Boekbespreking

Gorter SL. Rheumatologists' daily practice performance. A study with incognito standardized patients [dissertatie]. Maastricht: Universitaire Pers Maastricht; 2003.160 pagina's. ISBN 90 5278 393 4. Prijs € 24.75.

Assessment van klinische performance aan het einde van de opleiding tot arts of specialist, of van praktiserende specialisten vergt een geïntegreerde beoordeling van kennis, vaardigheden en attitude. Het is assessment van de top van de piramide van Miller, de *does fase*.¹ Deze assessment is wel beschreven als internationale uitdaging van de eeuw.² Een van de mogelijke oplossingen die door Miller wordt gegeven is assessment met behulp van undercover simulatie patiënten.¹ In het proefschrift van Simone Leonie Gorter wordt het onderzoek naar een variant hierop, assessment van de dagelijkse performance van reumatologen met behulp van incognito gestandaardiseerde patiënten (SP's), beschreven. Simulatiepatiënten en geïnstrueerde echte patiënten bezochten incognito het spreekuur van reumatologen die vooraf een informed consent hadden gegeven. De performance werd gescoord door de patiënten.

Voor een dergelijk onderzoek is het noodzakelijk als standaard een case-specifieke checklist te ontwikkelen, waarop vaardigheden als anamnese afnemen, lichamelijk onderzoek, case management en communicatie kunnen worden gescoord. In hoofdstuk 2 wordt een review van de literatuur gegeven, waaruit blijkt dat de ontwikkeling van een checklist nog geen gestandaardiseerde procedure is. In dit onderzoek is gekozen voor de samenstelling van een case-specifieke checklist door 11 reumatologen, die door de hierin getrainde SP wordt ingevuld. Attitude en communicatieve vaardigheden waren niet in de checklist opgenomen.

In hoofdstuk 3 wordt de logistieke uitdaging, die deze vorm van assessment is, beschreven. Dankzij deze inspanning bleef wel 98% van de SP's incognito. De belangrijkste conclusie is dat de procedure dermate arbeidsintensief is dat alleen gebruik op strikte indicatie mogelijk is.

Deze indicaties zouden kunnen zijn het assessment van de klinische performance van een individuele arts, gericht onderzoek naar verschillen in klinische praktijk of identificatie van onderwerpen in aanmerking komend voor scholing.

In dit onderzoek hebben 16 personen de rol van 8 SP's gespeeld. Elke SP bezocht 13 reumatologen. In totaal participeerden 27 reumatologen. Doordat de uitval gering was, werden hiermee 201 eerste kliniekbezoeken en 53 vervolgb bezoeken gerealiseerd. Er blijkt een factor 10 verschil te bestaan in de kosten van de aangevraagde diagnostiek.

Ook het therapeutisch beleid verschilt sterk per reumatoloog met aanzienlijke afwijkingen van het door een panel van 11 reumatologen voorgestelde beleid. Opvallend is dat met name werkervaring van de arts negatief correleert met de checklistscore. Dit laat zich niet eenduidig verklaren. De statistische onderbouwing is voldoende om de resultaten als betrouwbaar en reproduceerbaar te beschouwen.

Gezien de bewerkelijkheid van het assessment met behulp van SP's is ook onderzocht of de score op een gecomputeriseerde case-based test correleerde met de uitkomst van de SP-test. Er blijkt een negatieve correlatie te zijn. Merkwaardigerwijs blijkt er een relatie met ervaring. Reumatologen met meer ervaring scoren lager in de SP-gebaseerde assessment en hoger op de gecomputeriseerde test. Dit verschil wordt niet verklaard doordat ervaren reumatologen minder informatie nodig hebben, en dus lager op de checklist scoren, om tot een diagnose te komen. Ook op de key-items scoren zij lager. Beide testen meten waarschijnlijk verschillende eigenschappen.

Het belangrijke aspect van de dokter-patiënt interactie werd onderzocht met een focusgroep-interview met de meerderheid van de SP's. Het voordeel van SP's is dat zij geen afhankelijkheidsrelatie met de dokter hebben en dus objectiever hun mening geven. Ondanks de grote verschillen in benadering waren de SP's enthousiast over de reumatologen. Natuurlijk

waren er ook negatievere aspecten. Dit betrof vooral een soms gepercipieerd gebrek aan respect, wachtkamerervaringen en een grote variatie in gedrag tussen specialisten of tijdens een eerste en tweede bezoek. De verwijfsbrief van de huisarts was te vaak aanleiding tot een te gerichte anamnese zonder de mogelijkheid om daarin niet genoemde klachten naar voren te brengen.

Dit onderzoek naar de mogelijkheid van assessment van de dagelijkse performance van reumatologen legt een aantal problemen bloot. De performance is gemeten met de score op een checklist, ingevuld door gestandaardiseerde patiënten. De resultaten laten een grote variatie in dagelijkse performance zien. Deze uitslag is toch moeilijk te interpreteren. De kernvraag is of de score op een checklist wel meet wat ze zou moeten meten. Daaraan vooraf gaat nog de vraag of dat wat we willen weten wel zodanig gedefinieerd kan worden dat het zich leent voor objectieve meting. We willen weten of iemand een goede arts is. Zolang een definitie van wat een goede arts is discutabel is, zullen assessments ook vragen blijven, oproepen omdat altijd de vraag zal ontstaan of de meetwaarde wel iets zegt over kwaliteit. Intuïtief benadert de hier gebruikte manier van testen wel het beste de werkelijkheid, immers de reumatoloog zal zich gedragen zoals hij zich anders ook gedraagt. Misschien zit het probleem wel in de checklist. Het zou kunnen zijn dat de checklist die in alle rust is uitgediscussieerd niet zoveel zegt over de performance in de spreekkamer. Opvallend is ook dat de correlatie met de gecomputeriseerde patiënt-based test een negatieve is. Ervaren artsen scoren hoog op de gecomputeriseerde test en laag op de SP-assessment.

Als assessment van de individuele arts is deze test bij ontbreken van een goede standaard ongeschikt. Voor assessment van grotere groepen is bovendien de bewerkelijkheid een belemmerende factor. Afgeleide vragen zoals 'waarin zitten de grote verschillen in beleid, zodat hier over gediscussieerd kan worden om tot meer eenheid te komen' zijn wellicht wel te beantwoorden. Voorwaarde is dan wel dat er een algemeen erkende standaard bestaat. Bijvoorbeeld de tennisselleboog, de diagnose van één van de SP's, kent meerdere equivalenten behandelingen. Een geconstateerde afwijking van de standaard door het individu of de groep zou aanleiding kunnen zijn voor gerichte nascholing. In dat geval moet er wel een systeem voor implementatie en effectiviteitsmeting worden ontwikkeld. Ook dit is een zeer arbeidsintensieve procedure.

Concluderend wordt in dit proefschrift een degelijke poging gedaan om de uitdaging van de eeuw aan te gaan. Opnieuw blijkt dat naarmate de assessment beter is, het meer inspanning kost en dat daarmee de haalbaarheid in de praktijk moeilijk wordt. Bovendien blijft de vraag bestaan of de standaard, de checklist, wel representatief is voor wat we willen weten.

Literatuur

1. Miller GE. The assessment of clinical skills/competence/performance. *Acad Med* 1990;65(9 Suppl): S63-S67.
2. Wass V, Vleuten C van der, Shatzer J, Jones R. Assessment of clinical competence. *Lancet* 2001;357 (9260):945-9.

Dr. ing. J.G. Zijlstra

Internist-intensivist, Academisch Ziekenhuis Groningen