

Klinische bruikbaarheid en kwaliteit van spirometrie

Susanne van de Hei, Bertine Flokstra-de Blok, Hendrik-Jan Baretta, Nynke Doornewaard, Thys van der Molen, Niels Patberg, et al.

Spirometrietests in de huisartsenpraktijk blijken vaak niet te voldoen aan de ATS/ERS-kwaliteitscriteria. Betekent dat ook dat de kwaliteit van spirometrie voor dagelijks klinisch gebruik niet voldoende is? We probeerden die vraag te beantwoorden door een vergelijking te maken tussen de diagnoses van huisartsen en die van longartsen. Ook vroegen we longartsen de klinische bruikbaarheid van de spirometrietests te evalueren, en we vroegen longfunctieanalisten om de tests op basis van de ATS/ERS-kwaliteitscriteria te beoordelen. De klinische bruikbaarheid van spirometrie in de huisartsenpraktijk blijkt goed te zijn, hoewel er weinig overeenkomst is tussen de diagnoses van longartsen en die van huisartsen.

In Nederland zijn er naar schatting 1,8 miljoen mensen met astma (met een prevalentie in de huisartsenpraktijk van 102 patiënten per 1000 patiënten) en 580.000 mensen met COPD (met een prevalentie in de huisartsenpraktijk van 34 per 1000 patiënten).^{1,2} Spirometrie is essentieel voor het diagnosticeren en monitoren van luchtwegobstructie en wordt in nationale en internationale richtlijnen aanbevolen.^{3,4} In Nederland wordt een groot deel van de longpatiënten gediagnosticeerd en behandeld door de huisarts, en daarom wordt spirometrie veel gebruikt in de huisartsenpraktijk.⁵

De kwaliteit van spirometrie wordt beoordeeld aan de hand van de accepteerbaarheids- en herhaalbaarheids-criteria opgesteld door de American Thoracic Society (ATS) en de European Respiratory Society (ERS).⁶ Uit een Nederlands onderzoek blijkt dat slechts 31,9% van de spirometrietests in de Nederlandse huisartsenpraktijken aan de ATS/ERS-criteria voldeed, terwijl longfunctieanalisten 83,7% van deze tests op grond van hun eigen expertise als 'klinisch bruikbaar' beoordeelden.⁷

Wij evalueerden of de kwaliteit van spirometrie in de Nederlandse huisartsenpraktijk voldoende is voor dagelijks klinisch gebruik. Daarvoor keken we naar de overeenkomst in de diagnoses (geformuleerd op basis van de spirometrietest en klinische patiëntgegevens) van huisartsen

en longartsen. Zo'n beoordeling van de kwaliteit kan voor de dagelijkse praktijk relevanter zijn dan de criteria die de kwaliteit van de test zelf beoordelen. Ook onderzochten we de proportie tests van goede kwaliteit en klinisch bruikbare tests.

METHODE

Onderzoeksontwerp, setting en deelnemers

We voerden dit observationele onderzoek uit in de regio Zwolle in huisartsenpraktijken die in het bezit waren van een spirometer. Alle deelnemende praktijken vroegen we om 10 opeenvolgende spirometrietests te includeren die werden uitgevoerd als onderdeel van de standaardzorg, ongeacht de uiteindelijk gestelde diagnose. Patiënten van 18 jaar of ouder die een spirometrietest ondergingen vroegen we of ze aan het onderzoek wilden deelnemen. Huisartsen voerden reversibiliteit alleen uit als ze dit nodig vonden.

