BMJ

BMJ 2012;345:e8012 doi: 10.1136/bmj.e8012 (Published 17 December 2012)

RESEARCH

CHRISTMAS 2012: RESEARCH

Pain over speed bumps in diagnosis of acute appendicitis: diagnostic accuracy study

Helen F Ashdown *academic clinical fellow in general practice*[1], Nigel D'Souza *specialist registrar in general surgery*[2], Diallah Karim *foundation trainee*[2], Richard J Stevens *senior medical statistician*[1], Andrew Huang *consultant colorectal and general surgeon*[2], Anthony Harnden *university lecturer in general practice*[1]

[1]Department of Primary Care Health Sciences, University of Oxford, Oxford OX2 6GG, UK; [2]Department of Surgery, Stoke Mandeville Hospital, Aylesbury HP21 8AL, UK

Abstract

Objective To assess the diagnostic accuracy of pain on travelling over speed bumps for the diagnosis of acute appendicitis.

Design Prospective questionnaire based diagnostic accuracy study.

Setting Secondary care surgical assessment unit at a district general hospital in the UK.

Participants 101 patients aged 17-76 years referred to the on-call surgical team for assessment of possible appendicitis.

Main outcome measures Sensitivity, specificity, positive and negative predictive values, and positive and negative likelihood ratios for pain over speed bumps in diagnosing appendicitis, with histological diagnosis of appendicitis as the reference standard.

Results The analysis included 64 participants who had travelled over speed bumps on their journey to hospital. Of these, 34 had a confirmed histological diagnosis of appendicitis, 33 of whom reported increased pain over speed bumps. The sensitivity was 97% (95% confidence interval 85% to 100%), and the specificity was 30% (15% to 49%). The positive predictive value was 61% (47% to 74%), and the negative predictive value was 90% (56% to 100%). The likelihood ratios were 1.4 (1.1 to 1.8) for a positive test result and 0.1 (0.0 to 0.7) for a negative result. Speed bumps had a better sensitivity and negative likelihood ratio than did other clinical features assessed, including migration of pain and rebound tenderness.

Conclusions Presence of pain while travelling over speed bumps was associated with an increased likelihood of acute appendicitis. As a diagnostic variable, it compared favourably with other features commonly used in clinical assessment. Asking about speed bumps may contribute to clinical assessment and could be useful in telephone assessment of patients.

Introduction

Speed bumps are a commonly used traffic calming device to reduce the speed of vehicles.[1] Although controversial, traffic calming measures have been associated with a 70% decrease in injuries among child pedestrians in some areas,[2] and they may be a promising intervention for reducing the overall number of road traffic injuries and deaths.[3] However, speed bumps may have a useful alternative benefit in the diagnosis of acute appendicitis.

Acute appendicitis is the most common surgical abdominal emergency.[4] Rapid diagnosis is important, because increased time between onset of symptoms and surgical intervention is associated with increased risk of appendiceal perforation and therefore potential peritonitis, sepsis, and death.[5] However, the rate of negative appendicectomy (when appendicectomy is performed, but the appendix is found to be normal on histological evaluation[4]) ranges from 5% to 42%,[6] and this can be associated with considerable morbidity.[7] Clinical diagnosis can be challenging, particularly in the early stages of appendicitis when clinical manifestations may be quite non-specific or atypical. Different elements of history, examination, and laboratory findings have varying predictive power in the diagnosis of appendicitis,[6] and algorithms and scoring systems for clinical evaluation exist,[4] but appendicitis can nevertheless be easily missed.[8]

Correspondence to: H F Ashdown helen.ashdown@phc.ox.ac.uk

Extra material supplied by the author (see http://www.bmj.com/content/345/bmj.e8012?tab=related#webextra)

Patients with appendicitis have sometimes been noted to complain of a worsening of their abdominal pain when they travel over speed bumps.⇓ Some doctors ask about this routinely as part of history taking, believing it to be a highly diagnostic feature (personal communication). We sought to determine whether any evidence supports this practice and to determine its predictive power as a diagnostic sign.

Methods

We did a prospective study at a district general hospital in Buckinghamshire in the United Kingdom. Roads in the county of Buckinghamshire are almost universally surfaced in tarmac and are smooth, with any speed bumps raised from the road surface in a variety of designs and elevations. All patients aged 16 or over who had been referred to the on-call surgical team as part of their usual care, by either a general practitioner or an emergency department doctor, with suspected appendicitis were eligible. They were identified consecutively over a six month period between February and August 2012.

