



Accuraatheid en precisie van spirometers

Samenvatting

Schermer TR, Verweij EH, Cretier E, Pellegrino A, Poels PJ. Accuraatheid en precisie van spirometers. *Huisarts Wet* 2012;55(9):382-6

DOEL Spirometrie is een belangrijke test voor het diagnosticeren en monitoren van chronische luchtwegaandoeningen. Veel huisartsenpraktijken beschikken over een eigen spirometer en gebruiken deze geregeld in hun patiëntenzorg, maar het is niet bekend hoe deze apparaten presteren. Daarom wilden we de accuraatheid en precisie van spirometers in Nederlandse huisartsenpraktijken bepalen.

METHODE We hebben 49 spirometers uit een landelijke steekproef van 300 huisartsenpraktijken getest door met een computergestuurde ijkspuit acht gestandaardiseerde blaaspatronen toe te dienen. Meetprestaties drukten we uit als 'accuraatheid' (de afwijking van de ijkspuitwaarde) en 'precisie' (het verschil tussen de hoogste en laagste waarde) voor de FEV₁ (eensecondewaarde) en FVC (geforceerde vitale capaciteit). De turbinespirometers in de steekproef verdeelden we over drie leeftijdscategorieën (1-2, 3-5, ≥ 6 jaar oud). Aan de hand van variantieanalyse analyseerden we het verband tussen leeftijd van turbinespirometers en hun accuraatheid en precisie, evenals eventuele verschillen tussen turbinemeters en pneumotachografen.

RESULTATEN In de steekproef zaten 39 turbinespirometers en 8 pneumotachografen. Gemiddelde leeftijd van de spirometers was 4,3 (sd 3,7) jaar. De accuraatheid voor alle blaaspatronen samen was + 25 ml (95%-BI 12-39) voor de FEV₁ en + 27 ml (95%-BI 10-45) voor de FVC, met uitschieters in beide richtingen van > 200 ml voor FEV₁ en > 500 ml voor FVC. De gemiddelde precisie was 14 ml (95%-BI 11-17) voor FEV₁ en 35 ml (95%-BI 24-47) voor FVC. Voor de FEV₁ vertoonden de pneumotachografen qua accuraatheid grotere afwijkingen dan de turbinemeters ($p < 0,003$), voor de FVC was dat niet het geval. Bij de turbine-meters vonden we geen relatie tussen leeftijd en meetprestaties.

CONCLUSIE Spirometers in Nederlandse huisartsenpraktijken meten gemiddeld iets te hoge FEV₁- en FVC-waarden, al vertonen sommige apparaten aanzienlijke meetfouten. Turbinespirometers lijken accurater en preciezer te zijn dan pneumotachografen. Oudere turbinemeters presteren niet per se slechter dan jongere.

INLEIDING

In de afgelopen jaren is spirometrie voor huisartsen en praktijkondersteuners uitgegroeid tot een belangrijke test voor het diagnosticeren en monitoren van chronische luchtwegaandoeningen.^{1,2} Naar schatting 62% van de huisartsenpraktijken in Nederland beschikt over een eigen spirometer en het aantal in praktijken verrichte spirometrietests is in enkele jaren verdrievoudigd,³ al blijft de kwaliteit van uitvoering een belangrijk aandachtspunt.⁴

Er zijn in praktijken diverse merken en typen spirometers in gebruik.³ Eerder onderzoek heeft laten zien dat de meeste van deze apparaten de FEV₁ (geforceerd expiratoir volume in 1 seconde) en de FVC (geforceerde vitale capaciteit) nauwkeurig kunnen meten.⁵ De spirometers in het betreffende onderzoek kwamen echter rechtstreeks van de leveranciers, waardoor de bevindingen waarschijnlijk niet representatief zijn voor meters die (soms al jarenlang) in gebruik zijn. Onderhoud en kalibratie van spirometers in huisartsenpraktijken wijken immers vaak af van de aanbevelingen van de leveranciers,^{3,6,7} wat een negatieve invloed kan hebben op de meetprestaties van de apparatuur.⁸ Dit kan weer gevolgen hebben voor de uitslag van spirometrietests en de daaruit voortvloeiende behandeling van patiënten.

