

Een validatie onderzoek naar een nieuwe kwantificatie tool



Inleiding

De ziekte van Parkinson (PD) is de tweede meest voorkomende neurodegeneratieve aandoening met een incidentie van 1-1,5% wereldwijd (1). Een I-123-3-iodo-methoxybenzamide scintigrafie (ofwel een ¹²³I-IBZM-scan) is van diagnostische waarde voor de differentiatie tussen PD en Parkinsonisme (Multiple Systeem Atrofie (MSA) of Progressieve Supranucleaire Parese (PSP)) (2). De diagnostische beoordeling gebeurt in de praktijk op basis van een visuele en kwantitatieve beoordeling (2). Het is belangrijk dat de kwantificering zo nauwkeurig mogelijk gebeurt, omdat op basis hiervan de diagnose wordt gesteld en het behandelbeleid wordt gekozen. Tot op heden worden 2D Regions Of Interest (ROI's) gebruikt om de striatum-occipitaalkwab ratio te bepalen, welke tijdrovend en minder reproduceerbaar blijken te zijn (3,4). De 3D Voxel of Interest (VOI) kwantificatie tool is mogelijk een toekomstig alternatief voor de 2D ROI tool (4). Het is echter nog onduidelijk hoe de tools zich tot elkaar verhouden met betrekking tot de betrouwbaarheid en validiteit (2-4).

Onderzoeksvraag

In hoeverre is de 3D VOI kwantificatie tool een meer betrouwbaar en valide meetinstrument in vergelijking tot de 2D ROI kwantificatie tool voor de diagnosestelling bij patiënten met PD en MSA/PSP?

Methode

Kwantitatief retrospectief onderzoeksdesign: ten behoeve van (t.b.v.) de betrouwbaarheid.

- De inclusiecriteria zijn: datasets tussen 2010 en 2016 en patiënten tussen 45 en 80 jaar oud.
- Op de ruwe data zijn 2D ROI's (eSoft Master ROI) en 3D VOI's (Syngo.via) geplaatst door de onderzoeker. De striatum-occipitaalkwab ratio's zijn vervolgens vergeleken en getoetst op significantie middels een t-test.
- De intra- en inter-observer-variabiliteit is vergeleken tussen twee Medisch Beeldvormers en Bestralingsdeskundigen (MBB'ers) op de afdeling en de onderzoeker. Dit is getoetst middels een Intraclass Correlation Coëfficiënt (ICC) test.
- De 3D beelden (geblindeerd voor medische voorgeschiedenis en 2D diagnosestelling) zijn beoordeeld op diagnose door een Nucleair Geneeskundige (NG'er 1). De 2D beelden zijn in het verleden beoordeeld op diagnose door een andere NG'er (NG'er 2). De diagnosestellingen zijn getoetst op significantie middels de Fisher's Exact toets.

Prospectief experimenteel onderzoeksdesign: t.b.v. de validiteit.

- Bepaling verschil in activiteitsratio tussen striatum en occipitaalkwab in het fantoom versus de gemeten striatum-occipitaalkwab ratio's verkregen met beide tools.
- Het fantoomonderzoek werd uitgevoerd op de Ecam. Het rechter striatum werd gevuld met concentraties van 20 en 25 kBq/ml, het linker striatum met 25 kBq/ml en het achtergrond compartiment met 5 kBq/ml.



Afb. 1: Hersenfantoom met linker en rechter striatum (blauwe structuur) en het achtergrond compartiment (overige hersenen). Beide striata zijn onderverdeeld in de caudatus nucleus en het putamen.

Begeleider: Rudie van de Kolk, rudie.vandekolk@fontys.nl
Opdrachtgever: Dennis Dickerscheid, d.b.m.dickerscheid@asz.nl

Resultaten

- De onderzoekspopulatie bestond uit 54 datasets (♀22, ♂32) met een gemiddelde leeftijd van 68,85 jaar (±8,92).
- Berekend door de onderzoeker is bij de 3D tool de striatum-occipitaalkwab ratio 0,02% hoger dan bij de 2D tool. Hierbij is geen significant verschil aangetoond op basis van een ongepaarde t-test.
- Zoals weergegeven in tabel 1 is bij de 3D tool een significant hogere ICC (>0,9) aangetoond tussen de 2 MBB'ers en de onderzoeker (ofwel een kleinere inter-observer-variabiliteit).
- Tussen de diagnosestellingen is op basis van de Fisher's exact test geen significant verschil aangetoond tussen beide tools.
- De gemeten striatum-occipitaalkwab ratio komt bij de 3D tool bij beide activiteiten meer overeen met de activiteitsratio aanwezig in het fantoom, zoals weergegeven in tabel 2. $1-3 = \text{t.b.v. betrouwbaarheid}$
 $4 = \text{t.b.v. validiteit}$

Tabel 1: Resultaten intra-observer-variabiliteit van de onderzoeker¹ en inter-observer-variabiliteit tussen de onderzoeker en de twee MBB'ers².