We asked participants to complete a questionnaire survey about their symptoms, including four specific questions related to their journey into hospital: mode of transport, whether they had travelled over speed bumps, whether they had had pain on the journey, and whether the pain changed when they went over a speed bump. We defined patients as "speed bump positive" if they had a worsening of pain from baseline over speed bumps and as "speed bump negative" if their pain stayed the same, if they were unsure, or if their pain improved on going over speed bumps. To minimise recall bias, patients had to complete the questionnaire within 24 hours of arrival in hospital and before they had been to theatre. We also recorded examination findings on admission from their notes. Two of the authors entered data on to a spreadsheet, and a third author double checked transcription.

We then followed participants through their admission to determine the outcome and whether they were taken to theatre for presumed appendicitis. For those who had been to theatre, we obtained the subsequent histology report. We used histological diagnosis of appendicitis as the reference standard, which is the usual practice in studies of appendicitis.[6] One of the authors, who was blinded to all clinical details of the participants, corroborated interpretation of the histology findings. We also asked participants to provide contact details so that, if an alternative diagnosis or no diagnosis was made, we could contact them after their admission to ensure that their symptoms had resolved, to avoid missing cases of subacute or "grumbling" appendicitis. A positive or negative histological diagnosis of appendicitis was made in participants who went to theatre and had their appendix removed. We assumed participants whose symptoms resolved without surgery to have a negative diagnosis. We confirmed resolution of symptoms by telephone follow-up between two weeks and three months after admission.

In pilot data (11 cases and 21 controls) collected in 2009, the sensitivity was 82% (95% confidence interval 48% to 98%) and the specificity was 67% (43% to 85%). We used the R software package to simulate studies of varying sizes on the basis of these estimates. We calculated that 100-150 participants in the main study would be sufficient to show a likelihood ratio greater than 1.8-2.0.

We calculated the sensitivity, specificity, positive and negative predictive values, and positive and negative likelihood ratios, with 95% confidence intervals, for the outcome diagnosis of appendicitis. When a sign was recorded as "unsure," we considered it absent for the purposes of calculation. We restricted the primary analyses to those patients reported to have travelled over speed bumps on the route to the hospital. We also planned to compare the diagnostic accuracy of worsened pain over speed bumps with more conventional diagnostic features of appendicitis, such as migratory pain and rebound tenderness. We used the "diagt" command in Stata (Release 11) for calculations.

Results

One hundred and one patients were recruited into the study. The median age was 34 (range 17-76) years. Sixty one participants were taken to theatre for presumed appendicitis, of whom 54 had their appendix removed. Acute appendicitis was confirmed histologically in 43 of these, giving a negative appendicectomy rate of 20%.

Sixty eight participants had travelled over speed bumps. We excluded four patients from diagnostic accuracy analysis: one because histology was not available, and three because they were treated with antibiotics as an alternative to surgery, so diagnosis was not confirmed histologically. Of the 64 patients in the main analysis, 31 were recruited between 9 am and 5 pm, 24 between 5 pm and 10 pm, and nine between 10 pm and 9 am. Fifty eight patients travelled to the hospital by car and six by ambulance, of whom five had pain over speed bumps and a final diagnosis of appendicitis and one had no pain over speed bumps and no appendicitis.

Table 1⇓ shows pain over speed bumps in relation to diagnosis of appendicitis. Fifty four of 64 participants were "speed bump positive." Thirty four participants had a confirmed diagnosis of appendicitis, 33 of whom had worsened pain over speed bumps, giving a sensitivity of 97% (85% to 100%) and a specificity of 30% (15% to 49%). The positive predictive value was 61% (47% to 74%), and the negative predictive value was 90% (56% to 100%). The likelihood ratios were 1.4 (1.1 to 1.8) for a positive test result and 0.1 (0.0 to 0.7) for a negative result. Table 2⇓ shows how this compares with other clinical variables commonly used for diagnosis of appendicitis and also assessed in our sample.

Seven patients who were "speed bump positive" but did not have appendicitis had other important abdominal diagnoses, such as a ruptured ovarian cyst or diverticulitis. A post hoc secondary analysis of the diagnostic accuracy of pain over speed bumps for the diagnosis of important abdominal pathology requiring treatment (including appendicitis) increased the sensitivity to 98% (87% to 100%) and the specificity to 39% (20% to 61%).

