

# Medizinischer und gesundheitsökonomischer Nutzen der Untersuchung auf *Helicobacter pylori*-Besiedlung mittels <sup>13</sup>C-Harnstoff-Atemtest in der Primärdiagnostik im Vergleich zu invasiven und nichtinvasiven diagnostischen Verfahren

## Efficacy and cost-effectiveness of the <sup>13</sup>C-urea breath test as the primary diagnostic investigation for the detection of *Helicobacter pylori* infection compared to invasive and non-invasive diagnostic tests

### Abstract

#### Background

*Helicobacter pylori* (*H. pylori*) is one of the most common bacterial infections in humans. There is a risk factor for gastric or duodenal ulcers, gastric cancer and MALT (Mucosa Associated Lymphoid Tissue)-Lymphomas. There are several invasive and non-invasive methods available for the diagnosis of *H. pylori*. The <sup>13</sup>C-urea breath test is a non-invasive method recommended for monitoring *H. pylori* eradication therapy. However, this test is not yet used for primary assessment of *H. pylori* in Germany.

#### Objectives

What are the clinical and health economic benefits of the <sup>13</sup>C-urea breath test in the primary assessment of *H. pylori* compared to other invasive and non-invasive methods?

#### Methods

A systematic literature search including a hand search was performed for studies investigating test criteria and cost-effectiveness of the <sup>13</sup>C-urea breath test in comparison to other methods used in the primary assessment of *H. pylori*. Only studies that directly compared the <sup>13</sup>C-urea breath test to other *H. pylori*-tests were included. For the medical part, biopsy-based tests were used as the gold standard.

#### Results

30 medical studies are included. Compared to the immunoglobulin G (IgG) test, the sensitivity of the <sup>13</sup>C-urea breath test is higher in twelve studies, lower in six studies and one study reports no differences. The specificity is higher in 13 studies, lower in three studies and two studies report no differences. Compared to the stool antigen test, the sensitivity of the <sup>13</sup>C-urea breath test is higher in nine studies, lower in three studies and one study reports no difference. The specificity is higher in nine studies, lower in two studies and two studies report no differences. Compared to the urease test, the sensitivity of the <sup>13</sup>C-urea breath test is higher in four studies, lower in three studies and four studies report no differences. The specificity is higher in five studies, lower in five

Marc Nocon<sup>1</sup>  
Alexander Kuhlmann<sup>2</sup>  
Andreas Leodolter<sup>3</sup>  
Stephanie Roll<sup>1</sup>  
Christoph Vauth<sup>2</sup>  
Stefan N. Willich<sup>1</sup>  
Wolfgang Greiner<sup>4</sup>

1 Institut für Sozialmedizin,  
Epidemiologie und  
Gesundheitsökonomie,  
Charité - Universitätsmedizin  
Berlin, Berlin, Deutschland

2 Forschungsstelle für  
Gesundheitsökonomie,  
Leibniz Universität Hannover,  
Hannover, Deutschland

3 Ev. Jung-Stilling-  
Krankenhaus, Abt. Innere  
Medizin u. Intensivmedizin-  
konservat., Siegen,  
Deutschland

4 Gesundheitsökonomie und -  
management, Fakultät für  
Gesundheitswissenschaften,  
Universität Bielefeld,  
Bielefeld, Deutschland

studies and one study reports no difference. Compared to histology, the sensitivity of the  $^{13}\text{C}$ -urea breath test is higher in one study and lower in two studies. The specificity is higher in two studies and lower in one study. One study each compares the  $^{13}\text{C}$ -urea breath test to the  $^{14}\text{C}$ -urea breath test and the polymerase chain reaction (PCR) test, respectively, and reports no difference in sensitivity and specificity with the  $^{14}\text{C}$ -urea breath test, and lower sensitivity and higher specificity compared to PCR. The statistical significance of these differences is described for six of the 30 studies.

Nine health economic evaluations are included in the Health Technology Assessment (HTA) report. Among these studies, the test-and-treat strategy using the  $^{13}\text{C}$ -urea breath test is compared to test-and-treat using serology in six analyses and to test and treat using the stool antigen test in three analyses. Thereby, test-and-treat using the breath test is shown to be cost-effective over the serology based strategy in three models and is dominated by a test-and-treat strategy using the stool antigen test in one model. A cost-effectiveness comparison between the urea breath test approach and the empirical antisecretory therapy is carried out in four studies. Of these, two studies report that the strategy using the urea breath test is cost-effective over the empirical antisecretory therapy. In two studies, test-and-treat using the  $^{13}\text{C}$ -urea breath test is compared to the empirical eradication therapy and in five studies to endoscopy-based strategies. The breath test approach dominates endoscopy in two studies and is dominated by this strategy in one study.

## Discussion

All included medical and economic studies are limited to a greater or lesser extent. Additionally, the results of the studies are heterogeneous regarding medical and economic outcomes respectively. Thus, the majority of the medical studies do not report the statistical significance of the differences in sensitivity and specificity. In direct comparisons the  $^{13}\text{C}$ -urea breath test shows higher sensitivity and specificity than the IgG and stool antigen tests. In comparison to the urease test, results for sensitivity are inconsistent, and the specificity is slightly higher for the  $^{13}\text{C}$ -urea breath test. There are not enough results for comparisons between the  $^{13}\text{C}$ -urea breath test and the  $^{14}\text{C}$ -urea breath test, histology and PCR to describe tendencies.

The included economic studies suggest that the test-and-treat strategy using the  $^{13}\text{C}$ -urea breath test is cost-effective compared to test-and-treat using serology as well as empirical antisecretory therapies. Due to a lack of valid studies, it is not possible to assess the breath test approach in comparison to test-and-treat using the stool antigen test and the empirical eradication therapy respectively, regarding the cost-effectiveness. The results of economic analyses comparing test-and-treat using the breath test to endoscopy strategies are too heterogeneous to draw any conclusions. Overall, none of the included economic models is able to completely capture the complexity of managing patients with dyspeptic complaints.

## Conclusions/Recommendations

Based on available medical and economic studies, there is no sufficient evidence to recommend test and-treat using  $^{13}\text{C}$ -urea breath testing for the detection of *H. pylori* infection as the standard procedure for the management of uninvestigated dyspepsia in the German health care system. In addition, it must be considered that the DVGS guidelines of the Deutsche Gesellschaft für Verdauungs- und Stoffwechselkrankheiten

(DVGS) recommend endoscopy based methods for the management of patients with dyspeptic complaints.

## Zusammenfassung

### Hintergrund

*Helicobacter pylori* (*H. pylori*) zählt trotz abnehmender Inzidenz zu den häufigsten bakteriellen Infektionskrankheiten des Menschen. Die Infektion mit *H. pylori* ist ein Risikofaktor für Krankheiten wie gastroduodenale Geschwüre, Magenkarzinomen und MALT (Mucosa Associated Lymphoid Tissue)-Lymphomen. Zur Diagnostik von *H. pylori* stehen verschiedene invasive und nichtinvasive Verfahren zur Verfügung. Der <sup>13</sup>C-Harnstoff-Atemtest wird zur Kontrolle einer Eradikationstherapie empfohlen, kommt in der Primärdiagnostik von *H. pylori* derzeit jedoch nicht standardmäßig in Deutschland zum Einsatz.

### Fragestellung

Welchen medizinischen und gesundheitsökonomischen Nutzen hat die Untersuchung auf *H. pylori*-Besiedlung mittels <sup>13</sup>C-Harnstoff-Atemtest in der Primärdiagnostik im Vergleich zu invasiven und nichtinvasiven diagnostischen Verfahren?