Tot op heden is er geen onderzoek gedaan naar de prestaties van spirometers die in huisartsenpraktijken worden gebruikt als onderdeel van de dagelijkse patiëntenzorg. Doel van ons onderzoek was daarom het vaststellen van de accuraatheid en precisie van deze spirometers.

Wat is bekend?

- Spirometrie is een belangrijke test voor het diagnosticeren en monitoren van chronische luchtwegaandoeningen.
- Als ze nieuw zijn, meten spirometers die men in huisartsenpraktijken gebruikt de FEV₁ en de FVC nauwkeurig genoeg.
- Onderhoud en kalibratie van spirometers in huisartsenpraktijken wijken vaak af van de aanbevelingen van de leveranciers.
- Of spirometers in de dagelijkse patiëntenzorg qua nauwkeurigheid goed blijven presteren, is niet bekend.

Wat is nieuw?

- Spirometers in Nederlandse huisartsenpraktijken meten gemiddeld iets te hoge FEV₁- en FVC-waarden, maar sommige apparaten vertonen aanzienlijke meetfouten.
- Turbinespirometers lijken accurater en preciezer te zijn dan pneumotachografen.
- Oudere turbinemeters presteren niet per se slechter dan jongere.
- Huisartsenpraktijken zouden meer aandacht moeten besteden aan de controle van hun spirometer, liefst met een ijkspuit en door 'biologische controle'.

UMC St Radboud, afdeling Eerstelijns geneeskunde, Postbus 9101, 6500 HB Nijmegen: dr. T.R. Schermer, senior onderzoeker; E.H. Verweij, huisarts-onderzoeker; R. Cretier, longfunctieanalist, onderzoeksassistent. Afdeling Epidemiologie, Biostatistiek en Health Technology Assessment, UMC St Radboud, Nijmegen: A. Pellegrino, onderzoeksassistent. Huisartsenpraktijk Bles en Poels, Huissen: dr. P.J. Poels, huisarts-onderzoeker • Correspondentie: T.Schermer@elg.umcn.nl • Mogelijke belangenverstrengeling: niets aangegeven.

Dit artikel is een bewerkte vertaling van: Schermer TR, Verweij EH, Cretier R, Pellegrino A, Crockett AJ, Poels PJ. Accuracy and precision of desktop spirometers in general practices. *Respiration* 2012;83:344-52. Publicatie gebeurt met toestemming van de uitgever.

METHODE

Onderzoeksopzet

In 2007 trokken wij uit het landelijke huisartsenpraktijkenbestand van het Nivel ($n = 4533$ praktijken) een steekproef van 25 praktijken per provincie.³ Deze 300 praktijken stelden wij telefonisch de vraag of zij over een spirometer beschikten en zo ja, of zij bereid waren deze tijdelijk ter evaluatie aan ons af te staan. Met behulp van een vragenlijst registreerden wij van elke spirometer het merk, type en jaar van aankoop, evenals de frequentie van gebruik, reiniging en ijking. Ook inventariseerden we de geschiedenis van eventuele beschadigingen en reparaties. Vooraf berekenden we geen steekproefgrootte, maar het doel was 50 spirometers in het onderzoek op te nemen.

Evaluatie van de spirometers

Een getrainde medewerker testte de geworven spirometers op een vaste locatie (PT Medical bv, Leek) op een *waveform generator*: een computergestuurde ijkspuit die 24 gestandaardiseerde blaaspatronen kan genereren.⁹ We selecteerden acht van deze blaaspatronen, die samen een breed spectrum van FEV₁- en FVC-waarden vertegenwoordigen. De medewerker diende elk blaaspatroon drie keer achter elkaar aan een spirometer toe. De meetprestatie van een spirometer drukten we uit als de 'accuraatheid' en de 'precisie' van de door het apparaat gemeten FEV₁- en FVC-waarden. Voor elke spirometer berekenden we per blaaspatroon het gemiddelde van de drie toedieningen en vergeleken dat met de referentiewaarde zoals afgegeven door de ijkspuit. Voor zowel de FEV₁ als de FVC beschouwden we een spirometer als 'inaccuraat' voor een bepaald blaaspatroon als het verschil ≥ 100 ml was (positief of negatief).⁹ De precisie van de spirometers bepaalden we door per blaaspatroon het verschil tussen de hoogste en de laagste van de drie FEV₁- en FVC-waarden te berekenen. We beschouwden de precisie van een spirometer als onvoldoende als het verschil ≥ 100 ml was.⁹