Thirty three patients did not recall having travelled over speed bumps. A sensitivity analysis classifying those patients who did not recall travelling over speed bumps as having no pain over speed bumps had the effect of decreasing the sensitivity to 77% (61% to 88%) and increasing the specificity to 61% (47% to 74%), with a positive likelihood ratio of 2.0 (1.4 to 2.9) and a negative likelihood ratio of 0.4 (0.2 to 0.7) (see web extra data).

Discussion

Our results confirm that an increase in pain while travelling over speed bumps is associated with an increased likelihood of acute appendicitis. Absence of pain over speed bumps is associated with a significantly decreased likelihood of appendicitis. Although the specificity was relatively low, as a diagnostic variable pain over speed bumps compared favourably with other features commonly used in diagnostic assessment,

with a better sensitivity and negative likelihood ratio than all other features assessed. Moreover, some patients who were "speed bump positive" but did not have appendicitis had other important abdominal diagnoses, such as a ruptured ovarian cyst, diverticulitis, or pelvic inflammatory disease. We hypothesise that the worsening of pain when travelling over speed bumps in appendicitis may occur because the movement involved irritates the peritoneum in a similar way to that produced by testing for rebound tenderness on examination.

Strengths and limitations of study

Strengths of our study include the standardised approach to gathering information from patients by using a questionnaire and the obtaining of this information early in their admission and thus soon after their journey. A potential weakness is that although we recruited 101 patients as planned from our sample size calculation, only 68 recalled having travelled over speed bumps, a much lower rate than in our pilot study, which may be related to a redevelopment of the hospital site. Because of this, the number used for analysis (64 patients) was less than planned, leading to moderately large confidence intervals.

The presence of pain over speed bumps may have been overestimated in some patients owing to recall bias. Patients who had pain over speed bumps would be more likely to recall having travelled over them, whereas those who had no worsening of pain would not necessarily remember them. Although the sensitivity was 97% (table 2⇓) for patients who recalled speed bumps, because 33/97 (34%) patients did not travel (or did not recall travelling) over speed bumps, this diagnostic sign is not available in all patients and would therefore detect 77% (61% to 88%) of cases of appendicitis. This compares favourably with the other diagnostic features we assessed (see web extra data). Variable exposure to speed bumps would also occur in clinical practice, so ours is a pragmatic study that shows that pain over speed bumps can be a useful diagnostic sign when available, although availability will vary.

We used histological diagnosis of appendicitis as the reference standard for diagnosis. Three patients in our sample were treated with antibiotics for presumed appendicitis while waiting for surgery but went on to make a full recovery. A systematic review published during recruitment to our study has shown that antibiotics can lead to resolution of acute appendicitis.[9] We made the decision to exclude these patients from the analysis owing to the lack of a confirmed diagnosis, but a sensitivity analysis including these patients and classifying them in turn as positive or negative for a diagnosis of appendicitis made very little difference to overall results.

Comparison with other studies

Andersson (2004) did a meta-analysis of the diagnostic accuracy of clinical features of appendicitis.[6] Our finding of a negative likelihood ratio of 0.1 for pain over speed bumps in the diagnosis of appendicitis outperformed not only other clinical variables in our study (as shown in table 2⇓) but also those in Andersson's meta-analysis—migratory pain (0.52), nausea or vomiting (0.72), and rebound tenderness (0.39). Our positive likelihood ratio of 1.4 was similar to the findings of the meta-analysis for the above features. Another study, which also investigated the accuracy of various methods of diagnosis in 100 patients with possible appendicitis, found that the symptom of pain due to bumpiness in the road (which they termed the "cat's eye symptom") had a sensitivity of 80% and a specificity of 52%.[10] The "cat's eye symptom" had to be volunteered by the patient to be classed as positive, whereas in our study the response to speed bumps was solicited directly in a questionnaire. Our higher sensitivity of 97% but lower specificity of 30% may be related to the use of elicited rather than volunteered symptoms, for which one would predict exactly this difference in results.

Conclusions and implications

The high sensitivity of pain over speed bumps gives it a strong "rule-out value" and makes it a useful tool to use in excluding appendicitis and other important abdominal diagnoses. The low specificity, however, means that many patients with pain over speed bumps will not necessarily have appendicitis (that is, it is a poor "rule-in" test). Potential exists for it to be incorporated into clinical prediction rules for appendicitis. Our study was based in secondary care, so our results are not necessarily generalisable to a primary care population. However, pain over speed bumps could potentially have a useful role in primary care in assisting in the telephone assessment of patients with abdominal pain. As all our group of patients had already been assessed by a clinician who thought they might have appendicitis, the pre-test probability is quite high; the speed bump test might also be useful in assessment of all types of abdominal pain, not just when appendicitis is suspected. A history of pain on travelling over uneven road surfaces or potholes may provide a useful proxy for speed bumps in healthcare settings where speed bumps are less frequently found.

