

# Het nut van niet-invasieve testen voor de diagnose van *Helicobacter pylori*?

## Referentie

Best LM, Takwoingi Y, Siddique S, et al. Non-invasive diagnostic tests for *Helicobacter pylori* infection. *Cochrane Database Syst Rev* 2018, Issue 3. DOI: 10.1002/14651858.CD012080.pub2

## Duiding

Louis Ferrant, Vakgroep Eerstelijns en Interdisciplinaire Zorg, Centrum voor Huisartsgeneeskunde, Universiteit Antwerpen

## Klinische vraag

Wat is de diagnostische accuraatheid van niet-invasieve diagnostische testen om *Helicobacter pylori* op te sporen bij symptomatische of asymptomatische patiënten?

## Achtergrond

Uit verschillende studies is gebleken dat *Helicobacter pylori* gecorreleerd is met een aantal kwaadaardige aandoeningen zoals maagkanker, premaligne maagletsels, maaglymfoom (1) colonkanker (2,3), alsook goedaardige aandoeningen, zoals maagulcus (4), dyspepsie (5,6), recidiverende maagbloeding (7), onverklaarde ferriprievie anemie (8), idiopathische trombocytopenische purpura (9) en colorectaal adenoom (2). Het opsporen en eradiceren van *Helicobacter pylori* wordt daarom aanbevolen voor bepaalde risicogroepen. Niet-invasieve testen zouden enerzijds de drempel voor opsporing kunnen verlagen. Anderzijds is het, om overbehandeling te vermijden, ook belangrijk om het aantal vals-positieven zo klein mogelijk te houden. Vandaar het belang om de diagnostische accuraatheid van niet-invasieve testen voor de diagnose van *H. pylori* bij zowel symptomatische als asymptomatische personen te onderzoeken.

## Samenvatting

### Methodologie

Systematische review en meta-analyse

### Geraadpleegde bronnen

- MEDLINE, Embase, Science Citation Index Expanded, National Institute for Health Research Health Technology Assessment (tot maart 2016)
- gerelateerde artikels en referentielijsten met/van gevonden artikels (tot december 2016)
- geen restrictie voor taal of publicatiestatus.

### Geselecteerde studies

- inclusiecriteria: zowel prospectieve als retrospectieve studies die de diagnostische accuraatheid van een initieel uitgevoerde <sup>13</sup>C-ureum ademtest, <sup>14</sup>C-ureum ademtest, serologische test en/of stoelgangantigeentest (indextesten) onderzochten versus endoscopische biopsie met histologische kleuring teneinde *H. pylori* op te sporen bij symptomatische en asymptomatische personen waarbij men eradicatietherapie overwoog
- exclusiecriteria: case-control studies, onvoldoende gegevens over diagnostische accuraatheid (zoals aantal terecht en vals positieven/negatieven), patiënten met maagbloeding
- uit 23 896 gevonden referenties selecteerde men uiteindelijk 101 studies.

### Bestudeerde populatie

- in totaal 11 003 deelnemers waarvan 5 839 (53,1%) met een *H. pylori*-infectie (mediane prevalentie 53,7% met IQR 42,0% tot 66,5%)
- 14 studies met alleen kinderen en 87 studies met alleen of vooral volwassenen
- in 58 studies includeerde men alleen symptomatische patiënten (abdominale pijn of dyspepsie) en in 53 studies sloot men patiënten met recent gebruik van protonpompinhibitoren of antibiotica uit.