Gegevensverzameling

Op basis van de steekproefberekening besloten we 150 deelnemers te includeren. Alle deelnemers vulden gevalideerde vragenlijsten in over respiratoire symptomen en patiëntkenmerken.⁸⁻¹⁰ Daarna verrichtte de huisarts of de praktijkondersteuner (POH) de spirometriemeting. De POH selecteerde de 3 beste pre-luchtwegverwijdercurves en, wanneer verricht, de 3 beste post-luchtwegverwijdercurves. De huisarts formuleerde op basis van de spirometrietest en ingevulde vragenlijsten een diagnose en behandeladvies. Ook beoordeelde hij of de spirometrietest van goede, matige of slechte kwaliteit was, en of de test klinisch bruikbaar was (of de kwaliteit voldoende was

Dit is een bewerkte vertaling van: Van de Hei SJ, Flokstra-de Blok BMJ, Baretta HJ, Doornewaard NE, Van der Molen T, Patberg KW, et al. Quality of spirometry and related diagnosis in primary care with a focus on clinical use. *npj Prim Care Respir Med* 2020;30:22.

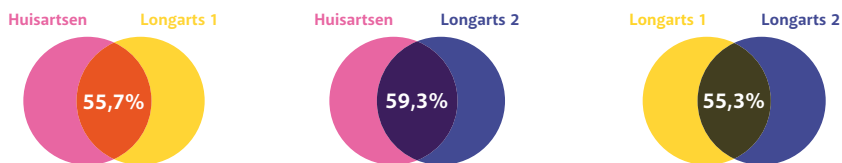
Kwaliteit van spirometrie en klinische bruikbaarheid in de huisartsenpraktijk

Vijftien huisartsen en 16 praktijkondersteuners werkten mee aan het onderzoek, van wie respectievelijk 80% en 93% heeft deelgenomen aan geaccrediteerd spirometrieonderwijs. In totaal werden 165 volwassenen benaderd voor deelname, 149 volwassenen werden geïncludeerd.

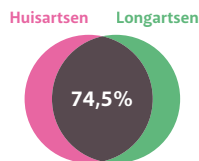


* astma/COPD-overlap, restrictieve longaandoening, onduidelijke diagnose en overige diagnose

Overeenkomst in diagnose



Post-hocanalyse overeenkomst in diagnose*

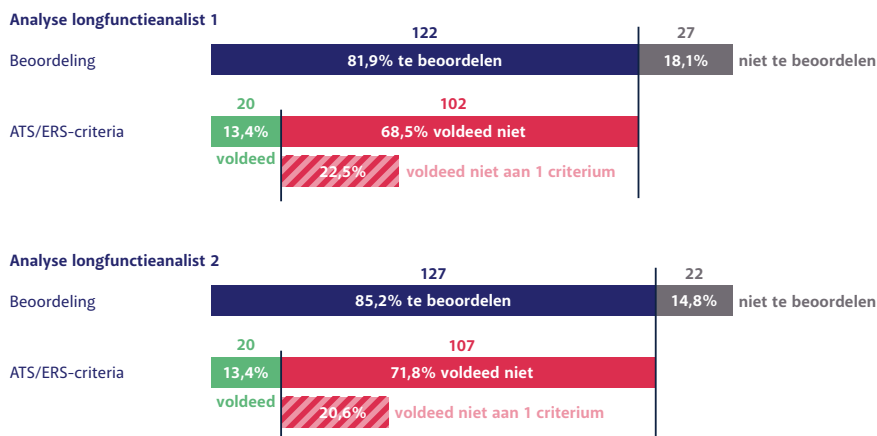


* Alleen de tests meegenomen waarbij de longartsen het eens waren over de diagnose (n = 55)

Kwaliteit van de tests



Analyse van de tests



Het vaakst werd niet voldaan aan de accepteerbaarheidscriteria over de expiratoire piekstrom (PEF):

- Goede start van expiratie
- Duidelijke piek (PEF maximaal krachtig)

voor het nemen van klinische beslissingen). Vervolgens beoordeelden 2 longartsen alle spirometrietests en de ingevulde vragenlijsten op dezelfde punten als de huisartsen. De longartsen hadden geen inzage in elkaars beoordelingen en ook niet in die van de huisartsen. Twee longfunctieanalisten beoordeelden alle spirometrietests op accepteerbaarheid en herhaalbaarheid volgens de ATS/ERS-criteria [kader].⁶ Wanneer het niet mogelijk was om alle ATS/ERS-criteria te beoordelen (bijvoorbeeld bij onjuiste software-instellingen), beschouwden we de test als 'niet te beoordelen'. Alle tests waarover de longfunctieanalisten het oneens waren, werden net zo lang in consensusbijeenkomsten besproken tot er consensus was.

