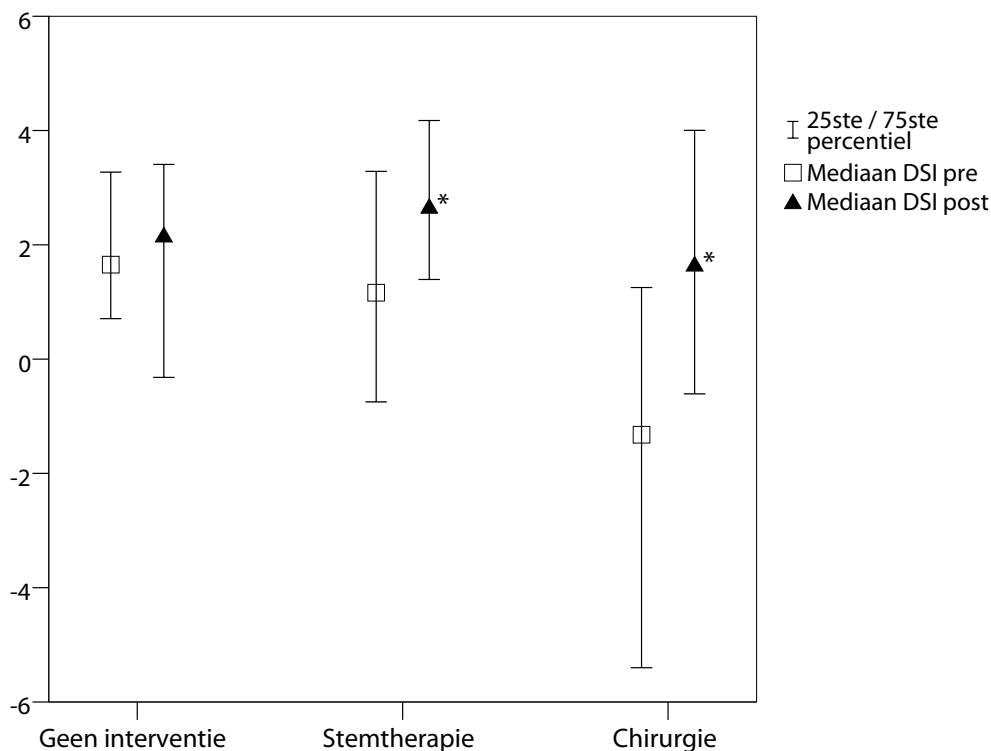
De invloed van de onderzoeker blijkt niet significant. De standaard deviatie van het verschil tussen de metingen was 1,27. Dit betekent dat een verschil van $> 2,5$ ($1,96 \times 1,27$) tussen 2 metingen van een persoon (bijvoorbeeld voor en na therapie) significant is.

Bespreking

De 'bandbreedte' van de variatie in de DSI blijkt vrij groot te zijn, want er moet een verschil van minimaal 2,5 op de DSI zijn om een verschil als significante verandering te kunnen beschouwen.

De toepasbaarheid van DSI en VHI bij het meten van effect van interventie bij stemstoornissen

Naast de DSI wordt in onze kliniek voor het evalueren van de mening van de patiënt gebruik gemaakt van de Voice Handicap Index (VHI) (Jacobson et al., 1997). Om te onderzoeken of de DSI en de VHI zijn te gebruiken voor het meten van effect van interventie (bijvoorbeeld logopedische of chirurgische behandeling), werden de resultaten van de metingen op twee verschillende momenten geanalyseerd van 171 patiënten. Hiervan waren er 45 niet behandeld in de tussenliggende periode, 34 ondergingen stemverbeterende chirurgie in verband met goedaardige afwijkingen en 92 werden logopedisch behandeld. Deze drie groepen werden gekozen om de verschillen in resultaten die we tussen deze groepen verwachtten. De verwachting was dat bij de onbehandelde groep weinig verschil zou zijn tussen de twee metingen en dat het verschil voor de chirurgisch behandelde groep groter zou zijn dan voor de logopedisch behandelde groep.