### Methodik

Basierend auf einer systematischen Literaturrecherche in Verbindung mit einer Handsuche werden Studien zur Testgüte und Kosten-Effektivität des <sup>13</sup>C-Harnstoff-Atemtests im Vergleich zu anderen diagnostischen Verfahren zum primären Nachweis von *H. pylori* identifiziert. Es werden nur medizinische Studien eingeschlossen, die den <sup>13</sup>C-Harnstoff-Atemtest direkt mit anderen *H. pylori*-Testverfahren vergleichen. Goldstandard ist eines oder eine Kombination der biopsiebasierten Testverfahren. Für die gesundheitsökonomische Beurteilung werden nur vollständige gesundheitsökonomische Evaluationsstudien einbezogen, bei denen die Kosten-Effektivität des <sup>13</sup>C Harnstoff-Atemtests direkt mit anderen *H. pylori*-Testverfahren verglichen wird.

### Ergebnisse

Es werden 30 medizinische Studien für den vorliegenden Bericht eingeschlossen. Im Vergleich zum Immunglobulin G (IgG)-Test ist die Sensitivität des <sup>13</sup>C-Harnstoff-Atemtests zwölfmal höher, sechsmal niedriger und einmal gleich, und die Spezifität 13-mal höher, dreimal niedriger und zweimal gleich. Im Vergleich zum Stuhl-Antigen-Test ist die Sensitivität des <sup>13</sup>C-Harnstoff-Atemtests neunmal höher, dreimal niedriger und einmal gleich, und die Spezifität neunmal höher, zweimal niedriger und zweimal gleich. Im Vergleich zum Urease-Schnelltest sind die Sensitivität des <sup>13</sup>C-Harnstoff-Atemtests viermal höher, dreimal niedriger und viermal gleich und die Spezifität fünfmal höher, fünfmal niedriger und einmal gleich. Im Vergleich mit der Histologie ist die Sensitivität des <sup>13</sup>C-Harnstoff-Atemtests einmal höher und zweimal niedriger und die Spezifität zweimal höher und einmal niedriger. In je einem Vergleich zeigt sich kein Unterschied zwischen <sup>13</sup>C-Harnstoff-Atemtest und <sup>14</sup>C-Harnstoff-Atemtest, sowie eine niedrigere Sensitivität und höhere Spezifität im Vergleich zur Polymerase-Kettenreaktion (PCR). Inwieweit die beschriebenen Unterschiede statistisch signifikant sind, wird in sechs der 30 Studien angegeben.

Es werden neun gesundheitsökonomische Evaluationen in dem vorliegenden Bericht berücksichtigt. Die Test-and-Treat-Strategie mittels

<sup>13</sup>C-Harnstoff-Atemtest wird in sechs Studien mit einem Test-and-Treat-Verfahren auf Basis der Serologie sowie in drei Studien mit einem Test-and-Treat-Verfahren auf Basis des Stuhl-Antigen-Tests verglichen. Dabei ist das Atemtestverfahren dreimal kosteneffektiv gegenüber der serologischen Methode und wird von der Stuhl-Antigen-Test-Strategie einmal dominiert. Vier Studien beinhalten einen Vergleich der Test-and-Treat-Strategie auf Basis des <sup>13</sup>C-Harnstoff-Atemtests mit einer empirischen antisekretorischen Therapie, wobei sich das Atemtestverfahren zweimal als kosteneffektive Prozedur erweist und zwei Studien einen Vergleich mit einer empirischen Eradikationstherapie. In fünf Studien wird das Test-and-Treat-Verfahren mittels <sup>13</sup>C-Harnstoff-Atemtest einer endoskopiebasierten Strategie gegenübergestellt. Zweimal dominiert die Atemteststrategie die endoskopische Prozedur und einmal wird sie von dieser Strategie dominiert.

## Diskussion

Sowohl die medizinischen als auch die ökonomischen Studien weisen mehr oder minder gravierende Mängel auf und liefern heterogene Ergebnisse. So werden in der Mehrzahl der medizinischen Studien keine Angaben zur statistischen Signifikanz der berichteten Unterschiede zwischen den jeweiligen Testverfahren gemacht. Im direkten Vergleich weist der <sup>13</sup>C-Harnstoff-Atemtest überwiegend eine höhere Testgüte als der IgG und der Stuhl-Antigen-Test auf. Aus den Vergleichen mit dem Urease-Schnelltest lassen sich keine Tendenzen bezüglich der Sensitivität ableiten, wohingegen die Spezifität des <sup>13</sup>C-Harnstoff-Atemtests höher einzuschätzen ist. Für die Vergleiche des <sup>13</sup>C-Harnstoff-Atemtest mit der Histologie, dem <sup>14</sup>C-Harnstoff-Atemtest und der PCR liegen zu wenige Ergebnisse vor.

In der eingeschlossenen ökonomischen Literatur deuten einige Studienergebnisse auf eine Kosten-Effektivität der Test-and-Treat-Strategie mittels <sup>13</sup>C-Harnstoff-Atemtest gegenüber dem Test-and-Treat-Verfahren auf Basis der Serologie und der empirischen antisekretorischen Therapie hin. Um Tendenzen bezüglich der Kosten-Effektivität der Atemteststrategie gegenüber der Test-and-Treat-Strategie mittels Stuhl-Antigen-Test sowie der empirischen Eradikationstherapie abzuleiten, mangelt es an validen Ergebnissen bzw. ökonomischer Evidenz. Die Untersuchungsergebnisse hinsichtlich eines Vergleichs mit endoskopiebasierten Verfahren fallen diesbezüglich zu heterogen aus. Insgesamt kann keines der ökonomischen Modelle der Komplexität des Managements von Patienten mit dyspeptischen Beschwerden gänzlich gerecht werden.

## Schlussfolgerungen/Empfehlungen

Zusammenfassend ist festzuhalten, dass die Studienlage zur medizinischen und ökonomischen Beurteilung des <sup>13</sup>C-Harnstoff-Atemtests im Vergleich zu anderen diagnostischen Methoden nicht ausreichend ist, um den Atemtest als primärdiagnostisches Standardverfahren im Rahmen einer Test-and-Treat-Strategie beim Management von Patienten mit dyspeptischen Beschwerden für die deutsche Versorgungslandschaft insbesondere vor dem Hintergrund der Leitlinien der Deutschen Gesellschaft für Verdauungs- und Stoffwechselkrankheiten (DGVS) anstelle einer endoskopiebasierten Methode zu empfehlen.

**Schlüsselwörter:** Helicobacter pylori, <sup>13</sup>C-Harnstoff-Atemtest, Diagnose, medizinische Effektivität, Kosteneffektivität, Sensitivität, Spezifität

## Executive Summary

### 1. Health political background

Diseases of the gastro-intestinal system are very common in Germany. About 28 million patients are treated for gastro-intestinal illnesses per year. In 2007, gastro-intestinal diseases accounted for 34.6 million prescriptions and a turnover of 1.363 billion Euro. A large number of gastroenterological diseases are caused by *Helicobacter pylori* (*H. pylori*). *H. pylori* infections in humans are one of the most common infections in developed countries, even though the incidence rates are decreasing. According to the Institute for Microbiology and Hygiene at the University of Freiburg, about 20 to 30 million people (30% of the German population) in Germany are infected. The infection of the mucus layer of the stomach with *H. pylori* causes usually a chronic inflammatory reaction, resulting in morbidity and mortality in four to six million people in Germany (about 20% of the infected population).