Analyse

De primaire uitkomst van dit onderzoek is de overeenkomst in de geformuleerde diagnose van huisartsen en longartsen. De spirometrietests van slechte kwaliteit en de tests die als klinisch onbruikbaar werden beoordeeld, namen we niet mee in de analyse. Overeenkomst in diagnose geven we weer in 'geobserveerde overeenkomst' (het aantal tests met overeenkomst in diagnose gedeeld door het totale aantal tests) en Cohen's kappa (κ). $\kappa > 0,81$ beschouwen we als een goede overeenkomst, $\kappa > 0,61$ als een voldoende tot goede overeenkomst, $\kappa > 0,41$ als een redelijke overeenkomst en $\kappa > 0,21$ als een matige overeenkomst. In een post-hocanalyse bekeken we de overeenkomst van de diagnoses van huisartsen en longartsen, waarin we alleen de tests meenamen waarbij de longartsen het over de diagnose eens waren. Andere uitkomsten zijn de overeenstemming tussen de beoordelaars, de proportie spirometrietests die aan ATS/ERS-criteria voldoen en de proportie tests die klinisch bruikbaar en kwalitatief goed waren.

RESULTATEN

Van juni 2017 tot september 2018 verzamelden we gegevens

WAT IS BEKEND?

- De kwaliteit van spirometrie wordt beoordeeld aan de hand van de ATS/ERS-criteria.
- In de Nederlandse huisartsenpraktijk voldoet ongeveer een derde van de spirometrietests aan deze criteria, maar wordt het merendeel als klinisch bruikbaar beschouwd.

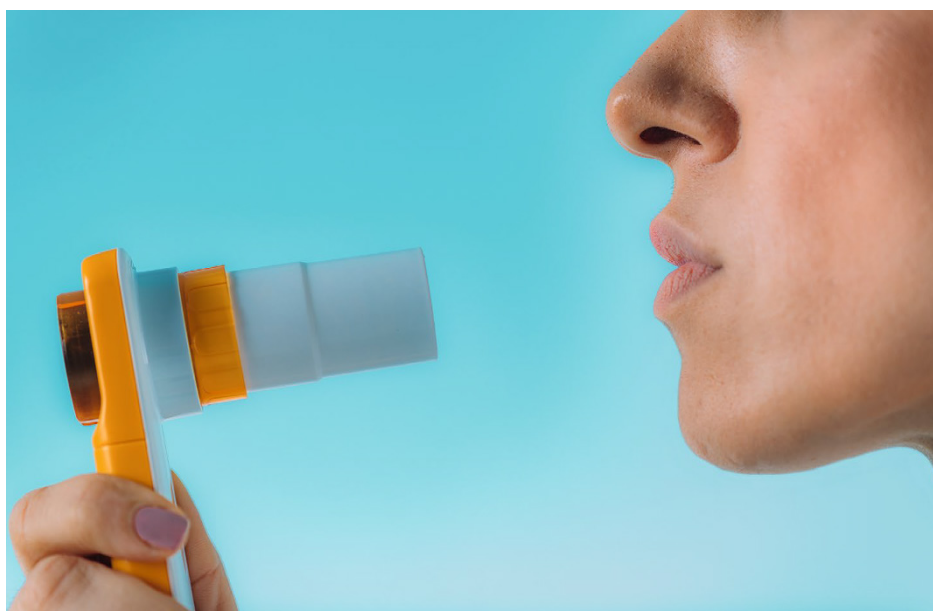
WAT IS NIEUW?