Analyse

Accuraatheid en de precisie drukten we uit als gemiddelde (95%-BI), zowel voor de acht blaaspatronen afzonderlijk als voor alle blaaspatronen gecombineerd. Voor verdere analyse van de accuraatheid zetten we in zogenaamde 'Bland-Altman-grafieken'¹⁰ het verschil tussen de door een spirometer gemeten waarde en de referentiewaarde van de ijkspuit uit tegen het gemiddelde van die beide waarden. De turbinespirometers in de steekproef verdeelden we over drie leeftijdscategorieën: 1 tot 2 jaar, 3 tot 5 jaar en 6 jaar en ouder. We gebruikten variantieanalyse (ANOVA) om het verband tussen de leeftijd van de turbinespirometers en hun accuraatheid en precisie voor de FEV₁ en FVC te analyseren. ANOVA gebruikten we ook om verschillen tussen typen spirometers (turbinespirometers versus pneumotachografen) te analyseren. Om te corrigeren voor herhaalde vergelijkingen definieerden we statistische significantie als $p < 0,0031$ (Bonferoni-correctie: $\alpha = 0,05/2 \times 8$ blaaspatronen).

RESULTATEN

Spirometers in het onderzoek

We konden met 269 van de 300 praktijken in de steekproef (90%) telefonisch contact leggen en er waren er 250 (83%) bereid om informatie voor het onderzoek te verschaffen.³ Honderdvierenvijftig praktijken (62%) gaven aan over een of meer spirometers te beschikken, waarvan er 63 (41%) bereid waren om hun spirometer te laten testen. Uiteindelijk werden er 50 (32%) spirometers opgestuurd naar de testlocatie, waar 1 spirometer niet kon worden getest omdat de benodigde software niet beschikbaar was. De belangrijkste redenen voor praktijken om hun spirometer niet te laten testen waren dat ze 'hun spirometer niet konden missen' (35%) of 'te druk waren om verzending van de spirometer te regelen' (21%).

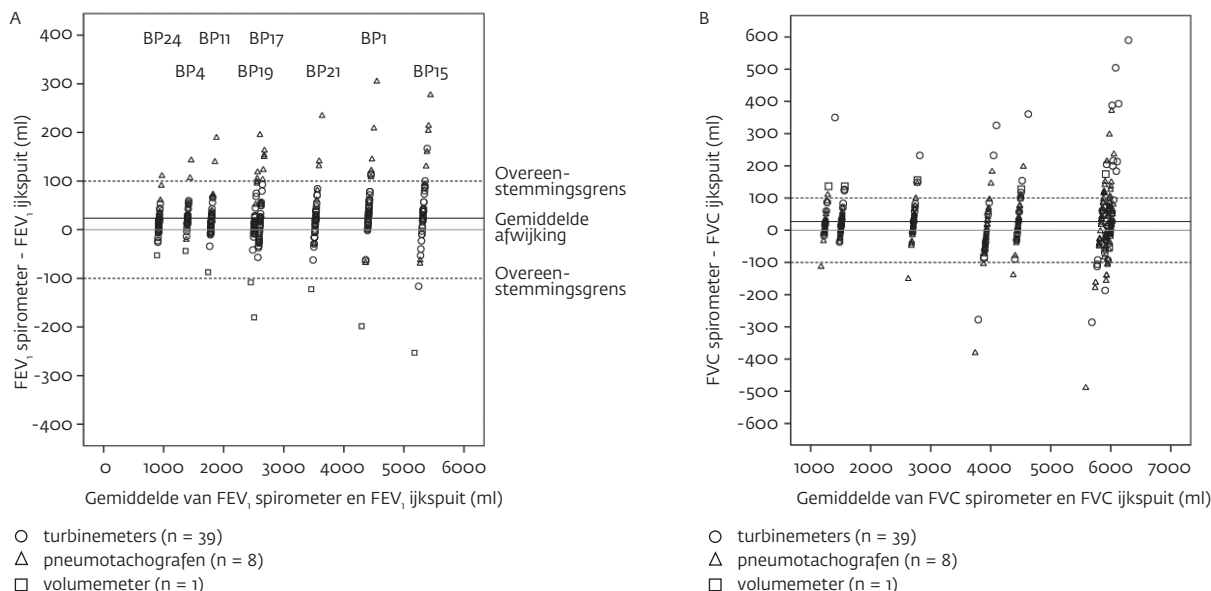
Van de 49 geteste spirometers waren er 38 (78%) van Micro-Medical (tegenwoordig CareFusion), 6 (12%) van Vitalograph, 3 (6%) van Clement Clarke International en 1 (2%) van Mijnhardt.