Although being "speed bump negative" offers some reassurance against a diagnosis of appendicitis, being "speed bump positive" certainly does not guarantee a diagnosis of appendicitis, so in this respect the myth is untrue. However, our findings suggest that questioning about speed bumps should form a routine part of the assessment of patients with possible appendicitis. Unanswered questions include whether the speed or manner of driving approach to a speed bump affects the diagnostic power.

We thank A K Allouni and S Kreckler for their contributions to the original idea for the project and pilot study. We also thank doctors in the surgical team at Stoke Mandeville Hospital for their assistance with recruitment.

Contributors: HFA developed the idea and designed the study, analysed and interpreted the data, and wrote the paper. ND'S and DK participated in study design, collected the data, and helped to write the paper. RJS assisted with study design, analysed and interpreted the data, and helped to write the paper. AHuang assisted with data interpretation and helped to write the paper. AHarnden supervised HFA, assisted with study design, and helped to write the paper. All authors approved the final version of the manuscript. HFA is the guarantor.

Funding: No formal funding. HFA is an academic clinical fellow, which is a scheme partially funded by the National Institute for Health Research. The University of Oxford acted as study sponsor and had no role in the study design, data collection, analysis or interpretation, writing of the paper, or decision to submit for publication. All authors were independent from funders and sponsors.

Competing interests: All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author) and declare: no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work.

Ethical approval: The study was approved by Oxford A Research Ethics Committee (reference 12/SC/0052). All participants gave informed consent before taking part.

Data sharing: Full data are available from the corresponding author on request. Consent for data sharing was not obtained, but the presented data are anonymised and the risk of identification is low.

What is already known on this topic

Speed bumps are a commonly used traffic calming device to reduce vehicle speeds

Clinical diagnosis of acute appendicitis can be difficult, and presence of various clinical features, such as migration of pain and rebound tenderness, can be used in assessment

Some doctors routinely ask about pain on travelling over speed bumps as part of their clinical assessment, but no evidence base exists for this

What this study adds

Pain on travelling over speed bumps had a high sensitivity (97%) but a low specificity (30%) for the diagnosis of appendicitis

It compared favourably with other clinical features used in diagnosis of appendicitis, and therefore provides a useful addition, particularly in terms of excluding appendicitis

It may also be useful for the diagnosis of other important abdominal conditions, and its use could be extended to all presentations of the "acute abdomen"

1 Department for Transport. Traffic calming. Local Transport Note 01/07, 2007.
2 Towner EML, Dowswell T, Mackereth C. What works in preventing unintentional injuries in children and young adolescents? An updated systematic review. Health Development Agency, 2001.
3 Bunn F, Collier T, Frost C, Ker K, Roberts I, Wentz R. Area-wide traffic calming for preventing traffic related injuries. *Cochrane Database Syst Rev* 2003;(1):CD003110.
4 Humes DJ, Simpson J. Acute appendicitis. *BMJ* 2006;333:530-4.
5 Bickell NA, Aufses AH Jr, Rojas M, Bodian C. How time affects the risk of rupture in appendicitis. *J Am Coll Surg* 2006;202:401-6.
6 Andersson RE. Meta-analysis of the clinical and laboratory diagnosis of appendicitis. *Br J Surg* 2004;91:28-37.
7 Walker SJ, West CR, Colmer MR. Acute appendicitis: does removal of a normal appendix matter, what is the value of diagnostic accuracy and is surgical delay important? *Ann R Coll Surg Engl* 1995;77:358-63.
8 Lewis SR, Mahony PJ, Simpson J. Appendicitis. *BMJ* 2011;343:d5976.
9 Varadhan KK, Neal KR, Lobo DN. Safety and efficacy of antibiotics compared with appendicectomy for treatment of uncomplicated acute appendicitis: meta-analysis of randomised controlled trials. *BMJ* 2012;344:e2156.
10 Golledge J, Toms AP, Franklin IJ, Scriven MW, Galland RB. Assessment of peritonism in appendicitis. *Ann R Coll Surg Engl* 1996;78:11-4.