- Er is weinig overeenkomst in diagnose [gebaseerd op spirometrie en patiëntgegevens] tussen huisartsen en longartsen, en tussen longartsen onderling.
- Er was een opmerkelijk groot verschil tussen spirometrietests die longartsen en huisartsen als klinisch bruikbaar beoordeelden [$> 87\%$] en het aantal tests dat aan de ATS/ERS-criteria voldeed [13%].

in 13 huisartsenpraktijken. De onderzoekspopulatie bestond voor 51,7% uit mannen, de gemiddelde leeftijd was 56,8 jaar en het gemiddelde FEV1 (% van voorspeld) was 79,1%. Het merendeel van de deelnemers was nog nooit bij een longarts geweest (65%) en ongeveer 20% rookte ten tijde van deelname. Bij 60% van de deelnemers werd de reversibiliteit getest. De overeenkomst in diagnoses tussen huisartsen en longarts 1 uitgedrukt in kappa (κ) was 0,392. Die tussen huisartsen en longarts 2 was 0,438 en tussen de longartsen onderling 0,382. De post-hocanalyse liet een betere overeenkomst zien met een kappa van 0,627 ($n = 55$). Zie [infographic] voor de resultaten.

BESCHOUWING

De overeenkomst in de diagnose gebaseerd op spirometrie en patiëntgegevens van de huisartsen en die van de longartsen is matig, net als de overeenkomst tussen de diagnoses van de



Er is weinig overeenkomst tussen de diagnose van huisartsen en die van longartsen bij astma/COPD. Ook tussen longartsen onderling verschilt de diagnose.

Foto: Shutterstock

KADER DE CRITERIA DIE GEBRUIKT WERDEN ALS BEORDELINGSINSTRUMENT VOOR DE LONGFUNCTIE- ANALISTEN⁶

Accepteerbaarheidscriteria*

Flow-volumecurve

- Goede start van expiratie – Peak Expiratory Flow (PEF) snel genoeg bereikt [geëxtrapoleerd volume < 5% of FVC < 0,15 L]
- Goede piek [PEF maximaal krachtig]
- Glad verloop expiratie [geen hoest gedurende de eerste seconde]
- Goede expiratie [onafgeknepen expiratie, geen twijfel in aanzet, glottis gaat niet dicht]
- Niet opnieuw ingeademd tijdens de manoeuvre

Volume-tijdcurve

- Duur expiratie ≥ 6 sec[†]
- Plateau [≥ 1 sec < 0,025 L verandering in volume]

Herhaalbaarheidscriteria*

- Verschil tussen de 2 hoogste FVC-waarden < 0,150 L
- Verschil tussen de 2 hoogste FEV1-waarden < 0,150 L

* Er moeten minimaal 3 accepteerbare curves worden verkregen.

† De duur is niet als criterium voor 3 accepteerbare curves gebruikt.

‡ Herhaalbaarheidscriteria worden pas toegepast nadat 3 accepteerbare curves zijn verkregen.

longartsen onderling. Een opmerkelijk groot verschil vonden we tussen spirometrietests die longartsen en huisartsen als klinisch bruikbaar beoordeelden (87%) en het aantal tests dat aan de ATS/ERS-criteria voldeed (13%).

Interpretatie

Twee eerdere onderzoeken evalueerden overeenkomst in diagnose, maar deze zijn niet te vergelijken met het onze. Het ene onderzoek includeerde immers alleen patiënten met een vermoeden van COPD, en het andere legde gestandaardiseerde casusbeschrijvingen voor aan huisartsen, die vervolgens vergeleken werden met een 'gouden standaard' (consensus binnen een expertpanel).^{11,12}

Onze bevinding dat een groot deel van de tests klinisch bruikbaar is komt overeen met een eerder Nederlands onderzoek.⁷ In dat onderzoek stelden longfunctieanalisten de bruikbaarheid vast, terwijl in ons onderzoek longartsen de bruikbaarheid vaststelden. Onze schatting lijkt representatiever voor de klinische bruikbaarheid, omdat longartsen dagelijks spirometrietests gebruiken bij het bepalen van hun beleid.