In this context, this Health Technology Assessment (HTA) report assesses the medical and health economic benefit of the <sup>13</sup>C-urea breath test compared to other invasive and non-invasive tests used in the primary assessment of *H. pylori*.

### 2. Scientific background

The infection with *H. pylori* occurs via the oral-oral or the faecal-oral route and is associated with low socioeconomic status, residential density and poor hygiene. An acute *H. pylori* infection usually causes no symptoms. A persisting infection causes a chronic inflammation of the mucus layer in the stomach. This chronic inflammation is a risk factor for gastro duodenal ulcers, gastric cancer and MALT-lymphomas (Mucosa Associated Lymphoid Tissue). The tests available for assessing an infection with *H. pylori* can be categorized into invasive and non-invasive methods. The invasive methods are based on the direct detection of *H. pylori* in biopsy samples taken during endoscopy. The invasive methods are: histology, urease tests, microbiological tests and the polymerase chain reaction (PCR). The non-invasive methods are based on the indirect detection of *H. pylori* in the blood, breath, urine or stool. The non-invasive methods are: the immunoglobulin G (IgG) test, the stool antigen test and the <sup>13</sup>C-urea breath test and <sup>14</sup>C-urea breath test. The <sup>13</sup>C-urea breath test is recommended for monitoring *H. pylori* eradication. The test is currently not used in the primary assessment of *H. pylori* in Germany.

### 3. Health economic background

The resources of health care systems are limited. Therefore it is crucial to allocate resources rationally within the health care system, i. e. by optimising the resource allocation. This implies that the medical benefit cannot be the only criterion for assessing medical procedures. Both, the effectiveness (. medical results) and the costs

of the medical procedures have to be regarded. This is the purpose of health economic evaluation studies. The price of the <sup>13</sup>C-urea breath tests are relatively high compared to other relevant tests. Thus, the objective of the health economic part of this report is to investigate whether the medical effectiveness as well as the other cost components of the test-and-treat strategy can compensate for the costliness of this test.

Health economic studies of high quality should be capable of capturing the complexity of treatment-oriented diagnostics of gastroenterological diseases. Therefore the following matters should be taken into account. The study should be based on comparative analysis. In order to include all relevant direct, indirect and intangible costs, a social perspective should be used. Furthermore multidimensional measures of effectiveness would be more appropriate for capturing all relevant outcomes. As the treatment of stomach and intestinal disorders are associated with several factors of uncertainty, complex multivariate and probabilistic sensitivity analysis should be carried out. Moreover, the time horizon of the model should be rather long-term.

### 4. Research questions

#### Medical questions:

**ME1** What is the sensitivity and specificity of the <sup>13</sup>C-urea breath test in the primary assessment of *H. pylori* compared to other methods?

#### Health economic questions:

**HE1** What is the cost-effectiveness of the <sup>13</sup>C-urea breath test for the diagnosis of *H. pylori* infection compared to that of other diagnostic methods?

**HE2** What factors should be considered in health economic evaluations to assess the cost-effectiveness of the <sup>13</sup>C-urea breath test compared to other invasive and non-invasive methods?

**HE3** What can be concluded from current health economic studies, especially concerning the reimbursement regulations for the <sup>13</sup>C-urea breath test?

**HE4** What further health economic research is needed?

#### Ethical questions:

**ETH1** What ethical aspects concerning the reimbursement regulations for the <sup>13</sup>C-urea breath test have to be considered?

**ETH2** What ethical aspects concerning the test-and-treat strategy using the <sup>13</sup>C-urea breath test in the management of dyspepsia have to be considered?

**ETH3** What ethical aspects concerning invasive diagnostic methods for *H. pylori* infections have to be considered?

### 5. Methods

A structured and sensitive search of the literature was performed on 03/09/2008 by the Deutsches Institut für Medizinische Dokumentation und Information (DIMDI) to assess the medical and health economic effectiveness

of the  $^{13}\text{C}$ -urea breath test for *H. pylori* for the primary assessment of *H. pylori* compared to other invasive and non-invasive methods. The following databases were used:

Deutsches Ärzteblatt (AR96); CCMed (CC00); NHS-CRD-DARE (CDAR94); DAHTA (DAHTA); gms Meetings (GM03); Hogrefe-Verlagsdatenbank und Volltexte (HG05); NHS-CRD-HTA (INAHTA); Kluwer-Verlagsdatenbank (KL97); Krause & Pachernegg-Verlagsdatenbank (KP05); Karger-Verlagsdatenbank (KR03); MEDIKAT (MK77); SOMED (SM78); Springer-Verlagsdatenbank (SP97); Thieme-Verlagsdatenbank (TV01); Cochrane Library – Central (CC-TR93); MEDLINE (ME60); CAB Abstracts (CV72); NHS Economic Evaluation Database (NHSEED); GLOBAL Health (AZ72); AMED (CB85); IPA (IA70); Derwent Drug Backfile (DH64); EMBASE (EM74); EMBASE Alert (EA08); Derwent Drug File (DD83); ISTEP + ISTEP/ISSHP (II78); SciSearch (IS74); BIOSIS Previews (BA26).

Titles and abstracts of the studies were assessed for relevance by two independent reviewers. The quality evaluation of the medical studies was done in accordance with the standards for reporting of diagnostic accuracy checklist. Only primary studies and systematic reviews describing sensitivity and specificity were included. For the health economic assessment only evaluation studies describing cost-effectiveness, incremental cost-effectiveness ratio, cost benefit ratio or cost value benefit ratio were included. Special attention was paid to the method of analysis and the quality of the health economic models.

## 6. Results

### 6.1 Quantitative results

Using the defined search terms 1,035 medical, 117 economic and one ethical/legal publication are identified. Thereof, 30 medical, five economic and no ethical/legal publications are included. The hand search resulted in four other relevant economic publications so that a total of nine economic publications are included.

### 6.2 Qualitative results

#### 6.2.1 Included medical publications

Using the search terms 1,035 publications are identified. After titles/abstracts are reviewed, 99 publications are ordered as full texts. Of these 99 publications, 30 meet the inclusion criteria.

The studies include between 22 and 316 participants. In total 3,415 patients take part in 30 studies. 15 studies include adults, 14 studies comprehend children and youths and one study does not report the age of the patients involved.

The  $^{13}\text{C}$ -urea breath test is compared to the IgG test 18 times, 13 times with the stool antigen test, eleven times with the urease test, three times with histology, and one time each with PCR and the  $^{14}\text{C}$ -urea breath test. The sensitivity of the  $^{13}\text{C}$ -urea breath test is between 75% and

100%, the specificity between 55% and 100%. The sensitivity of the IgG test is between 50% and 100%, the specificity between 52% and 100%. The sensitivity of the stool antigen tests is between 50% and 98%, the specificity between 63% and 100%. The sensitivity of the urease tests is between 79 % and 100 %, the specificity between 59% and 100%. Sensitivity and specificity higher than 90% are found in 84% of the studies for the  $^{13}\text{C}$ -urea breath test. Sensitivity and specificity higher than 90% are found in 62% of the studies for the stool antigen test, for the IgG test in 56% (sensitivity) and 44% (specificity) of the studies, for the urease test in 73% (sensitivity) and 55% (specificity) of the studies. Compared to the IgG, the sensitivity of the IgG test is higher in twelve studies, lower in six studies and one study reports no differences. The specificity is higher in 13 studies, lower in three studies and two studies do not find any differences. Compared to the stool antigen test, the sensitivity of the  $^{13}\text{C}$ -urea breath test is higher in nine studies, lower in three studies and one study reports no difference. The specificity is higher in nine studies, lower in two studies and two studies do not inform about any differences. Compared to the urease tests, the sensitivity of the  $^{13}\text{C}$ -urea breath test is higher in four studies, lower in three studies and four studies report no variations. The specificity grows in five studies, decreases in five studies and one study reports no difference. Compared to histology, the sensitivity of the  $^{13}\text{C}$ -urea breath test is higher in one study and lower in two studies. The specificity is higher in two studies and lower in one study. The  $^{13}\text{C}$ -urea breath test has a lower sensitivity and a higher specificity compared to the PCR. There is no difference between the  $^{13}\text{C}$ -urea breath test and the  $^{14}\text{C}$ -urea breath test.