Accepted: 16 November 2012

Cite this as: *BMJ* 2012;345:e8012

Tables

Table 1| Pain over speed bumps in relation to appendicitis

	Appendicitis		
Pain over speed bumps	**Positive**	**Negative**	**Total**
Positive	33	21	54
Negative	1	9	10
Total	34	30	64

Table 2| **Diagnostic performance (with 95% CI) of pain over speed bumps compared with other clinical diagnostic variables for appendicitis**

Diagnostic variable	Sensitivity (%)	Specificity (%)	Positive predictive value (%)	Negative predictive value (%)	Positive likelihood ratio	Negative likelihood ratio
Pain over speed bumps	97 (85 to 100)	30 (15 to 49)	61 (47 to 74)	90 (56 to 100)	1.4 (1.1 to 1.8)	0.1 (0.0 to 0.7)
Migratory pain	65 (46 to 80)	33 (17 to 53)	52 (36 to 68)	45 (24 to 68)	1.0 (0.7 to 1.4)	1.1 (0.5 to 2.1)
Nausea or vomiting	79 (62 to 91)	17 (5.6 to 35)	52 (38 to 66)	42 (15 to 72)	1.0 (0.8 to 1.2)	1.2 (0.4 to 3.5)
Rebound tenderness	71 (53 to 85)	50 (31 to 69)	62 (45 to 77)	60 (39 to 79)	1.4 (0.9 to 2.2)	0.6 (0.3 to 1.1)

Figure

[Image: Ian Williams]

Critical Appraisal Skills Programme (CASP)

making sense of evidence

10 questions to help you make sense of randomised controlled trials

How to use this appraisal tool

Three broad issues need to be considered when appraising the report of a randomised controlled trial:

- **Is the trial valid?**
- **What are the results?**
- **Will the results help locally?**

The 10 questions on the following pages are designed to help you think about these issues systematically.

The first two questions are screening questions and can be answered quickly. If the answer to both is "yes", it is worth proceeding with the remaining questions.

You are asked to record a "yes", "no" or "can't tell" to most of the questions. A number of italicised prompts are given after each question.

These are designed to remind you why the question is important. Record your reasons for your answers in the spaces provided.

The 10 questions are adapted from Guyatt GH, Sackett DL, and Cook DJ, Users' guides to the medical literature. II. How to use an article about therapy or prevention. *JAMA* 1993; 270 (21): 2598-2601 and *JAMA* 1994; 271(1): 59-63

Screening Questions

1. Did the study ask a clearly-focused question? ❑ Yes ❑ Can't tell ❑ No

Consider if the question is 'focused' in terms of:

– the population studied

– the intervention given

– the outcomes considered

---

2. Was this a randomised controlled trial (RCT) and was it appropriately so? ❑ Yes ❑ Can't tell ❑ No

Consider:

– why this study was carried out as an RCT

– if this was the right research approach for the question being asked

---

Is it worth continuing?

Detailed Questions

3. Were participants appropriately allocated to intervention and control groups? ❑ Yes ❑ Can't tell ❑ No

Consider:

– how participants were allocated to intervention and control groups. Was the process truly random?

– whether the method of allocation was described. Was a method used to balance the randomization, e.g. stratification?

– how the randomization schedule was generated and how a participant was allocated to a study group

– if the groups were well balanced. Are any differences between the groups at entry to the trial reported?

– if there were differences reported that might have explained any outcome(s) (confounding)

4. Were participants, staff and study personnel 'blind' to participants' study group? ❑ Yes ❑ Can't tell ❑ No

Consider:

– the fact that blinding is not always possible

– if every effort was made to achieve blinding

– if you think it matters in this study

– the fact that we are looking for 'observer bias'

5. Were all of the participants who entered the trial accounted for at its conclusion? ❑ Yes ❑ Can't tell ❑ No

Consider:

– if any intervention-group participants got a control-group option or vice versa

– if all participants were followed up in each study group (was there loss-to-follow-up?)

– if all the participants' outcomes were analysed by the groups to which they were originally allocated (intention-to-treat analysis)

– what additional information would you liked to have seen to make you feel better about this

6. Were the participants in all groups followed up and data collected in the same way? ❑ Yes ❑ Can't tell ❑ No

Consider:

– if, for example, they were reviewed at the same time intervals and if they received the same amount of attention from researchers and health workers. Any differences may introduce performance bias.