In eerdere eerstelijns-onderzoeken werd vaker voldaan aan de ATS/ERS-criteria dan in ons onderzoek (32-40% versus 13%).^{7,13,14} Hiervoor kunnen we verschillende verklaringen geven: samenwerkingsverbanden met ziekenhuizen waarbij huisartsen geregeld spirometrietraining en support kregen, in-

tensieve spirometrietraining bij deelname aan het onderzoek, en de vereiste van 2 accepteerbare curves (in plaats van 3).^{7,14} We vonden maar 1 onderzoek waarin spirometrietests in de tweede lijn op ATS/ERS-criteria beoordeeld werden en daaruit bleek dat 41% van de tests aan de criteria voldeed.¹⁵ Voor veel zorggroepen in Nederland behoort het volgen van de CASPIR-cursus tot de vereisten om spirometrie in de huisartsenpraktijk te mogen uitvoeren. Ons onderzoek laat zien dat de gewenste kwaliteit van spirometrietests op basis van de strikte ATS/ERS-kwaliteitseisen vaak niet gehaald wordt. Deze kwaliteitseisen zijn inmiddels aangepast.¹⁶ Naar verwachting zal zonder CASPIR of andere vergelijkbare cursussen niet alleen de kwaliteit, maar ook de klinische bruikbaarheid afnemen. Hiervoor is tot op heden geen wetenschappelijke onderbouwing.

Diagnose 'onduidelijk'

In totaal beschouwden de longartsen de diagnose bij 1 op de 3 spirometrietests als 'onduidelijk'. Maar longarts 1 en longarts 2 vonden respectievelijk 75% en 100% van de tests met de diagnose 'onduidelijk' wel klinisch bruikbaar. We denken niet dat de longartsen de diagnose 'onduidelijk' minder vaak zouden kiezen wanneer ze de patiënten 'live' gezien hadden. Immers eerder onderzoek, dat gebruikmaakte van een vergelijkbare vragenlijst, vond een goede overeenkomst tussen 'live' en 'papier' beoordeling (κ 0,82).¹⁷ De frequente 'onduidelijke' diagnose laat vooral zien hoe lastig het kan zijn om een diagnose te stellen met de beperkte diagnostische faciliteiten in de huisartsenpraktijk. Dit zou voor een derde van de patiënten een verwijzing voor aanvullend onderzoek rechtvaardigen. In de herziene NHG-Standaard Astma wordt gesuggereerd om te verwijzen bij diagnostische problemen. De anderhalvelijnszorg zou ook meer gebruik kunnen maken van provocatietests (bijvoorbeeld in huisartslaboratoria).¹⁸ Op dit moment loopt er een onderzoek dat een diagnostisch algoritme gebaseerd op *machine learning* moet valideren op basis van de gegevens van ons onderzoek.¹⁹ Zulke algoritmes, die klinische gegevens en spirometriegegevens combineren, zouden de kwaliteit van de diagnostiek kunnen vergroten en daarmee de variatie tussen zorgverleners kunnen verkleinen.

Overeenstemming tussen de beoordelaars

De bedoeling van ons onderzoek was dat de diagnose die de longarts formuleerde de gouden standaard zou zijn. We kozen ervoor om de beoordelingen te laten doen door 2 longartsen en 2 longfunctieanalisten, omdat we er zeker van wilden zijn dat de gouden standaard juist vertegenwoordigd werd. De overeenstemming tussen de longfunctieanalisten was goed (κ 0,67 voor de consensusbijeenkomsten, κ 0,81 na de consensusbijeenkomsten). De overeenstemming tussen de longartsen was echter matig (κ 0,38), ook al beschouwden ze de meeste tests als klinisch bruikbaar. De overeenkomst tussen de longartsen was het grootst wanneer we alleen naar de diagnose COPD (κ 0,60) keken. Toch is ook deze overeenkomst lager dan verwacht, omdat COPD een ziekte is die gedefinieerd wordt op basis van spirometriebevindingen. Van astma is