#### 6.2.2 Included economic publications

On the basis of defined inclusion and exclusion criteria, nine economic studies are included in this HTA report. Four of these studies report a comparison of cost-effectiveness between test-and-treat using the  $^{13}\text{C}$ -urea breath test and test-and-treat using other non-invasive diagnostic tests for the detection of *H. pylori*. In three studies, a cost-effectiveness analysis is performed comparing a test-and-treat approach based on the  $^{13}\text{C}$ -urea breath test to other management strategies for dyspepsia. A discrete event simulation incorporating first and second order simulation to determine the benefits and costs of a test-and-treat strategy using the urea breath test as well as other strategies for the management of uninvestigated dyspepsia is carried out in another study. In one case, the model only compares the costs of test-and-treat on the basis of urea breath test to endoscopy assuming medical equivalence of both methods. Test-and-treat using the  $^{13}\text{C}$ -urea breath test is compared to test-and-treat procedures on the basis of serology in six studies and to test-and-treat based on the stool antigen test in three models (but only one study reports the outcomes of both strategies). Additionally, test-and-treat strategies using several diagnostic tests (as confirmation) are evaluated in two studies. Fur-

thermore, test-and-treat based on the urea breath test is compared to different antisecretory therapies in four models, to empirical eradication therapy in two models and to endoscopy based procedures in five models. If necessary the endoscopy based strategies include invasive testing for H. pylori infection.

According to the results of the economic studies, the test-and-treat strategy based on the  $^{13}\text{C}$ -urea breath test for the management of patients with unexamined dyspepsia is neither dominated by serological strategy nor by an empirical anti-secretory therapy or an empirical eradication therapy. Test-and-treat using the breath test turns out to be cost-effective over test-and-treat using serology in three studies and over the empirical antisecretory therapy in two studies but is dominated in another study by the test-and-treat strategy on the basis of the stool antigen test. The results of two models show a domination of the breath test approach over endoscopy based strategies whereas one study reports a domination of endoscopy over test-and-treat using the  $^{13}\text{C}$ -urea breath test.

### 6.2.3 Included ethical, social and judicial publications

No publications could be found that described ethical, social or judicial aspects of the primary assessment methods of H. pylori infection. Therefore it is not possible to evaluate the ethical, social and judicial implications of the assessment methods for H. pylori infection in this HTA report.

## 7. Discussion

### 7.1 Discussion of medical aspects

The results of the included studies are heterogeneous with regard to the sensitivity and specificity of the tests. Possible explanations are the differences in the study populations, the choice of the reference tests and the different ways of conducting the H. pylori tests. Also, the majority of the studies do not report the statistical significance of the differences in sensitivity and specificity, i. e. based on the results of the included studies, only tendencies can be described for the test quality of the  $^{13}\text{C}$ -urea breath test and alternative testing methods for H. pylori.

In direct comparisons the  $^{13}\text{C}$ -urea breath test shows higher sensitivity and specificity than the IgG and stool antigen tests. In comparison to the urease test, results for sensitivity are inconsistent, and the specificity is slightly higher for the  $^{13}\text{C}$ -urea breath test. There are not enough results for comparisons between the  $^{13}\text{C}$ -urea breath test and the  $^{14}\text{C}$ -urea breath test, histology and PCR to describe tendencies.

### 7.2 Discussion of economic aspects

The cost-effectiveness of testing for H. pylori infection using the  $^{13}\text{C}$ -urea breath test compared to other diagnostic methods must be assessed using management

strategies, because H. pylori assessment should only be performed if there are clear strategies for dealing with the test results. To identify an optimal method, strategies that do not incorporate a H. pylori test are also to be taken into consideration.

All included studies had limitations to a greater or lesser extent. None of the included models is able to completely capture the complexity of managing patients with dyspeptic complaints. This is due to the use of one-dimensional outcome measures, insufficient consideration of therapeutic implications and follow-ups, time horizons which are too short or unverifiable determination of utility values and social costs. Furthermore, it must be considered that a cost-effectiveness comparison between alternative management strategies is limited in several economic analyses. In the majority of the included studies, disease specific outcome measures are used. Due to the lack of expert literature on threshold values for incremental outcome gains, the reported incremental cost-effectiveness ratios cannot be assessed. Therefore, conclusions on the cost-effectiveness of test-and-treat using the urea breath test in comparison to other strategies can only be drawn for dominant cost-effectiveness ratios.

According to the results of the included economic studies, the test-and-treat strategy using the  $^{13}\text{C}$ -urea breath test seems to be cost-effective compared to test-and-treat based on serology and compared to the empirical antisecretory therapy. Due to a lack of valid studies, it is not possible to assess the breath test approach in comparison to test-and-treat using the stool antigen test and the empirical eradication therapy respectively, regarding the cost-effectiveness. The results of economic analyses comparing test-and-treat using the breath test to endoscopy strategies are too heterogeneous to draw any conclusions.

### 7.3 Discussion of ethical, judicial and social aspects

From a social and ethical perspective, available resources should be effectively used to prevent rationing. Because there are no high-quality health economic studies that compare test-and-treat strategies using the  $^{13}\text{C}$ -urea breath test to alternative methods explicitly for the German population, the effectiveness of the current resource allocation remains uncertain.

A limitation of test-and-treat strategies is that they cannot be used to diagnose gastric cancer, ulcers etc. Endoscopy on the other hand, is associated with significant discomfort and a very low risk of serious complications. Currently, the only fully reimbursed alternative to endoscopy for the primary assessment of H. pylori is IgG testing. The IgG test is less effective than the  $^{13}\text{C}$ -urea breath test and other methods. This regulation will become more important, as the incidence of H. pylori infections continues to increase in persons with lower socioeconomic status.

## 8. Conclusions/Recommendations

The results of the included medical and health economic studies assessing the  $^{13}\text{C}$ -urea breath test in comparison to other diagnostic methods for *H. pylori* detection are not sufficient to recommend the  $^{13}\text{C}$ -urea breath test as a standard primary assessment method in the context of a test-and-treat strategy for managing patients with dyspeptic disorders in Germany.

Because of the public health relevance of the topic, high-quality economic and medical studies are needed in order to assess the effectiveness of the  $^{13}\text{C}$ -urea breath test for *H. pylori* infection. In addition the cost-effectiveness of serology test should be studied in more detail.