### Uitkomstmeting

- primaire uitkomstmaat: diagnostische accuraatheid van de <sup>13</sup>C-ureum ademtest, de <sup>14</sup>C-ureum ademtest, de serologische test en de stoelgangantigeentest (indextesten) versus endoscopische biopsie met histologische kleuring (referentietest) om *H. pylori* op te sporen
- secundaire uitkomstmaat: invloed van potentiële bronnen van heterogeniteit zoals type referentietest, risico van bias, publicatiestatus, prospectieve versus retrospectieve studies, symptomatische versus asymptomatische deelnemers, recent of huidig gebruik van protonpompinhibitoren of antibiotica, verschillende subtypes van testen, tijdsinterval tussen uitvoering van index- en referentietest
- verschil in diagnostische accuraatheid visueel weergegeven in **sROC-curve**n en numeriek uitgedrukt in **diagnostische odds ratio's**.

### Resultaten

- van de primaire uitkomstmaat:
  - er is een statistisch significant verschil in diagnostische accuraatheid tussen de <sup>13</sup>C-ureum ademtest, de <sup>14</sup>C-ureum ademtest, de serologische test en de stoelgangantigeentest ( $p=0,024$ ) ( $N=99$  studies)
  - diagnostische odds ratio's (DOR) voor de <sup>13</sup>C-ureum ademtest, de <sup>14</sup>C-ureum ademtest, de serologische test en de stoelgangantigeentest bedragen respectievelijk 153 (95% BI van 73,7 tot 316), 105 (95% BI van 74,0 tot 150), 47,4 (95% BI van 25,5 tot 88,1) en 45 (95% BI van 24,2 tot 84,1)
  - met een specificiteit vastgelegd op 0,90 (mediane waarde van alle studies) bedraagt de sensitiviteit 0,94 (95% BI van 0,89 tot 0,97) voor de <sup>13</sup>C-ureum ademtest, 0,92 (95% BI van 0,89 tot 0,94) voor de <sup>14</sup>C-ureum ademtest, 0,84 (95% BI van 0,74 tot 0,91) voor de serologische test en 0,83 (95% BI van 0,73 tot 0,90) voor de stoelgangantigeentest
  - met deze specificiteit van 0,90 en een prevalentie voor *H. pylori* van 53,7% (mediane waarde van alle studies) zou men op 1 000 geteste personen 46 vals-positieven hebben alsook 30 (95% BI van 15 tot 58) vals negatieven met de <sup>13</sup>C-ureum ademtest, 42 (95% BI van 30 tot 58) met de <sup>14</sup>C-ureum ademtest, 86 (95% BI van 50 tot 140) met de serologische test en 89 (95% BI van 52 tot 146) met de stoelgangantigeentest
  - op basis van alleen head-to-head-studies kon men voor de <sup>13</sup>C-ureum ademtest versus de serologische test ( $N=7$  studies) en voor de <sup>14</sup>C-ureum ademtest versus de stoelgangantigeentest ( $N=7$  studies) geen statistisch significant verschil in DOR berekenen
- van de secundaire uitkomstmaten: te weinig informatie om de invloed van deze factoren op de diagnostische accuraatheid te bepalen.

### Besluit van de auteurs

De auteurs besluiten dat bij personen zonder voorgeschiedenis van gastrectomie en zonder recent gebruik van antibiotica of protonpompinhibitoren, ureum ademtesten een hoge diagnostische accuraatheid hadden, terwijl de serologische test en de stoelgangantigeentest minder accuraat waren voor de diagnose van een *Helicobacter pylori*-infectie. Dit besluit is gebaseerd op een indirecte vergelijking tussen de testen (met een potentieel risico van bias door confounding) omdat evidentie uit directe vergelijkingen beperkt of niet beschikbaar is. De gebruikte afkapwaarden voor deze testen waren zeer variabel en het was onmogelijk om specifieke afkapwaarden te bepalen die nuttig zouden kunnen

zijn in de klinische praktijk. Er zijn nieuwe vergelijkende studies met hoge methodologische kwaliteit nodig om meer betrouwbare evidentie te bekomen over de relatieve accuraatheid van deze testen. Het zou moeten gaan om prospectieve studies met een representatief spectrum van deelnemers en een duidelijke rapportage teneinde het risico van bias zo laag mogelijk te houden. Het belangrijkste is echter nog dat studies op voorhand de gebruikte afkapwaarden duidelijk vermelden en vermijden om onterecht deelnemers uit te sluiten.