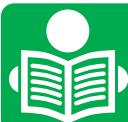
bekend dat er meer variatie is tussen artsen, omdat het een diagnose is die behalve op spirometrieresultaten ook gebaseerd is op de anamnese en ander aanvullend onderzoek.²⁰

CONCLUSIE

In dit onderzoek vonden we weinig overeenkomst tussen de diagnose (gebaseerd op spirometrie en een uitgebreide inventarisatie van patiëntkenmerken en klachten) van huisartsen en die van longartsen, en tussen die van longartsen onderling. Wanneer we alleen de tests meenamen waarbij de longartsen het eens waren over de diagnose, was de overeenkomst in diagnose tussen huisartsen en longartsen een stuk groter. Slechts een beperkt aantal tests voldeed aan de ATS/ERS-criteria, maar de longartsen en huisartsen vonden het merendeel van de tests wel klinisch bruikbaar. Dit kan erop wijzen dat voldoen aan de spirometriekwaliteitscriteria geen vereiste hoeft te zijn om een diagnose te stellen op basis van spirometrie in combinatie met (voldoende) patiëntgegevens. Het is nog onduidelijk of de overeenkomst in diagnose, en daarmee de klinische besluitvorming, beter is wanneer spirometrietests wel aan de spirometriekwaliteitscriteria voldoen. ■

LITERATUUR

1. Volksgezondheidszorg.info. Prevalentie astma in de huisartsenpraktijk. Utrecht: Nivel, 2020. Geraadpleegd op 4 januari 2021.
2. Volksgezondheidszorg.info. Prevalentie en nieuwe gevallen COPD in huisartsenpraktijk. Utrecht: NIVEL, 2020. Geraadpleegd op 4 januari 2021.
3. Global Initiative for Chronic Obstructive Lung Disease. Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease. Fontana: Global Initiative for Chronic Obstructive Lung Disease, 2020. Geraadpleegd op 13 oktober 2020.
4. Global Initiative for Asthma. Global strategy for asthma management and prevention. Fontana: Global Initiative for Asthma, 2020. Geraadpleegd op 13 oktober 2020.
5. Schellekens D, Poels P, Pellegrino A, Cretier R, Smeele I, Schermer T. Spirometrie in de Nederlandse huisartsenpraktijk. *Huisarts Wet* 2008;51:434-9.
6. Miller MR, Hankinson J, Brusasco V, Burgos F, Casaburi R, Coates A, et al. Standardisation of spirometry. *Eur Respir J* 2005;26:319-38.
7. Landman M, Gilissen T, Grootens-Stekelenburg J, Akkermans R, Schermer T. Kwaliteit van spirometrie in de eerste lijn. *Huisarts Wet* 2011;54:536-42.
8. Bestall JC, Paul EA, Garrod R, Garnham R, Jones P, Wedzicha J. Usefulness of the Medical Research Council (MRC) dyspnoea scale as a measure of disability in patients with chronic obstructive pulmonary disease. *Thorax* 1999;54:581-6.
9. Juniper EF, O'Byrne P, Guyatt G, Ferrie PJ, King DR. Development and validation of a questionnaire to measure asthma control. *Eur Respir J* 1999;14:902-7.
10. Van der Molen T, Willemse BWM, Schokker S, Ten Hacken NHT, Postma DS, Juniper EF. Development, validity and responsiveness of the clinical COPD questionnaire. *Health Qual Life Outcomes* 2003;1:13.
11. White P, Wong W, Fleming T, Gray B. Primary care spirometry: test quality and the feasibility and usefulness of specialist reporting. *Br J Gen Pract* 2007;57:701-5.
12. Chavannes N, Schermer T, Akkermans R, Jacobs JE, Van de Graaf G, Bollen R, et al. Impact of spirometry on GPs' diagnostic differentiation and decision-making. *Respir Med* 2004;98:1124-30.
13. Schermer TRJ, Crockett AJ, Poels PJP, Van Dijke JJ, Akkermans RP, Vlek HF, et al. Quality of routine spirometry tests in Dutch general practices. *Br J Gen Pract* 2009;59:921-6.
14. Schneider A, Gindner L, Tilemann L, Schermer T, Dinant G-J, Meyer FJ, et al. Diagnostic accuracy of spirometry in primary care. *BMC Pulm Med* 2009;9:31.
15. Spiegelaar J, Steenbruggen I, Meulenbelt J, Grotjohan H. Does feedback improve compliance to the ATS/ERS 2005 acceptability criteria in our lung function laboratory? *Eur Respir J* 2008;32:552s.
16. Graham BL, Steenbruggen I, Barjaktarevic IZ, Barjaktarevic IZ, Cooper BG, Hall GL, et al. Standardization of spirometry 2019 update. An official American Thoracic Society and European Respiratory Society technical statement. *Am J Respir Crit Care Med* 2019;200:E70-88.
17. Lucas A, Smeenk FJWM, Smeele I, Brouwer T, Van Schayck O. The validity of diagnostic support of an asthma/COPD service in primary care. *Br J Gen Pract* 2007;57:892-6.
18. Bins JE, Metting EI, Muilwijk-Kroes JB, Kocks JWH, In 't Veen JCCM. The use of a direct bronchial challenge test in primary care to diagnose asthma. *npj Prim Care Respir Med* 2020;30.
19. Kaplan A, Cao H, Fitzgerald JM, Yang E, Iannotti N, Kocks JWH, et al. Asthma/COPD Differentiation Classification (AC/DC): machine learning to aid physicians in diagnosing asthma, COPD and Asthma-COPD Overlap (ACO) *Am J Respir Crit Care Med* 2020;201:A6285.
20. Aaron SD, Boulet LP, Reddel HK, Gershon AS. Underdiagnosis and overdiagnosis of asthma. *Am J Respir Crit Care Med* 2018;198:1012-20.