7. Did the study have enough participants to minimise the play of chance? ❑ Yes ❑ Can't tell ❑ No

Consider:

– if there is a power calculation. This will estimate how many participants are needed to be reasonably sure of finding something important (if it really exists and for a given level of uncertainty about the final result).

8. How are the results presented and what is the main result?

Consider:

– if, for example, the results are presented as a proportion of people experiencing an outcome, such as risks, or as a measurement, such as mean or median differences, or as survival curves and hazards

– how large this size of result is and how meaningful it is

– how you would sum up the bottom-line result of the trial in one sentence

9. How precise are these results?

Consider:

– if the result is precise enough to make a decision

– if a confidence interval were reported. Would your decision about whether or not to use this intervention be the same at the upper confidence limit as at the lower confidence limit?

– if a p-value is reported where confidence intervals are unavailable

10. Were all important outcomes considered so the results can be applied?

❑ Yes ❑ Can't tell ❑ No

Consider whether:

– the people included in the trail could be different from your population in ways that would produce different results

– your local setting differs much from that of the trial

– you can provide the same treatment in your setting

Consider outcomes from the point of view of the:

– individual

– policy maker and professionals

– family/carers

– wider community

Consider whether:

– any benefit reported outweighs any harm and/or cost. If this information is not reported can it be filled in from elsewhere?

– policy or practice should change as a result of the evidence contained in this trial

Rapid critical appraisal of a diagnostic test accuracy study

Step 1: What question did the study ask?

Population/problem: .....

Index case: .....

Comparison: .....

Outcome(s): .....

Step 2: How well was the study done? (internal validity)

Recruitment — Was the diagnostic test evaluated in a representative spectrum of patients (like those in whom it would be used in practice)?

What is best?	**Where do I find the information?**
It is ideal if the diagnostic test is applied to the full spectrum of patients — those with mild, severe, early and late cases of the target disorder. It is also best if the patients are randomly selected or consecutive admissions so that selection bias is minimised.	The **Methods** section should tell you how patients were enrolled and whether they were randomly selected or consecutive admissions. It should also tell you where patients came from and whether they are likely to be representative of the patients in whom the test is to be used.

This paper: Yes ☐ No ☐ Unclear ☐ Comment: .....

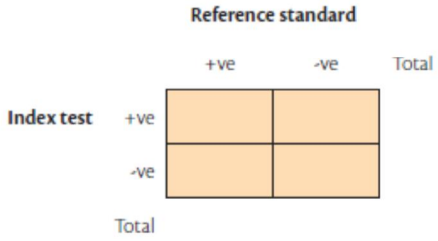
Maintenance — Was the endpoint of the reference standard obtained for all the subjects?

What is best?	**Where do I find the information?**
The endpoint of the reference standard (ie whether the subjects are positive or negative for the condition) should be measured for all the subjects. In cases where this depends on the follow-up of people for a period of time (dependent on the disease in question) to see whether they are truly negative, this follow-up should be long enough to be certain of the outcome.	The **Methods** section should indicate whether the endpoint of the reference standard was obtained for all subjects.

This paper: Yes ☐ No ☐ Unclear ☐ Comment: .....

***M**easurement* — Were the assessors kept **b**lind to the results of each test and/or were the reference standard endpoints **o**bjective?	
What is best?	**Where do I find the information?**
The reference standard and the index test being assessed should be applied to each patient independently and blindly. Those who interpreted the results of one test should not be aware of the results of the other test. Finally, the paper should also have sufficient description of the index test to allow its replication and also interpretation of the results.	The **Methods** section should describe who conducted the two tests and whether each was conducted independently and blinded to the results of the other. The **Methods** section should describe the tests in detail.
This paper: Yes ☐ No ☐ Unclear ☐ Comment:	

Step 3: What do the results mean?



Measure	Result
Sensitivity (Sn)	
Specificity (Sp)	
Positive predictive value (PPV)	
Negative predictive value (NPV)	

دانشگاه علوم پزشکی تبریز

دانشکده پزشکی– گروه پزشکی اجتماعی

فرم ارزیابی نقادانه مقالات کارآزمایی بالینی

(Critical Appraisal Tool for clinical trials)

مرحله 1: سوال مطالعه چه بود؟

P:

I:

C:

O:

مرحله 2: مطالعه تا چه حد درست انجام شده؟(بررسی سوگرایی، عوامل مخدوش کننده ، اعتبار درونی مطالعه) **RAMMbo**