## Kurzfassung

### 1. Gesundheitspolitischer Hintergrund

Erkrankungen des Magen-Darm-Systems zählen in Deutschland zu den häufigsten Erkrankungen. Pro Quartal werden ca. sieben Mio. Patienten wegen Magen-Darm-Erkrankungen behandelt. Mit einem Volumen von 34,6 Mio. Verordnungen und einem Umsatz von 1,363 Mrd. Euro liegt diese Indikationsgruppe 2007 innerhalb der gesetzlichen Krankenversicherung (GKV) auf Platz 6 der meistverschriebenen Medikamente und auf Platz 7 der umsatzstärksten Indikationsgruppen. Ein großer Anteil der jährlichen gastroenterologischen Krankheitsfälle wird durch eine *Helicobacter pylori* (*H. pylori*)-Infektion verursacht. Diese ist trotz deutlich abnehmender Inzidenzraten in den Industrienationen eine der am häufigsten auftretenden bakteriellen Infektionskrankheiten des Menschen. Über die Hälfte der Weltbevölkerung ist mit dem Keim infiziert. Für Deutschland werden in epidemiologischen Studien Gesamtprevalenzraten zwischen 20 bis 40% der Bevölkerung ermittelt. Laut dem Nationalen Referenzzentrum (NRZ) für *H. pylori* (Institut für medizinische Mikrobiologie und Hygiene der Universität Freiburg) sind etwa 20 bis 30 Mio. Menschen in Deutschland mit dem Bakterium infiziert (ca. 30% der deutschen Bevölkerung). Die Besiedlung der Magenschleimhaut durch *H. pylori* führt bei allen Infizierten zu einer zumeist lebenslangen entzündlichen Reaktion, die bei etwa vier bis sechs Mio. deutschen *H. pylori*-Infizierten (ca. 20% der *H. pylori*-Infizierten) mit erheblicher Mortalität und Morbidität verbundene Folgeerkrankungen auslöst.

Vor diesem Hintergrund untersucht der vorliegende Health Technology Assessment (HTA)-Bericht den medizinischen und gesundheitsökonomischen Nutzen des  $^{13}\text{C}$ -Harnstoff-Atemtests im Vergleich zu invasiven und nichtinvasiven Testverfahren zur Untersuchung auf *H. pylori*-Besiedlung im Rahmen der Primärdiagnostik.

### 2. Wissenschaftlicher Hintergrund

Die Infektion mit *H. pylori* erfolgt vermutlich auf oral-oralen oder fäkal-oralen Weg, begünstigt durch beengte

Wohnverhältnisse und mangelhafte Hygiene. Eine akute Infektion mit *H. pylori* verläuft meistens beschwerdefrei. Eine persistierende Infektion hat eine chronische Entzündung der Magenschleimhaut zur Folge. Eine solche chronische Entzündung stellt einen Risikofaktor dar für Erkrankungen wie gastroduodenale Geschwüre, Magenkarzinome und MALT (Mucosa Associated Lymphoid Tissue)-Lymphome. Zur Diagnose einer *H. pylori*-Infektion stehen verschiedene Testverfahren zur Verfügung, die sich in invasive und nichtinvasive Methoden einteilen lassen. Die invasiven Methoden basieren auf einer direkten Identifikation von *H. pylori* in Biopsieproben, die im Rahmen einer Endoskopie entnommen werden. Zu den invasiven Verfahren zählen die Histologie, der Urease-Schnelltest, der mikrobiologische Test sowie die Polymerase-Kettenreaktion (PCR). Die nichtinvasiven Methoden basieren auf einem indirekten Nachweis von *H. pylori* aufgrund bestimmter Eigenschaften des Bakteriums oder der Reaktion des Immunsystems. Zu den nichtinvasiven Verfahren zählen der Immunglobulin G (IgG)-Nachweis, der Stuhl-Antigen-Test sowie der  $^{13}\text{C}$ - und der  $^{14}\text{C}$ -Harnstoff-Atemtest. Der  $^{13}\text{C}$ -Harnstoff-Atemtest wird zur Kontrolle einer Eradikationstherapie empfohlen, kommt aufgrund der Erstattungsregelungen und den Empfehlungen der Deutschen Gesellschaft für Verdauungs- und Stoffwechselkrankheiten (DGVS) in der Primärdiagnostik von *H. pylori* derzeit nicht standardmäßig in Deutschland zum Einsatz.

### 3. Gesundheitsökonomischer Hintergrund

Die dem Gesundheitswesen zur Verfügung stehenden Mittel sind begrenzt. Deshalb ist es entscheidend, die eingesetzten Ressourcen im Gesundheitssystem rational zu verteilen, d. h. eine optimale Ressourcenallokation zu gewährleisten. Dies ist nur möglich, sofern Über-, Unter- und Fehlversorgungen vermieden werden. Diese Forderung impliziert, dass der rein medizinische Nutzen als alleiniges Kriterium zur Beurteilung einer Maßnahme nicht ausreichend ist. Vielmehr müssen der Effektivität (= medizinisches Ergebnis) die Kosten der medizinischen Maßnahme gegenübergestellt werden, so dass eine ökonomische Bewertung realisiert werden kann. Dies ist das Anliegen von gesundheitsökonomischen Evaluationsstudien. Da der Preis für die Durchführung und Auswertung des  $^{13}\text{C}$ -Harnstoff-Atemtests im Vergleich zu anderen relevanten Tests relativ hoch ist, gilt es im gesundheitsökonomischen Teil des vorliegenden Berichts anhand der aktuellen Studienlage zu untersuchen, ob die medizinische Effektivität sowie die übrigen Kostenkomponenten der Test-and-Treat-Strategie auf Basis des Atemtests diesen Nachteil kompensieren können.

Zur Erfassung der Komplexität einer therapieorientierten Diagnostik gastroenterologischer Erkrankungen sollte das ideale gesundheitsökonomische Studiendesign auf einer vergleichenden Analyse basieren. Dabei wäre eine gesellschaftliche Perspektive einzunehmen, die alle rele-



vanten direkten, indirekten und intangiblen Kosten berücksichtigt. Das abschließende Effektivitätsmaß müsste multidimensional sein, um alle relevanten Outcomes zu messen. Da die Behandlung von Magen- und Darmerkrankungen mit vielen Unsicherheitsfaktoren behaftet ist, wären komplexe multivariate und probabilistische Sensitivitätsanalysen ideal. Zudem müsste der Zeithorizont des Modells relativ langfristig sein.

## 4. Forschungsfragen

### Medizinische Fragestellung:

**ME1** Wie effektiv ist die Untersuchung auf *H. pylori* mittels <sup>13</sup>C-Harnstoff-Atemtest in der primären Diagnostik hinsichtlich der Sensitivität und Spezifität im Vergleich zu anderen diagnostischen Testverfahren?

### Gesundheitsökonomische Fragestellungen:

**ÖK1** Wie ist die Kosten-Effektivität der Untersuchung auf *H. pylori*-Besiedlung mittels <sup>13</sup>C-Harnstoff-Atemtest in der Primärdiagnostik im Vergleich zu anderen diagnostischen Methoden?

**ÖK2** Welche Kriterien und Parameter sollten in gesundheitsökonomischen Evaluationsstudien zur Kosten-Effektivität des <sup>13</sup>C-Harnstoff-Atemtests in der Primärdiagnostik im Vergleich zu anderen invasiven und nichtinvasiven diagnostischen Verfahren berücksichtigt werden, damit die Ergebnisse als Entscheidungshilfe bei der Auswahl der kosteneffektivsten therapieorientierten Diagnostik im Einzelfall dienlich sind?