Lees ook: Van Nederveen-Bendien SA, et al. Is goede astma-/COPD-zorg mogelijk met niet-optimale spirometrie? *Huisarts Wet* 2021;64:DOI:10.1007/s12445-021-1112-4.

En: Honkoop PJ, et al. Het gebruik van fractional exhaled nitric oxide [FeNO] bij astmamonitoring. *Huisarts Wet* 2021;64:DOI:10.1007/s12445-021-1124-0.

Van de Hei SJ, Flokstra-de Blok BMJ, Baretta HJ, Doornwaard NE, Van der Molen T, Patberg KW, Ruberg ECM, Schermer TRJ, Steenbruggen I, Van den Berg JWK, Kocks JWH. Klinische bruikbaarheid en kwaliteit van spirometrie. *Huisarts Wet* 2021;64:DOI:10.1007/s12445-021-1123-1.

UMC Groningen, afdeling Huisartsgeneeskunde en Ouderengeneeskunde, Groningen: S.J. van de Hei, aioto, s.j.van.de.hei@umcg.nl; H.J. Baretta, aios; N.E. Doornwaard, huisarts; prof. dr. T. van der Molen, huisarts [n.p.] en emeritus hoogleraar Huisartsgeneeskunde. UMC Groningen Research Institute for Asthma and COPD [GRIAC], Groningen: dr. B.M.J. Flokstra-de Blok, onderzoeker; dr. J.W.H. Kocks, huisarts. Isala Ziekenhuis, afdeling Longgeneeskunde, Zwolle: dr. K.W. Patberg, longarts; dr. J.W.K. van der Berg, longarts; E.C.M. Ruberg, longfunctie-analist; I. Steenbruggen, longfunctie-analist. Radboud UMC, afdeling Eerstelijngeneeskunde, Nijmegen: dr. T.R.J. Schermer, epidemioloog. Mogelijke belangenverstremgeling: dit onderzoek is mogelijk gemaakt door een 'unrestricted grant' van Chiesi B.V.