Tabel Accuraatheid en precisie per blaaspatroon⁹ en totaal voor de steekproef van 49 spirometers. Getallen zijn gemiddelden (95%-BI)

Blaaspa- troon	Ijkspuut* (ml)	FEV ₁				FVC				
		Accuraatheid (ml)		Precisie (ml)		Ijkspuut* (ml)	Accuraatheid (ml)		Precisie (ml)	
1	4395	42	(21-62)	20	(15-26)	6000	7	(-26 - -40)	29	(18-40)
4	1387	25	(17-33)	10	(6-13)	1500	26	(15-37)	16	(12-20)
11	1784	20	(9-32)	10	(6-14)	2701	35	(18-51)	29	(20-38)
15	5303	39	(15-63)	22	(17-27)	5934	48	(22-73)	36	(22-50)
17	2597	18	(1-35)	14	(8-19)	5829	49	(3-95)	52	(31-73)
19	2505	20	(7-33)	13	(10-17)	3931	-9	(-39-22)	32	(21-43)
21	3519	24	(8-39)	16	(10-21)	4446	34	(12-57)	38	(16-60)
24	916	15	(7-23)	9	(7-12)	1230	29	(11-46)	24	(5-43)
Totaal		25	(12-39)	14	(11-17)		27	(10-45)	35	(24-47)

Accuraatheid: verschil tussen het gemiddelde van drie bepalingen van de FEV₁ respectievelijk FVC en de referentiewaarde zoals afgegeven door de ijkspuit
 Precisie: verschil tussen de hoogste en de laagste waarde van de drie FEV₁- en FVC-bepalingen voor het betreffende blaaspatroon BI: betrouwbaarheidsinterval; FEV₁: geforceerde expiratoire volume in een seconde; FVC: geforceerde vitale capaciteit * Computergestuurde ijkspuit die 24 gestandaardiseerde blaaspatronen kan genereren⁹

Figuur 1 Bland-Altman-grafiek voor de accuraatheid van de FEV₁ (A) en de FVC (B) voor de spirometers in de steekproef. Elk symbool is het resultaat voor een van de acht toegepaste blaaspatronen (gemiddelde van drie toedieningen per blaaspatroon) van één spirometer. De stippellijnen zijn de 'grenzen van overeenstemming' (+/- 100 ml⁹)



BP: blaaspatroon (de door de computergestuurde ijkspuit afgegeven FEV₁- en FVC-waarden per blaaspatroon staan in de tabel op de vorige pagina)

Voor 1 spirometer ontbrak informatie over merk en type sensor. Negenendertig spirometers (80%) waren turbinespirometers, 8 (16%) waren pneumotachografen en 1 (2%) was een volumespirometer. De gemiddelde leeftijd van alle spirometers was 4,3 (sd 3,7) jaar. De 'jongste' spirometer was 4 maanden oud, de oudste 18 jaar. De gemiddelde leeftijd van de turbinespirometers was 3,7 (sd 2,9) jaar.

Nauwkeurigheid en precisie van de spirometers

De gemiddelde afwijking van de spirometers ten opzichte van de referentiewaarden van de ijkspuit (accuraatheid) voor alle blaaspatronen samen bedroeg 25 ml (95%-BI 12-39) voor de FEV₁ en 27 ml (95%-BI 10-45) voor de FVC. Het gemiddelde verschil tussen de laagste en de hoogste waarde (precisie) was 14 ml (95%-BI 11-17) voor de FEV₁ en 35 ml (95%-BI 24-47) voor de FVC. De [tabel] toont de nauwkeurigheid en precisie per blaaspatroon.