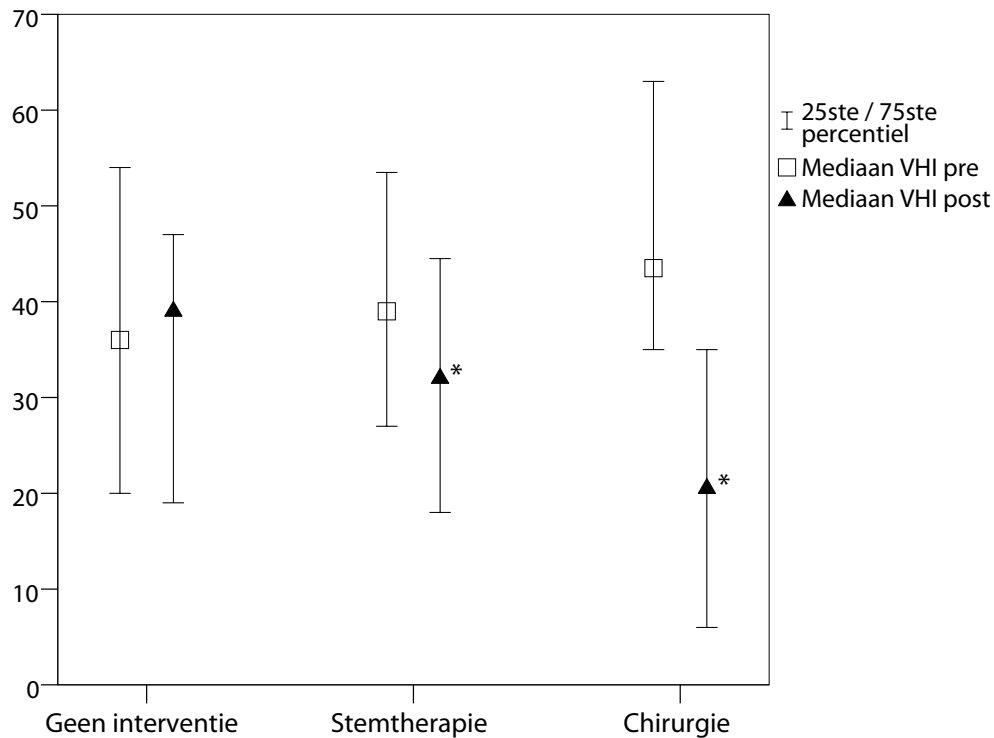
Figuur 4: Resultaten voor de groepen patiënten voor en na interventie, met mediane scores en 25ste en 75ste percentiel voor de DSI. Een hogere DSI correspondeert met een betere stemkwaliteit. * Verschil tussen pre and post is significant ($p < 0,05$).

Resultaten

De resultaten werden uitgebreid elders beschreven (Hakkesteeft, Brocaar, & Wieringa, 2010), hier volgt een beknopte weergave. De resultaten voor de drie groepen patiënten voor de DSI en de VHI zijn weergegeven in respectievelijk figuur 4 en 5. Bij de groep zonder interventie is er geen significant verschil tussen de twee meetmomenten, bij de twee groepen met interventie wel, waarbij het verschil bij de groep na chirurgie groter is dan bij de groep die logopedisch werd behandeld. Tevens werd naar de gecombineerde resultaten van DSI en VHI van de individuele patiënten gekeken (figuur 6).

Bespreking

De DSI blijkt bruikbaar om voor groepen patiënten te onderzoeken of een bepaalde interventie effect had, ook als dit effect mogelijk niet heel groot is op individueel niveau (zoals vaak bij logopedische stemtherapie het geval is). Op individueel niveau moet het verschil minimaal 2,5 op de DSI zijn om dit als verandering te beschouwen. Dit is vooral voor een deel van de groep die stemtherapie kreeg niet het geval. Alle patiënten vulden ook een VHI in. Voor een groot deel van de patiënten komen de resultaten op de DSI en de VHI overeen (allebei beter