**ÖK3** Welche gesundheitspolitischen Schlussfolgerungen, insbesondere hinsichtlich der Erstattungsfähigkeit des <sup>13</sup>C-Harnstoff-Atemtests, ergeben sich aus den vorliegenden gesundheitsökonomischen Informationen zur Kosten-Effektivität des Atemtests im Vergleich zu anderen diagnostischen Methoden?

**ÖK4** Wo besteht weiterer gesundheitsökonomischer Forschungsbedarf?

### Ethische Fragestellungen:

**ETH1** Welche ethischen Aspekte sind hinsichtlich der Erstattungsfähigkeit des <sup>13</sup>C-Harnstoff-Atemtests zu berücksichtigen?

**ETH2** Welche ethischen Aspekte sind hinsichtlich einer Test-and-Treat-Strategie mittels <sup>13</sup>C-Harnstoff-Atemtest beim Management der Dyspepsie zu berücksichtigen?

**ETH3** Welche ethischen Aspekte sind im Zusammenhang mit invasiven Diagnoseverfahren zur Untersuchung auf *H. pylori*-Besiedlung zu berücksichtigen?

## 5. Methodik

Um den medizinischen und gesundheitsökonomischen Nutzen der Untersuchung auf *H. pylori*-Besiedlung mittels <sup>13</sup>C-Harnstoff-Atemtest in der Primärdiagnostik im Vergleich zu invasiven und nichtinvasiven diagnostischen Verfahren beurteilen zu können, wurde vom Deutschen Institut für Medizinische Dokumentation und Information (DIMDI) am 03.09.2008 eine strukturierte und hochsen-

sitive Literaturrecherche in folgenden 28 Datenbanken durchgeführt:

Deutsches Ärzteblatt (AR96); CCMed (CC00); NHS-CRD-DARE (CDAR94); DAHTA-Datenbank (DAHTA); gms Meetings (GM03); Hogrefe-Verlagsdatenbank und Volltexte (HG05); NHS-CRD-HTA (INAHTA); Kluwer-Verlagsdatenbank (KL97); Krause & Pacherneegg-Verlagsdatenbank (KPO5); Karger-Verlagsdatenbank (KRO3); MEDIKAT (MK77); SOMED (SM78); Springer-Verlagsdatenbank (SP97); Thieme-Verlagsdatenbank (TV01); Cochrane Library – Central (CCTR93); MEDLINE (ME60); CAB Abstracts (CV72); NHS Economic Evaluation Database (NHSEED); GLOBAL Health (AZ72); AMED (CB85); IPA (IA70); Derwent Drug Backfile (DH64); EMBASE (EM74); EMBASE Alert (EA08); Derwent Drug File (DD83); ISTPB + ISTP/ISSHP (II78); SciSearch (IS74); BIOSIS Previews (BA26).

Auf Basis der Suchbegriffe (siehe Anhang) werden vier Recherchen ohne Einschränkung der Sprache und des Zeitraums durchgeführt sowie durch Handsuchen ergänzt. Die Literaturtitel und Abstracts identifizierter Publikationen werden auf der Basis der Themenstellung systematisch durch zwei unabhängige Gutachter auf Relevanz für das vorliegende Thema überprüft. Die Bewertung der Qualität medizinischer Studien erfolgt anhand der „Standards for Reporting of Diagnostic Accuracy“ Checkliste des „Standards for Reporting of Diagnostic Accuracy Steering Committee“<sup>8</sup>. Aus dieser Checkliste folgen Mindestanforderungen an die methodische Qualität der Publikationen. Es werden nur Primärstudien und systematische Reviews einbezogen. In diesen müssen die Zielgrößen Sensitivität und Spezifität der verglichenen diagnostischen Methoden angegeben sein. Bei den gesundheitsökonomischen Studien wird geprüft, ob die Fragestellung und der Evaluationsrahmen den ökonomischen Problemkontext ausreichend darstellen. Betrachtet werden im Besonderen die Analysemethoden und die Güte der Modellierung. Für die gesundheitsökonomische Beurteilung werden nur vollständige gesundheitsökonomische Evaluationsstudien mit den Zielparametern Kosten-Effektivität, inkrementelles Kosten-Effektivitäts-Verhältnis, Kosten-Nutzen- und Kosten-Nutzwert-Beziehungen einbezogen.

## 6. Ergebnisse

### 6.1 Quantitative Ergebnisse

Insgesamt ergibt die systematische Datenbankrecherche des DIMDI 1.153 Treffer, die in Form einer Reference-Manager-Datei digital zur Verfügung gestellt werden. Auf der Grundlage der definierten Suchbegriffe wurden 1.035 medizinische, 117 ökonomische und ein ethischer/rechtlicher Treffer identifiziert. Nach zwei Selektionsprozessen werden insgesamt 30 medizinische, fünf ökonomische und null ethische/rechtliche/soziale Publikationen einbezogen. Im Rahmen der Handrecherche können vier weitere relevante ökonomische Publikationen identifiziert

werden, so dass insgesamt neun ökonomische Studien berücksichtigt werden.

## 6.2 Qualitative Ergebnisse

### 6.2.1 Eingeschlossene medizinische Publikationen

Aufgrund der Suchbegriffe werden 1.035 medizinische Publikationen identifiziert. Davon werden aufgrund von Titel und Abstract 99 medizinische Publikationen im Volltext bestellt. Von diesen 99 Veröffentlichungen werden 30 für den vorliegenden Bericht berücksichtigt.

In 15 Studien werden die Testkriterien für volljährige Personen angegeben, in 14 Studien für Kinder sowie Jugendliche. In einer weiteren Studie werden keine Angaben zum Altersbereich der Patienten gemacht. Insgesamt haben 3.415 Personen an den 30 Studien teilgenommen. Die Fallzahl liegt zwischen 22 und 316 Teilnehmern.

Der <sup>13</sup>C-Harnstoff-Atemtest wird 18-mal mit dem IgG-Nachweis, 13-mal mit dem Stuhl-Antigen-Test, elfmal mit dem Urease-Schnelltest, dreimal mit der Histologie, und je einmal mit der PCR und dem <sup>14</sup>C-Harnstoff-Atemtest verglichen. Die Sensitivität des <sup>13</sup>C-Harnstoff-Atemtests liegt zwischen 75% und 100%, die Spezifität zwischen 55% und 100%. Die Sensitivität des IgG-Tests liegt zwischen 50% und 100%, die Spezifität zwischen 52% und 100%. Die Sensitivität des Stuhl-Antigen-Tests liegt zwischen 50% und 98%, die Spezifität zwischen 63% und 100%. Die Sensitivität des Urease-Schnelltests liegt zwischen 79% und 100%, die Spezifität zwischen 59% und 100%. Sensitivität und Spezifität von 90% oder höher werden vom <sup>13</sup>C-Harnstoff-Atemtest in 84% der Studien erreicht. Der Stuhl-Antigen-Test weist in 62% der Studien eine Sensitivität und eine Spezifität von über 90% auf, der IgG-Test in 56% bzw. 44% der Studien und der Urease-Schnelltest in 73% bzw. 55% der Studien. Im Vergleich zum IgG-Test sind die Sensitivität des <sup>13</sup>C-Harnstoff-Atemtests zwölfmal höher, sechsmal niedriger und einmal gleich und die Spezifität 13-mal höher, dreimal niedriger und zweimal gleich. Im Vergleich zum Stuhl-Antigen-Test sind die Sensitivität des <sup>13</sup>C-Harnstoff-Atemtests neunmal höher, dreimal niedriger und einmal gleich, und die Spezifität neunmal höher, zweimal niedriger und zweimal gleich. Im Vergleich zum Urease-Schnelltest ist die Sensitivität des <sup>13</sup>C-Harnstoff-Atemtests viermal höher, dreimal niedriger und viermal gleich sowie die Spezifität fünfmal höher, fünfmal niedriger und einmal gleich. Im Vergleich mit der Histologie ist die Sensitivität des <sup>13</sup>C-Harnstoff-Atemtests einmal höher und zweimal niedriger sowie die Spezifität zweimal höher und einmal niedriger. In je einem Vergleich zeigt sich kein Unterschied zwischen <sup>13</sup>C-Harnstoff-Atemtest und <sup>14</sup>C-Harnstoff-Atemtest, sowie eine niedrigere Sensitivität und höhere Spezifität im Vergleich zur PCR.