[Figuur 1] laat voor alle spirometers de afwijkingen zien tussen de gemeten FEV₁ en FVC, en de ijkspuitwaarden voor de geselecteerde blaaspatronen. Voor de FEV₁ [figuur 1a] laat de grafiek zien dat er sprake is van enige nonlineariteit van de accuraatheid: afwijkingen van de ijkspuitwaarden worden groter naarmate de FEV₁ hoger is. De meeste sterk afwijkende FEV₁-waarden waren toe te schrijven aan 5 van de 8 pneumotachografen in de steekproef en aan de ene volumespirometer [figuur 1a]. Ook voor de FVC zagen we nonlineariteit van de afwijkingen [figuur 1b], met maximale afwijkingen van +590 en -490 ml.

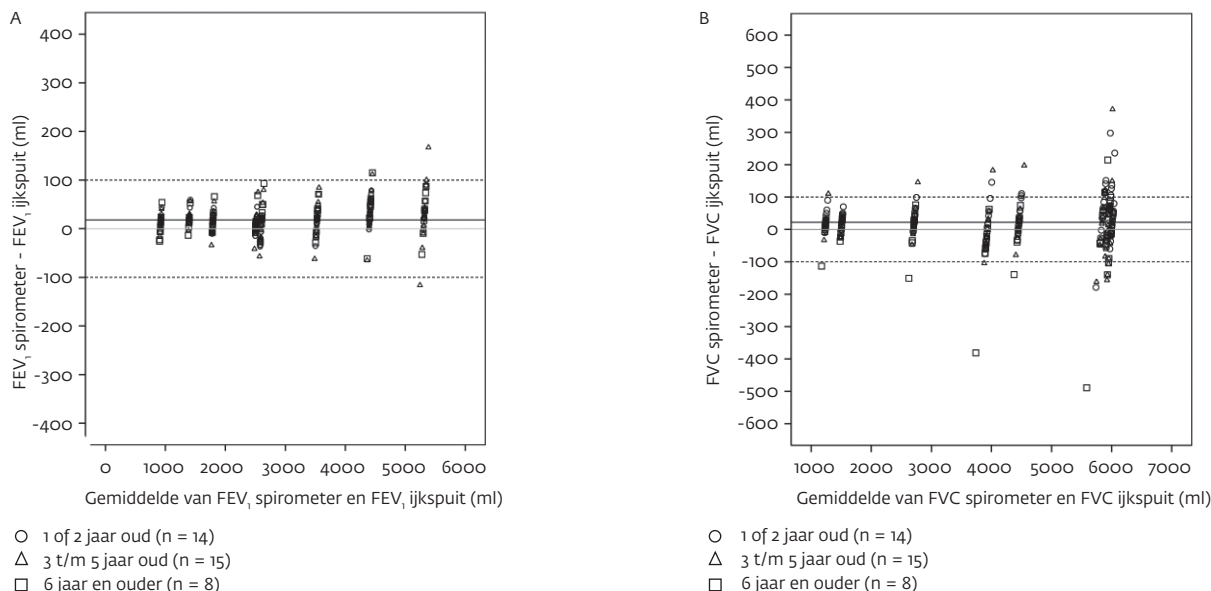
Voor de accuraatheid van de FEV₁ gold voor alle blaaspatronen dat de pneumotachografen statistisch significant grotere afwijkingen vertoonden dan de turbinespirometers. Voor de FVC was dit bij géén van de blaaspatronen het geval. Wat betreft de precisie lieten de pneumotachografen eveneens significant grotere verschillen zien dan de turbinespirometers: bij de FEV₁ voor vijf blaaspatronen (nummers 1, 4, 11, 17 en 19) en bij de FVC voor drie blaaspatronen (nummers 1, 4 en 19).