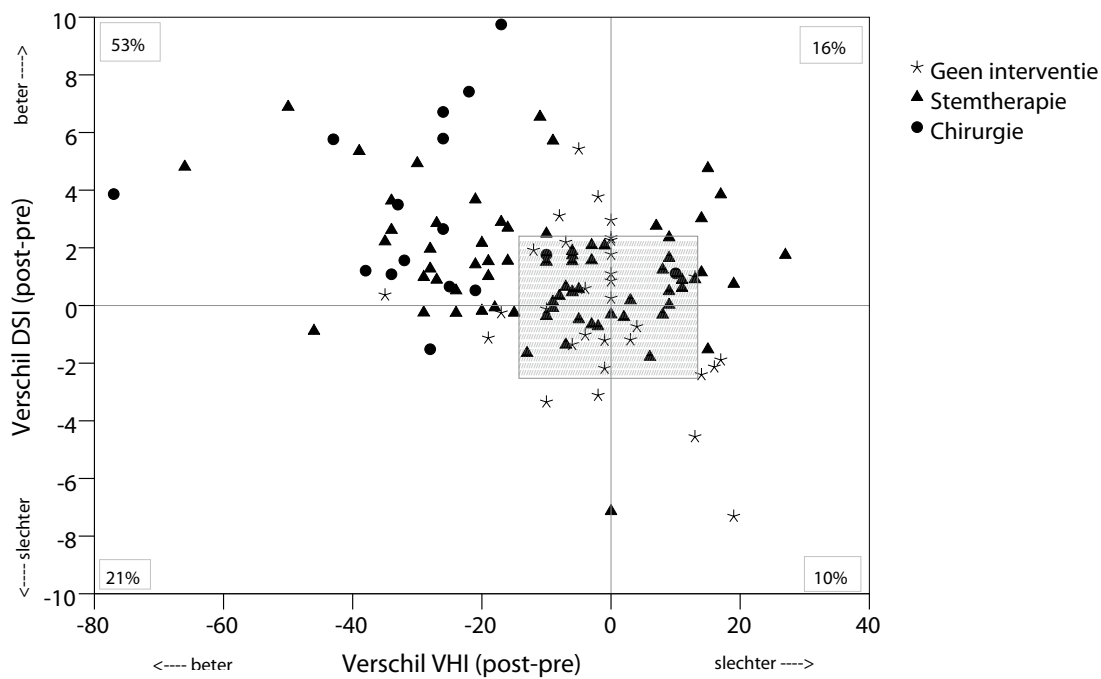


Figuur 5: Resultaten voor de groepen patiënten voor en na interventie, met mediane scores en 25ste en 75ste percentiel voor de VHI. Een lagere VHI correspondeert met minder stemklachten. * Verschil tussen pre and post is significant ($p < 0,05$).

of allebei slechter). Maar er zijn ook mensen bij wie de resultaten op DSI en VHI elkaar lijken tegen te spreken, vooral in de groep die stemtherapie kreeg. Dit kan te maken hebben met de aard van logopedische stemtherapie: er is vaak sprake van een geleidelijke verbetering van de stem, en mensen zijn vaak intensief met de stem bezig. Hierdoor zijn ze zich misschien juist wel méér bewust van de beperkingen die ze ondervinden, terwijl ze een verbetering van de stem nauwelijks opmerken. Het is ook mogelijk dat de stemkwaliteit al wel is verbeterd, maar dat dit in de dagelijkse praktijk nog niet goed kan worden toegepast, waardoor de klachten nog wel aanwezig zijn. Anderzijds zijn er mensen bij wie de stemkwaliteit niet beter wordt, maar die op de VHI minder klachten aangeven. Misschien zijn dit wel de mensen die bij de stemtherapie hebben geleerd beter om te gaan met hun beperkingen.