Recruitment /Representative:R آیا نمونه مورد مطالعه نماینده

جامعه مورد نظر است؟ بلی☐ خیر☐ مشخص نشده است☐

نظر کلی:

در مقاله بهتر است چه مواردی مشخص شود؟	در کدام قسمت مقاله باید به دنبال جواب این سوال باشیم؟
محل انجام مطالعه(جامعه، کلینیک تخصصی، منطقه خاصی از یک شهر ، نمونه ها از تمام مناطق شهر و تمام کلینیک ها؟) معیار های ورود و خروج در مطالعه مشخص شده اند؟نمونه ها بطور تصادفی انتخاب شده اند؟یا بیماران به طور داوطلبانه وارد مطالعه شده اند؟	در بخش روش کار یا method در قسمت setting و participants یاpatients

Allocation /Adjustment:A آیا تخصیص افراد به دوگروه تصادفی و

مخفی بوده است؟بلی☐ خیر☐ مشخص نشده است☐

نظر کلی:

به جدول صفحه بعد مراجعه کنید:

در مقاله بهتر است چه مواردی مشخص شود؟	در کدام قسمت مقاله باید به دنبال جواب این سوال باشیم؟
تخصیص تصادفی به دو گروه مداخله توسط برنامه های کامپیوتری ترجیحا کارآزمایی های بالینی به صورت چند مرکزی انجام شده باشد. Blocked Randomization انجام شده باشد.در مورد پلاسبو هم باید توضیح داده شود که از نظر ظاهری مشابه داروی مداخله باشد.یا در مورد shamکه مشابه اقدام یا پروسیجر مداخله باشد. سوال بعدی در این قسمت این است که آیا گروه های تحت مطالعه در ابتدای مطالعه با هم مشابه بوده اند؟ هرچه گروه های مداخله و مقایسه در ابتدای مطالعه بیشتر شبیه به هم باشند بهتر است.	در قسمت روش کارmethod، به این دقت می نیم که آیا اشاره ای به مخفی بودن تخصیص تصادفی شده است؟ Concealment؟ به جدول خصوصیات پایه ای بیماران مراجعه می کنیمBasic characteristicsکه دو گروه را از نظر برخی عوامل مانند سن، جنس ، عوامل خطر با هم مقایسه کرده است.

M: Maintenance آیا گروه ها طی مطالعه یکسان مانده اند؟بلی☐ خیر☐

مشخص نشده است☐ نظر کلی:

در مقاله بهتر است چه مواردی مشخص شود؟	در کدام قسمت مقاله باید به دنبال جواب این سوال باشیم؟
گروه ها باید از هر نظر بجز مداخله با هم یکسان باشند. آیا مدت پیگیری follow up کافی بوده است؟ افرادی که طی مدت مطالعه از دست می روند(از مطالعه خارج می شوند یا به هر علتی تا پایان مطالعه نمی مانند) بهتر است کمتر از 20٪ باشد بیماران در همان گروهی که در ابتدا به آن تعلق داشته اند باید محاسبه شوند(مخرج کسر تعداد اولیه گروه باشد.قبل از ترک مطالعه) Intention to treat Analysis	در قسمت روش کاردر پروتکل های درمانی برای هر گروه (درمان هایی غیر از مداخله که بر هرگزوه داده شده است. در قسمت نتایج: در ابتدا چند نفر به طور تصادفی تقسیم شدند؟ چند نفر مورد آنالیز قرار گرفتند؟ این موارد معمولا در فلوچارت ها در مقاله نشان داده می شوند.

Measurement(blinding, Objective outcome) :M(bo)

آیا افراد تحت مطالعه و محققینی که نتایج را ارزیابی می کنندنسبت به قرار گیری افراد در گروه ها کورسازی شده اند؟ (نمی دانند کدام داروی اصلی و کدام پلاسبو گرفته) آیا نتایج مورد بررسی بصورت عینی است ؟بلی□ خیر□ مشخص نشده است□ نظر کلی:

در مقاله بهتر است چه مواردی مشخص شود؟	در کدام قسمت مقاله باید به دنبال جواب این سوال باشیم؟
برای نتایجو پیامد های عینی مانند مرگ ، کورسازیblinding کمتر اهمیت دارد ولی برای نتایج ذهنی (مانند احساس درد، عملکرد یا رضایت)کورسازی لازم است	در قسمت روش کار که در کورد سنجش پیامد (outcome) توضیح داده شده است.