Invloed van leeftijd op meetprestaties van turbinespirometers [figuur 2] toont voor de FEV₁ en FVC Bland-Altman-grafieken voor de accuraatheid van de turbinespirometers, verdeeld in drie leeftijdscategorieën. Voor de FEV₁ [figuur 2a] bleken vier meters (alle ouder dan 2 jaar) afwijkingen voor een of twee blaaspatronen te vertonen. Voor de FVC [figuur 2b] gold dat zeventien meters (evenredig verdeeld over de drie leeftijdscategorieën) voor een of twee blaaspatronen afwijkingen vertoonden. Voor accuraatheid noch precisie vonden we statistisch significante verschillen tussen de drie leeftijdscategorieën.

BESCHOUWING

In dit onderzoek hebben we 49 spirometers van huisartsenpraktijken uit een landelijke steekproef op hun accuraatheid en precisie getest met behulp een computergestuurde ijkspuit. Qua aantal domineerden turbinespirometers de steekproef, wat overeenkomt met het landelijke beeld.³ Voor alle spirometers samen was de gemiddelde afwijking van de ijkspuit (accuraatheid) ongeveer +25 milliliter voor zowel de FEV₁ als de FVC. Het gemiddelde verschil tussen de hoogste en de laagste van

Figuur 2 Bland-Altman-grafiek voor de accuraatheid van de FEV₁ (A) en de FVC (B) per leeftijdscategorie voor de turbinespirometers (n = 37) in de steekproef. Elk symbool is het resultaat voor een van de acht toegepaste blaaspatronen (gemiddelde van drie toedieningen per blaaspatroon) van één spirometer. De stippellijnen zijn de 'grenzen van overeenstemming' (+/- 100 ml⁹).



drie metingen op de ijkspuit (precisie) was 14 milliliter voor de FEV₁ en 35 milliliter voor de FVC. Hoewel het aantal pneumotachografen in de steekproef klein was (n = 8), suggereren onze bevindingen dat dit type spirometer gevoeliger is voor afwijkingen dan turbinemeters. In de 38 geteste turbinemeters vonden we geen verband tussen de leeftijd van apparaten en de meetprestaties.

Sterke en zwakte punten van het onderzoek

Van alle praktijken die aangaven over een eigen spirometer te beschikken was ongeveer een derde bereid hun spirometer te laten testen. De precieze overwegingen om wel of niet aan het onderzoek deel te nemen kennen we niet, maar we kunnen niet uitsluiten dat praktijken die meededen vaker (terecht of onterecht) twijfelden aan de betrouwbaarheid van hun spirometer en deelname aan het onderzoek zagen als een mogelijkheid om hun apparatuur te laten controleren. Dit kan tot een groter aandeel minder betrouwbare spirometers in onze steekproef hebben geleid, met als mogelijk gevolg een te negatief beeld van de gemiddelde accuraatheid en precisie.

Vergelijking met eerder onderzoek

Uit evaluatie van tien nieuwe elektronische spirometers bij personen die ook met een geavanceerde spirometer in een longfunctielaboratorium werden gemeten, concludeerden Belgische onderzoekers enkele jaren geleden dat nieuwe turbinespirometers qua accuraatheid van FEV₁ en FVC over het algemeen goed presteren, terwijl dat bij sommige pneumotachografen in mindere mate het geval was.⁵ Onze bevindingen

bij spirometers die al (geruime tijd) in praktijken in gebruik zijn lijken deze conclusie te ondersteunen. Een ouder onderzoek (uit 1996) naar het verloop van de meetprestaties van turbinemeters in de loop der tijd wijst ook op een goede stabiliteit van dit type spirometer.¹¹

Implicaties voor de praktijk

De meeste praktijken in het onderzoek bleken weinig of geen aandacht te besteden aan controle en ijking van hun spirometers.³ Voor sommige bleek dit niet gerechtvaardigd: bij evaluatie op de ijkspuit lieten de spirometers gemiddeld weliswaar een beperkte (circa 25 milliliter) afwijking in FEV₁ en FVC zien, maar sommige apparaten – in het bijzonder pneumotachografen – vertoonden aanzienlijke afwijkingen. We hebben geen gegevens van patiënten die met de betreffende spirometers zijn gemeten, maar het is zeer wel mogelijk dat de onnauwkeurigheid van deze apparaten zich in de betreffende praktijken heeft vertaald in onjuiste uitslagen en tot verkeerde interpretaties van spirometrietests heeft geleid.