#### **Financiering van de studie**

Financiering niet vermeld.

#### **Belangenconflicten van de auteurs**

Geen vermelding van belangenconflicten.

## **Bespreking**

### **Methodologische beschouwingen**

Deze systematische review vertrok van een adequate onderzoeksvraag en zocht in 4 verschillende databanken naar bruikbare literatuur. De onderzoekers zochten echter niet specifiek naar ‘grijze literatuur’. De selectie van studies op basis van vooraf vastgelegde in-en exclusiecriteria, alsook de data-extractie en het bepalen van de methodologische kwaliteit van alle weerhouden studies gebeurde telkens door 2 onafhankelijke auteurs. Geen enkele studie had een hoge methodologische kwaliteit (=laag risico van bias voor alle domeinen van de Cochrane risk of bias tool). Voor 15 studies was er een hoog risico van selectiebias omdat ze patiënten niet op een consecutieve of ad random manier hadden geïnccludeerd. Voor alle studies was het onduidelijk of de interpretatie van de indextest gebeurde zonder kennis van het resultaat van de referentietest en of er op voorhand een afkapwaarde voor de indextest was bepaald. Ook wat betreft de referentietest was het risico van bias in 27 studies onduidelijk omdat het niet zeker was of men het resultaat van deze test had geïnterpreteerd zonder het resultaat van de indextest te kennen. In 72 studies was dit risico zelfs hoog omdat men niet de vooropgestelde immunohistochemische kleuring had gebruikt. Voor 74 studies was het interval tussen de uitvoering van de index- en de referentietest onduidelijk of was het niet duidelijk in hoeverre men alle patiënten includeerde. In 20 studies sloot men patiënten uit zonder rapportering van het resultaat van de referentietest.

Door de hoge heterogeniteit in afkapwaarden was het niet mogelijk om de sensitiviteit en de specificiteit voor de verschillende testen te poolen in een **ROC-curve**. We kunnen dus niet achterhalen in welke richting de diagnostische accuraatheid van een test zal evolueren afhankelijk van de gekozen afkapwaarde.

Tot slot is het belangrijk te weten dat slechts weinig niet-invasieve testen rechtstreeks met elkaar vergeleken zijn.

### **Interpretatie van de resultaten**

De diagnostische accuraatheid van de <sup>13</sup>C-ureum ademtest en de <sup>14</sup>C-ureum ademtest was hoger dan van de serologische test en de stoelgangantigeentest. Wanneer men de aanwezigheid van *Helicobacter pylori* wil opsporen, kan men dus kiezen tussen een voor de patiënt meer belastende en duurdere <sup>13</sup>C- of <sup>14</sup>C-ureum ademtest met hoge diagnostische accuraatheid tegenover de goedkopere en gemakkelijker uitvoerbare serologische test en stoelgangantigeentest met lagere diagnostische accuraatheid.

Na het kiezen van een test is ook de keuze van de afkapwaarde voor een bepaalde test belangrijk. Enerzijds zou men kunnen kiezen voor een afkapwaarde met hogere sensitiviteit maar met lagere specificiteit en dus meer vals-positieven waardoor meer gastroscopieën met biopsie nodig zullen zijn om de diagnose te bevestigen en overbehandeling met eradicatietherapie te vermijden. Anderzijds zou men kunnen kiezen voor een afkapwaarde met hogere specificiteit maar met lagere sensitiviteit en dus meer vals-negatieven waardoor veel patiënten ten onrechte een eradicatietherapie ontzegd zullen worden. Omdat eradicatietherapie niet zeer belastend is voor de patiënt kan men de keuze voor