Algemene bespreking

Stemproblemen zijn vaak complex. De ernst van het probleem wordt niet alleen bepaald door hoe de stem klinkt, maar ook (en misschien wel vooral) door de eisen die iemand aan zijn of haar stem stelt. Voor een zanger is een geringe afwijking van de stem al een groot probleem, terwijl eenzelfde afwijking door een niet-professioneel stemgebruiker misschien



Figuur 6: De relatie tussen het verschil in DSI en VHI voor en na interventie voor de 3 interventiegroepen. De punten die zich in het grijze vlak bevinden, representeren niet-significante verschillen voor zowel DSI als VHI. De percentages van de totale groep per quadrant staan in de hoeken vermeld. Een hogere DSI correspondeert met een betere stemkwaliteit ('better'), een lagere DSI dus met een slechtere stemkwaliteit ('worse'). Een lagere VHI correspondeert met minder stemklachten ('better'), een hogere VHI dus met meer klachten ('worse').

niet eens zou worden opgemerkt. Voor de onderzoeker is de stemkwaliteit hetgeen we direct waarnemen en dus kunnen beoordelen en willen objectiveren. Het is bij het evalueren van stemproblemen in ieder geval van groot belang de verschillende aspecten te evalueren. Natuurlijk voor het stellen van een diagnose, maar ook voor het evalueren van een eventuele interventie (bijvoorbeeld logopedie, chirurgie).

Om te onderzoeken wat het effect van interventie is, kun je groepen patiënten vergelijken. Dan wordt het gemiddelde van de groep van bijvoorbeeld de DSI vergeleken met het gemiddelde van dezelfde groep na interventie. Dan weet je wat gemiddeld het effect van die interventie is. Maar het is ook belangrijk te weten hoe dat op individueel niveau ligt: hoeveel patiënten zijn verbeterd en hoeveel verbetering was er, zijn er misschien patiënten niet verbeterd of zelfs verslechterd. Dan wordt mogelijk ook inzichtelijk bij welke patiënten een bepaalde interventie zinvol zal zijn, en bij welke niet, of minder. Omdat we van de DSI weten wat het verschil op individueel niveau moet zijn om te mogen spreken van een verandering, zou de DSI als objectieve maat hiervoor kunnen worden toegepast. Het voordeel van de DSI is daarbij dat deze vrij snel en eenvoudig te meten is en onafhankelijk is van de onderzoeker die de meting uitvoert. Een nadeel in de dagelijkse praktijk blijkt echter dat ernstig dysfone stem-

men niet goed te meten te zijn. Dit geldt zowel voor de meting van de hoogste frequentie als voor de minimale intensiteit, maar in nog veel sterkere mate voor de jitter. Het is bekend dat de jitter geen geschikte parameter is voor ernstig dysfone stemmen (Ma & Yiu, 2005). Het is daarom het overwegen waard om de jitter te begrenzen tot een maximale waarde van bijvoorbeeld 5 of 10%: alle metingen groter dan die waarde duiden op een ernstig dysfone stem. Voor de DSI betekent dat er dan iets meer begrenzing van de mogelijke uitkomsten is. Het verschil tussen een DSI van -10 of -20 is namelijk klinisch irrelevant. Een ander punt van aandacht bij het vergelijken van DSI's van verschillende instellingen is de apparatuur die is gebruikt voor het bepalen van de jitter, omdat de uitkomst sterk daarvan afhankelijk is (Maryn, Corthals, De Bodt, Van Cauwenberge, & Deliyski, 2009).

Conclusie

In de dagelijkse praktijk blijft de perceptuele beoordeling van een stem een belangrijk instrument, maar deze beoordeling kent vanwege een matige betrouwbaarheid vooral beperkingen in de toepassing voor wetenschappelijk onderzoek. De DSI is een objectieve en kwantitatieve maat voor stemkwaliteit, geschikt voor het evalueren van de resultaten van interventie (zoals stemtherapie). Logopedische evaluatie van stemstoornissen dient multidimensioneel te zijn, naast een objectieve maat zoals de DSI kan voor het evalueren van de mening van de patiënt een vragenlijst zoals de VHI worden gebruikt.