### 6.2.2 Eingeschlossene ökonomische Publikationen

Auf Basis der definierten Ein- und Ausschlusskriterien werden neun ökonomische Studien für den vorliegenden

HTA-Bericht berücksichtigt. Davon beinhalten sieben Studien eine Analyse der Kosten-Effektivität, die sich bei vier Studien allein auf die jeweiligen diagnostischen, nichtinvasiven Methoden im Rahmen einer Test-and-Treat-Strategie zum Management der Dyspepsie bezieht und bei drei Studien auf unterschiedliche Diagnose-Therapie-Strategien zum Management der Dyspepsie. In einer Studie werden die Kosten und entsprechenden Nutzwerte verschiedener Strategien zum Management nichtuntersuchter Dyspepsie auf Basis einer Discrete-event-Simulation der ersten und zweiten Ordnung analysiert und in einer weiteren Studie ein reiner Kostenvergleich einer Test-and-Treat-Strategie mittels <sup>13</sup>C-Harnstoff-Atemtest gegen Endoskopie zum Management nichtuntersuchter Dyspepsie durchgeführt, da in diesem Fall von einer identischen medizinischen Effektivität beider Technologien ausgegangen wird. Die Test-and-Treat-Strategie mittels <sup>13</sup>C-Harnstoff-Atemtest wird in sechs Studien mit einem Test-and-Treat-Verfahren auf Basis der Serologie, in drei mit einem Test-and-Treat-Verfahren auf Basis des Stuhl-Antigen-Tests (davon werden allerdings nur in einer Studie die Ergebnisse beider Strategien dargestellt) verglichen. Zusätzlich evaluieren zwei Studien Test-and-Treat-Strategien mit mehreren Tests (als Bestätigung des Befundes). Des Weiteren wird die Test-and-Treat-Strategie auf Basis des <sup>13</sup>C-Harnstoff-Atemtests in vier Studien mit unterschiedlichen empirischen antisekretorischen Therapien, in zwei mit empirischen Eradikationstherapien und in fünf Modellen mit endoskopiebasierten Verfahren, die ggf. einen invasiven Test einschließen, verglichen. Die Test-and-Treat-Strategie mittels <sup>13</sup>C-Harnstoff-Atemtest wird in keinem Modell weder von einer Test-and-Treat-Strategie auf Basis der Serologie noch von einer empirischen antisekretorischen Therapie oder einer empirischen Eradikationstherapie dominiert. Das Atemtestverfahren erweist sich in drei Modellen als kosteneffektiv gegenüber der serologischen Prozedur und wird in einem Modell von der Test-and-Treat-Strategie auf Basis des Stuhl-Antigen-Tests dominiert. Des Weiteren zeigen die Ergebnisse von zwei Studien eine Kosten-Effektivität der Atemteststrategie gegenüber der empirischen antisekretorischen Therapie. Endoskopiebasierte Verfahren werden in zwei Modellen von der Test-and-Treat-Strategie mittels <sup>13</sup>C-Harnstoff-Atemtest und die Atemtestmethode in einer Untersuchung von der endoskopischen Prozedur dominiert.

### 6.2.3 Eingeschlossene ethische, soziale und juristische Publikationen

Im Rahmen des vorliegenden HTA-Berichts können keine Publikationen identifiziert werden, die sich mit ethischen, sozialen oder juristischen Aspekten der Primärdiagnostik einer H. pylori-Infektion beschäftigen. Die Bearbeitung ethischer, sozialer oder juristischer Implikationen diagnostischer Verfahren zur Untersuchung auf H. pylori-Besiedlung anhand der Literatur ist deshalb nicht möglich.

## 7. Diskussion

### 7.1 Diskussion medizinischer Aspekte

Die Ergebnisse der berücksichtigten Studien sind heterogen, sowohl hinsichtlich der Angaben zu Sensitivität und Spezifität der jeweiligen Tests als auch hinsichtlich der Vergleiche des <sup>13</sup>C-Harnstoff-Atemtests mit den alternativen diagnostischen Verfahren. Mögliche Ursachen hierfür sind Unterschiede in den untersuchten Populationen, den gewählten Referenztests sowie Unterschiede in der Durchführung der H. pylori-Tests.

Zudem macht die Mehrzahl der Studien keine Angaben zur statistischen Signifikanz der beschriebenen Unterschiede der Sensitivität und Spezifität. Somit können basierend auf den Ergebnissen der eingeschlossenen Studien allenfalls Tendenzen hinsichtlich der Testgüte des <sup>13</sup>C-Harnstoff-Atemtest im Vergleich zu alternativen Nachweisverfahren auf H. pylori angegeben werden. Im direkten Vergleich weist der <sup>13</sup>C-Harnstoff-Atemtest überwiegend eine bessere Testgüte als der IgG- und der Stuhl-Antigen-Test auf. Die Vergleiche mit dem Urease-Schnelltest sind hinsichtlich der Sensitivität uneinheitlich, hinsichtlich der Spezifität weist der <sup>13</sup>C-Harnstoff-Atemtest in den hier beschriebenen Studien tendenziell höhere Werte auf als der Urease-Schnelltest. Für die Vergleiche des <sup>13</sup>C-Harnstoff-Atemtests mit der Histologie, dem <sup>14</sup>C-Harnstoff-Atemtest und der PCR liegen nur wenige Ergebnisse vor, aus denen sich keine Tendenzen ableiten lassen.

### 7.2 Diskussion ökonomischer Aspekte

Im gesundheitsökonomischen Kontext muss die Kosten-Effektivität der Untersuchung auf H. pylori-Besiedlung mittels <sup>13</sup>C-Harnstoff-Atemtest in der Primärdiagnostik im Vergleich zu anderen diagnostischen Methoden auf Basis von Managementstrategien erfolgen, da eine H. pylori-Diagnostik nur dann durchzuführen ist, wenn eindeutige Strategien zum Umgang mit dem Testergebnis vorhanden sind. Um ein gesundheitsökonomisch optimales Verfahren zu ermitteln, sind in diese Betrachtung auch Strategien einzubeziehen, die keinen H. pylori-Test beinhalten.