Een van de problemen met spirometrie in huisartsenpraktijken is dat er slechts beperkte mogelijkheden zijn om de accuraatheid en precisie van de apparatuur te controleren. Met een ijkspuit kan men de luchtstroomsnelheden die een spirometer meet (waaronder de FEV₁) niet controleren.¹² Regelmatige 'biologische controle' (een respiratoir gezonde praktijkmedewerker maximaal laten blazen) is wel een haalbare en zinvolle optie.¹³ Bij de keuze voor een spirometer zou de praktijk rekening moeten houden met de robuustheid van de verschillende soorten sensoren en de mate van onderhoud en

ijking die zij nodig hebben, naast andere belangrijke factoren, zoals hygiëne,¹⁴ gebruiksvriendelijkheid en kosten van apparaat en mondstukken.

CONCLUSIE

Dit onderzoek liet zien dat spirometers in Nederlandse huisartsenpraktijken gemiddeld genomen iets te hoge FEV₁- en FVC-waarden meten, maar ook dat sommige apparaten aanzienlijke meetfouten vertonen. Onze steekproef bestond vooral uit turbinespirometers, maar het waargenomen gebrek aan accuraatheid zagen we vooral bij pneumotachografen. Bij de turbinespirometers in de steekproef vonden we geen relatie tussen leeftijd en meetprestaties. In het algemeen geldt dat huisartsenpraktijken meer aandacht zouden moeten besteden aan de controle van hun spirometer, bij voorkeur zowel met een ijspuut als door 'biologische controle'. ■

LITERATUUR

- 1 Smeele IJM, Van Weel C, Van Schayck CP, Van der Molen T, Thoonen B, Schermer T, et al. NHG-Standaard COPD (tweede herziening). Huisarts Wet 2007;50:362-79.
- 2 Levy ML, Fletcher M, Price DB, Hausen T, Halbert RJ, Yawn BP. International Primary Care Respiratory Group (IPCRG) Guidelines: diagnosis of

- respiratory diseases in primary care. Prim Care Respir J 2006;15:20-34.
- 3 Schellekens D, Poels P, Pellegrino A, Cretier R, Smeele I, Schermer T. Spirometrie in de Nederlandse huisartsenpraktijk. Resultaten van een landelijke survey. Huisarts Wet 2008;51:434-9.
- 4 Landman M, Gilissen T, Grootens-Stekelenburg J, Akkermans R, Schermer T. Kwaliteit van spirometrie in de eerste lijn. Huisarts Wet 2011;54:536-42.
- 5 Liistro G, Vanwelde C, Vincken W, Vandevoorde J, Verleden G, Buffels J. Technical and functional assessment of 10 office spirometers: A multicenter comparative study. Chest 2006;130:657-65.
- 6 Johns DP, Burton D, Walters JA, Wood-Baker R. National survey of spirometer ownership and usage in general practice in Australia. Respirology 2006;11:292-8.
- 7 Dowson LJ, Yeung A, Allen MB. General practice spirometry in North Staffordshire. Monaldi Arch Chest Dis 1999;54:186-8.
- 8 Townsend MC, Hankinson JL, Lindesmith LA, Slivka WA, Stiver G, Ayres GT. Is my lung function really that good? Flow-type spirometer problems that elevate test results. Chest 2004;125:1902-9.
- 9 American Thoracic Society. Standardization of spirometry, 1994 Update. Am J Respir Crit Care Med 1995;152:1107-36.
- 10 Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986;1:307-10.
- 11 Dirksen A, Madsen E, Pedersen OF, Vedel AM, Kok-Jensen A. Long-term performance of a hand held spirometer. Thorax 1996;51:973-6.
- 12 Van den Boom G, Van der Star LM, Folgering H, Van Schayck CP, Van Weel C. Volume calibration alone may be misleading. Respir Med 1999;93:643-7.
- 13 Ninaber MK, Schot R, Fregonese L, Stolk J. A syringe simulation of biological controls for quality assessment of prospective lung volume measurements. Respiration 2008;76:187-92.
- 14 Hancock KL, Schermer TR, Holton C, Crockett AJ. Microbiological contamination of spirometers – an exploratory study in general practice. Aust Fam Physician 2012;41:63-4.



Foto: Cuus Herbschleb