de meest gevoelige afkapwaarde verantwoorden. Echter is het ook zo dat deze keuze het onnodige gebruik van antibiotica en dus ook het risico op resistentie kan doen toenemen, wat niet alleen voor het individu nu maar ook voor latere generaties gevolgen kan hebben. In dat geval zou een meer specifieke afkapwaarde te verkiezen zijn. De heterogeniteit van de gebruikte afkapwaarden in de verschillende studies laat niet toe om deze afweging voor de verschillende testen te maken. Uit het feit dat de onderzoekers alleen de sensitiviteit van de verschillende testen vergelijken, kunnen we echter wel afleiden dat ze in de eerste plaats op zoek zijn naar de meest gevoelige test. Waarschijnlijk is dat terecht wegens de hoge mediane prevalentie die ze voor alle studies samen konden vaststellen. Deze bevinding strookt dan weer volledig met het feit dat slechts één studie alleen asymptomatische patiënten includeerde. Tot slot moet nog vermeld worden dat de meeste studies patiënten met een gastrectomie of met gebruik van PPI's of antibiotica in de voorgeschiedenis excludeerden.

## Besluit van Minerva

Deze systematische review van 101 diagnostische studies met matig tot hoog risico van bias toont aan dat ureum ademtesten een hogere diagnostische accuraatheid hebben dan serologische testen en stoelgangantigeentesten om een *Helicobacter pylori*-infectie op te sporen bij symptomatische patiënten zonder gastrectomie in de voorgeschiedenis en zonder recent antibioticagebruik of inname van PPI's. Door de hoge heterogeniteit is het niet mogelijk om voor deze testen de meeste gevoelige of meest specifieke afkapwaarde te bepalen.

## Voor de praktijk

Het niet-invasief opsporen van *Helicobacter pylori* is aanbevolen bij patiënten afkomstig uit mediterrane landen, het Midden of Verre Oosten, Afrika, Midden- en Zuid-Amerika en een eerste episode van maagklachten (wegens een prevalentie tot 75% in deze groep); bij persisterende of recidiverende maagklachten zonder alarmsymptomen; na elke *H. pylori*-eradicatorbehandeling; bij noodzakelijk NSAID-gebruik indien men ooit een ulcus diagnosticeerde en de *H. pylori*-status onbekend is (10). Men kan hiervoor een ureum ademtest, serologische test of stoelgangantigeentest aanbevelen (11). De NHG-standaard beveelt de stoelgangantigeentest of een <sup>13</sup>C-ureum ademtest aan (10). De <sup>13</sup>C-ureum ademtest heeft de voorkeur boven de <sup>14</sup>C-ureum ademtest omdat de <sup>14</sup>C-ureum ademtest radioactief en duurder is. Voorafgaand aan deze testen mag gedurende 2 weken geen PPI en gedurende 4 weken geen antibioticum gebruikt worden. Een PPI kan eventueel worden vervangen door een H<sub>2</sub>-receptorantagonist, die één dag voor de test moet worden gestopt of door een antacidum, dat niet gestopt hoeft te worden. Serologische testen worden niet aanbevolen tenzij ureum adem- of stoelgangantigeentesten niet beschikbaar zijn of als een patiënt het PPI-gebruik niet kan stoppen. Serologische sneltesten (via vingerprik) worden ontraden. Ondanks de lagere gevoeligheid van de stoelgangantigeentest, wat nogmaals werd aangetoond in de hoger besproken systematische review, is het verdedigbaar om deze test omwille van zijn lage kostprijs en zijn gebruiksgemak aan te bevelen. In afwachting van eventuele nieuwere methodologisch correct uitgevoerde diagnostische studies is dit echter een voorlopig besluit na afweging van voor- en nadelen van de beschikbare niet-invasieve testen om *Helicobacter pylori* op te sporen. Er is geen wetenschappelijk bewijs om een asymptomatische populatie te screenen op en te behandelen voor *H. pylori* (11), wat ook door deze systematische review bevestigd wordt.

**Referenties** zie website