Alle eingeschlossenen ökonomischen Studien weisen mehr oder minder gravierende Limitationen auf und liefern teils sehr heterogene Ergebnisse. Insgesamt kann keines der ökonomischen Modelle der Komplexität des Managements von Patienten mit dyspeptischen Beschwerden gänzlich gerecht werden. Gründe hierfür liegen insbesondere in den verwendeten Effektivitätsmaßen, der unzureichenden Berücksichtigung therapeutischer Konsequenzen und Follow-ups sowie eines zu geringen Zeithorizonts oder einer nicht nachvollziehbaren Generierung der Nutzwerte und sozialen Kosten. Des Weiteren gilt es zu berücksichtigen, dass ein Kosten-Effektivitäts-Vergleich der einzelnen Managementstrategien bei mehreren Studien nur eingeschränkt möglich ist. In der Mehrzahl der eingeschlossenen Untersuchungen werden spezielle krankheitsspezifische Effektivitätsmaße verwendet, für

die in der Literatur keine Analysen zu Zahlungsbereitschaften für zusätzliche Outcomeeinheiten vorliegen, so dass das entsprechende inkrementelle Kosten-Effektivitäts-Verhältnis nicht bewertet werden kann. In diesem Zusammenhang können Aussagen zur Kosten-Effektivität der Atemteststrategie im Vergleich zu anderen Verfahren nur dann getroffen werden, wenn Dominanzen bestehen.

Insgesamt deuten die Ergebnisse einiger qualitativ hochwertiger Studien eine Kosten-Effektivität der Test-and-Treat-Strategie mittels <sup>13</sup>C-Harnstoff-Atemtest gegenüber einer Test-and-Treat-Strategie auf Basis der Serologie und der empirisch antisekretorischen Therapie an. Bezüglich eines Vergleichs zwischen der Atemteststrategie und der Test-and-Treat-Strategie mittels Stuhl-Antigen-Test sowie der empirischen Eradikationstherapie mangelt es an validen Ergebnissen bzw. ökonomischer Evidenz, um auf die ökonomische Vorteilhaftigkeit eines der Verfahren zu schließen. Aussagen zur Kosten-Effektivität der Test-and-Treat-Strategie mittels <sup>13</sup>C-Harnstoff-Atemtests gegenüber endoskopiebasierten Verfahren sind anhand der vorliegenden Studienergebnisse nicht möglich, da diese zu heterogen sind.

Für spezielle krankheitsspezifische Outcomes, die in den Untersuchungen häufig verwendet werden, liegen in der Literatur keine Analysen zu Zahlungsbereitschaften für zusätzliche Outcomeeinheiten vor. Die berechneten inkrementellen Kosten-Effektivitäts-Verhältnisse können deshalb nicht bewertet werden. Demnach ist ein eindeutiger Rückschluss auf die Kosten-Effektivität einzelner Verfahren, sofern keine Dominanzen bestehen, dann nicht möglich.

### 7.3 Diskussion ethischer, juristischer und sozialer Aspekte

Aus sozialem ethischer Sicht ist ein rationaler Ressourceneinsatz zu fordern, da dieses Rationierungen sowie Über-, Unter- und Fehlversorgungen in gewissem Maß vorbeugen kann. Da qualitativ hochwertige gesundheitsökonomische Studien fehlen, die die Test-and-Treat-Strategie mittels <sup>13</sup>C-Harnstoff-Atemtests mit entsprechenden alternativen Verfahren explizit für den deutschen Versorgungskontext vergleichen, bestehen derzeit Zweifel an einer rationalen Ressourcenallokation. Mit Test-and-Treat-Strategien sind Magenkarzinome, Ulkuserkrankungen etc. nicht direkt nachzuweisen. Dies führt zu einer Benachteiligung der Patienten, die eine entsprechende Erkrankung aufweisen. Demgegenüber werden bei einer Anwendung endoskopiebasierter Verfahren beim Management von Patienten mit dyspeptischen Beschwerden all jene benachteiligt, die von diesen Krankheiten eben gerade nicht betroffen sind, da diese Methoden mit signifikanten Unannehmlichkeiten verbunden sind und mit sehr geringer Wahrscheinlichkeit zu schwerwiegenden Komplikationen führen können. Patienten, die aus triftigen Gründen eine Endoskopie ablehnen, bekommen in der Primärdiagnostik von den GKV nur einen IgG-Test erstattet, der nachweislich weniger effektiv ist als der <sup>13</sup>C-Harnstoff-Atemtest und verschiedene invasive Verfahren. Folglich werden Personen, die

einen gleichen Bedarf haben, in einem unnötigen Maß ungleich behandelt. Dieser Sachverhalt gewinnt vor dem Hintergrund einer zunehmenden *H. pylori*-Prävalenz in sozialschwächeren Schichten zusätzlich an Bedeutung.

## 8. Schlussfolgerungen/Empfehlungen

Zusammenfassend ist festzuhalten, dass die Studienlage zur medizinischen und ökonomischen Beurteilung des <sup>13</sup>C-Harnstoff-Atemtests im Vergleich zu anderen diagnostischen Methoden nicht ausreichend ist, um den Atemtest als primärdiagnostisches Standardverfahren im Rahmen einer Test-and-Treat-Strategie beim Management von Patienten mit dyspeptischen Beschwerden für die deutsche Versorgungslandschaft insbesondere vor dem Hintergrund der DGVS-Leitlinien anstelle einer endoskopiebasierten Methode zu empfehlen.

Aufgrund der Public-health-Relevanz der Thematik werden deshalb qualitativ hochwertige ökonomische und medizinische Studien, die der Komplexität einer therapieorientierten Diagnostik gastroenterologischer Erkrankungen Rechnung tragen, für die deutsche Versorgungslandschaft benötigt, um den medizinischen und gesundheitsökonomischen Nutzen der Untersuchung auf *H. pylori*-Besiedlung mittels <sup>13</sup>C-Harnstoff-Atemtest in der Primärdiagnostik im Vergleich zu anderen diagnostischen Methoden und Managementstrategien zu bewerten. Neben einem Vergleich von Test-and-Treat-Verfahren gegen endoskopiebasierte Strategien sollte in Anbetracht der derzeitigen Erstattungsregelungen unbedingt geprüft werden, ob die Serologie im Vergleich zum <sup>13</sup>C-Harnstoff-Atemtest und

zum Stuhl-Antigen-Test aus gesellschaftlicher Perspektive ein kosteneffektiveres Verfahren ist.

### Korrespondenzadresse:

Alexander Kuhlmann  
Forschungsstelle für Gesundheitsökonomie, Leibniz Universität Hannover, Königsworther Platz 1, 30167 Hannover, Tel.: + 49-511/762-19446  
ak@ivbl.uni-hannover.de

### Bitte zitieren als

Nocon M, Kuhlmann A, Leodolter A, Roll S, Vauth C, Willich SN, Greiner W. Medizinischer und gesundheitsökonomischer Nutzen der Untersuchung auf *Helicobacter pylori*-Besiedlung mittels <sup>13</sup>C-Harnstoff-Atemtest in der Primärdiagnostik im Vergleich zu invasiven und nichtinvasiven diagnostischen Verfahren. *GMS Health Technol Assess.* 2009;5:Doc14.

### Artikel online frei zugänglich unter

<http://www.egms.de/en/journals/hta/2009-5/hta000076.shtml>

### Der vollständige HTA-Bericht steht zum kostenlosen Download zur Verfügung unter:

[http://portal.dimdi.de/de/hta/hta\\_berichte/hta257\\_bericht\\_de.pdf](http://portal.dimdi.de/de/hta/hta_berichte/hta257_bericht_de.pdf)

### Copyright

©2009 Nocon et al. Dieser Artikel ist ein Open Access-Artikel und steht unter den Creative Commons Lizenzbedingungen (<http://creativecommons.org/licenses/by-nc-nd/3.0/deed.de>). Er darf vervielfältigt, verbreitet und öffentlich zugänglich gemacht werden, vorausgesetzt dass Autor und Quelle genannt werden.