Referenties

- De Bodt, M.S., Wuyts, F.L., Van de Heyning, P.H., & Croux, C. (1997). Test-retest study of the GRBAS scale: influence of experience and professional background on perceptual rating of voice quality. *Journal of Voice*, 11(1), 74-80.
- Dejonckere, P.H., Remacle, M., Fresnel-Elbaz, E., Woisard, V., Crevier, L., & Millet, B. (1998). Reliability and clinical relevance of perceptual evaluation of pathological voices. *Revue de Laryngologie Otologie Rhinologie*, 119(4), 247-248.
- Hakkesteeft, M.M., Brocaar, M.P., & Wieringa, M.H. (2010). The applicability of the dysphonia severity index and the voice handicap index in evaluating effects of voice therapy and phonosurgery. *Journal of Voice*, 24(2), 199-205.
- Hakkesteeft, M.M., Brocaar, M.P., Wieringa, M.H., & Feenstra, L. (2006). Influence of age and gender on the dysphonia severity index. A study of normative values. *Folia Phoniatica et Logopaedica*, 58(4), 264-273.
- Hakkesteeft, M.M., Brocaar, M.P., Wieringa, M.H., & Feenstra, L. (2008). The relationship between perceptual evaluation and objective multiparametric evaluation of dysphonia severity. *Journal of Voice*, 22(2), 138-145.
- Hakkesteeft, M.M., Wieringa, M.H., Brocaar, M.P., Mulder, P.G., & Feenstra, L. (2008). The interobserver and test-retest variability of the dysphonia severity index. *Folia Phoniatica et Logopaedica*, 60(2), 86-90.

- Hirano, M. (1981). *Clinical examination of voice*. Wien; New York: Springer.
- Honjo, I., & Isshiki, N. (1980). Laryngoscopic and voice characteristics of aged persons. *Archives of Otolaryngology*, 106(3), 149-150.
- Jacobson, B.H., Johnson, A., Grywalski, C., Silbergleit, A., Jacobson, G., Benninger, M.S., et al. (1997). The Voice Handicap Index (VHI): Development and Validation. *American Journal of Speech-Language Pathology*, 6(3), 66-70.
- Kreiman, J., & Gerratt, B.R. (1998). Validity of rating scale measures of voice quality. *Journal of the Acoustical Society of America*, 104(3 Pt 1), 1598-1608.
- Kreiman, J., Gerratt, B.R., Kempster, G.B., Erman, A., & Berke, G.S. (1993). Perceptual evaluation of voice quality: review, tutorial, and a framework for future research. *Journal of Speech, Language and Hearing Research*, 36(1), 21-40.
- Ma, E.P.M., & Yiu, E.M.L. (2005). Suitability of acoustic perturbation measures in analysing periodic and nearly periodic voice signals. *Folia Phoniatica et Logopaedica*, 57(1), 38-47.
- Ma, E.P.M., & Yiu, E.M.L. (2006). Multiparametric Evaluation of Dysphonic Severity. *Journal of Voice*, 20(3), 380-390.
- Maryn, Y., Corthals, P., De Bodt, M., Van Cauwenberge, P., & Deliyski, D. (2009). Perturbation measures of voice: a comparative study between Multi-Dimensional Voice Program and Praat. *Folia Phoniatica et Logopaedica*, 61(4), 217-226.
- Ramig, L.O., Gray, S., Baker, K., Corbin-Lewis, K., Buder, E., Luschei, E., et al. (2001). The aging voice: a review, treatment data and familial and genetic perspectives. *Folia Phoniatica et Logopaedica*, 53(5), 252-265.
- Sataloff, R.T., Rosen, D.C., Hawkshaw, M., & Spiegel, J.R. (1997). The aging adult voice. *Journal of Voice*, 11(2), 156-160.
- Teles-Magalhaes, L.C., Pegoraro-Krook, M.I., & Pegoraro, R. (2000). Study of the elderly females' voice by phonetography. *Journal of Voice*, 14(3), 310-321.
- Webb, A.L., Carding, P.N., Deary, I.J., MacKenzie, K., Steen, N., & Wilson, J.A. (2004). The reliability of three perceptual evaluation scales for dysphonia. *European Archives of Otorhino-Laryngology*, 261(8), 429-434.
- Wuyts, F.L., De Bodt, M.S., Molenberghs, G., Remacle, M., Heylen, L., Millet, B., et al. (2000). The dysphonia severity index: an objective measure of vocal quality based on a multiparameter approach. *Journal of Speech, Language and Hearing Research*, 43(3), 796-809.