Intra-variabiliteit ¹	Waarde ICC test	P-waarde
Links 2D	ICC=0,926; F(53)=25,98	p<0,001
Rechts 2D	ICC=0,927; F(53)=26,29	p<0,001
Links 3D	ICC=0,971; F(53)=67,46	p<0,001
Rechts 3D	ICC=0,965; F(53)=55,39	p<0,001
Inter-variabiliteit ²		
Links 2D	ICC=0,825; F(53)=10,42	p<0,001
Rechts 2D	ICC=0,805; F(53)=9,27	p<0,001
Links 3D	ICC=0,917; F(53)=23,04	p<0,001
Rechts 3D	ICC=0,915; F(53)=22,40	p<0,001

ICC = Intraclass Correlation Coëfficiënt, 2D = 2-Dimensionaal, 3D = 3-Dimensionaal, F = onsystematische variatie.

Tabel 2: Werkelijke ratio in % aanwezig in het fantoom en gemiddelde gemeten striatum-occipitaalkwab ratio in % voor de 2D en 3D kwantificatie tool.

BG = 5 kBq/ml Li = 25 kBq/ml	Fantoom Ratio in %	2D tool S-O ratio in %	3D tool S-O ratio in %	Verskil fantoom ratio vs 2D tool ratio in %	Verskil fantoom ratio vs 3D tool ratio in %
Re 20 kBq/ml	1,11	1,14	1,12	0,03	0,01
Re 25 kBq/ml	1,00	1,04	1,02	0,04	0,02

kBq/ml = kiloBecquerel per milliliter, BG = background, li = links, re = rechts, 2D = 2-Dimensionaal, 3D = 3-Dimensionaal, S-O = striatum-occipitaalkwab, vs = versus.

Discussie

- De 3D tool geeft een kleinere inter-observer-variabiliteit wat resulteert in meer consistentie en een hogere betrouwbaarheid (3).
- De verschillen in diagnosestelling zijn niet te wijden aan het verschil in kwantificatie tool. NG'er 2 had wel zicht op de medische voorgeschiedenis, maar NG'er 1 niet. Hierdoor is de diagnosestelling niet valide, maar wel klinisch relevant. De niet valide diagnosestelling geeft reden voor een vervolgonderzoek naar het bepalen van de inter-beoordelaars betrouwbaarheid tussen meerdere NG'ers, met daarbij de follow-up van de patiënt voor het berekenen van de sensitiviteit en specificiteit (2).
- De verschillen van het fantoomonderzoek zijn klinisch relevant door het verschil van 0,02% tussen beide tools, omdat op basis hiervan overstapt zou kunnen worden op de 3D tool. De 3D tool is meer valide omdat de gemeten ratio meer overeenkomt met de werkelijke ratio dan bij de 2D tool. De validiteit van het fantoomonderzoek is verlaagd door een lek in de linker caudatus. De betrouwbaarheid is verlaagd doordat twee scans zijn vervaardigd.

Conclusie en aanbeveling

De 3D tool lijkt hoger te scoren op zowel betrouwbaarheid als validiteit. Op basis van de interpretatie van de diagnosestellingen kan niet met volledige zekerheid geconcludeerd worden dat de 3D tool een meer betrouwbaar en valide meetinstrument is t.o.v. de 2D tool bij het diagnosticeren van PD en MSA/PSP. Het advies voor de praktijk is om tijdelijk beide tools in gebruik te nemen en over een langere periode een besluit te nemen welke tool gebruikt dient te worden in de toekomst.

Referenties

- Fang Y-HD et al. Fully Automated Quantification of the Striatal Uptake Ratio of [99mTc]-TRODAT with SPECT Imaging: Evaluation of the Diagnostic Performance in Parkinson's Disease and the Temporal Regression of Striatal Tracer Uptake. BioMed Research International. 2015;2015:461625.
- Heinzel A. et al. Comparison of automatic versus manual procedures for the quantification of dopamine D2 receptor availability using I-123-IBZM-SPECT. Nuclear medicine communications. 2015;36(11):1120.
- Pöppel G et al. Diagnostic performance of a 3-D automated quantification method of dopamine D2 receptor SPECT studies in the differential diagnosis of parkinsonism. Nuclear medicine communications. 2005;26(1):39.
- Buchert R et al. IBZM tool: a fully automated expert system for the evaluation of IBZM SPECT studies. European Journal of Nuclear Medicine and Molecular Imaging. 2006;33(9):1073-83.