S100A8/A9: A Potential New Diagnostic Aid for Acute Appendicitis

John F. Bealer, MD and Mark Colgin, PhD

Abstract

Objectives: Diagnosing acute appendicitis is a daunting clinical challenge, as there is no single test that reliably distinguishes acute appendicitis from other etiologies of acute abdominal pain. In this study, the authors examined whether circulating levels of S100A8/A9 could be useful as a marker to aid in the diagnosis of acute appendicitis.

Methods: Plasma samples from emergency department (ED) patients with acute abdominal pain (n = 181) were tested using an immunoassay for S100A8/A9.

Results: The sensitivity and specificity for S100A8/A9 in diagnosing acute appendicitis were estimated to be 93% (95% confidence interval [CI] = 81% to 97%) and 54% (95% CI = 45% to 62%), respectively. Negative predictive value (NPV) was 96% (95% CI = 89% to 99%), and positive predictive value (PPV) was 37% (95% CI = 28% to 47%). Performance characteristics of elevated white blood cell (WBC) count were also estimated: sensitivity 63% (95% CI = 47% to 76%), specificity 67% (95% CI = 59% to 75%), NPV 86% (95% CI = 78% to 91%), and PPV 36% (95% CI = 26% to 47%).

Conclusions: This is the first report exploring the relationship between circulating S100A8/A9 and acute appendicitis and establishes proof of concept for this biomarker as a diagnostic test for acute appendicitis. Further studies are indicated to optimize the use of this biomarker, in conjunction with other established approaches.

ACADEMIC EMERGENCY MEDICINE 2010; 17:333–336 © 2010 by the Society for Academic Emergency Medicine

Keywords: S100A8/A9, calprotectin, appendicitis, diagnosis

Acute appendicitis is the most common surgical condition that produces abdominal pain. The diagnostic differentiation of acute appendicitis from other causes of abdominal pain has increasingly utilized computed tomography (CT) scanning. However, a number of authors have raised significant questions regarding the utility of CT in this clinical setting.[1–4] Ideally, CT would only be used as a diagnostic test for appendicitis, after a screening test has defined a patient at risk. Given these concerns, it would be desirable to have a noninvasive and inexpensive approach to ruling out acute appendicitis, which could eventually be utilized as a screening tool. S100A8 (also named calgranulin A; myeloid-related protein 8 [MRP8]) and S100A9 (calgranulin B; MRP14) are intracellular calcium-binding proteins, which are key to the transduction of calcium signaling during inflammation.[5] These two independent proteins have a tissue-specific pattern of expression and readily form dimers that, when combined, are commonly known as either S100A8/A9 or calprotectin.[6] S100A8/A9 is constitutively expressed in neutrophils, monocytes, some epithelial cells, and the keratinocytes of inflamed tissues, while not generally expressed in tissue macrophages or lymphocytes.[5,6] Most of S100A8/A9's proinflammatory functions require extracellular release, but the exact secretory mechanism is not fully understood. What is known about this mechanism is that the secretion of S100A8/A9 *in vivo* is tightly controlled and requires concomitant activation of two independent signal pathways. Signal one is activation of protein kinase C that can be induced by many different inflammatory stimuli.[7] The second signal is provided by contact of phagocytes with activating surfaces, such as extracellular matrix (ECM) proteins or tumor necrosis factor (TNF)-stimulated endothelium, but not by interaction with resting endothelial cells. Thus, secretion of S100A8/A9 is restricted to the sites of monocyte–endothelial or

From AspenBio Pharma, Inc., Castle Rock, CO.

Received May 15, 2009; revision received August 10, 2009; accepted August 11, 2009.

Presented at the 2009 Society for Academic Emergency Medicine annual meeting, New Orleans, LA, May 2009.

Drs. Bealer and Colgin have financial relationships with AspenBio Pharma, Inc.: Dr. Colgin is the Chief Scientific Officer and Dr. Bealer is a consultant for the company.

AspenBio Pharma, Inc., is the company that developed the enzyme-linked immunosorbent assay tested in this article.

Address for correspondence and reprints: Mark Colgin, PhD; e-mail: mcolgin@aspenbiopharma.com.

doi: 10.1111/j.1553-2712.2010.00663.x
ISSN 1069-6563
PII ISSN 1069-6563583

monocyte–ECM interactions during inflammatory conditions.[7,8] The overexpression of S100A8 and S100A9 at these types of inflammatory sites are well recognized and there is growing evidence that S100A8/A9 could be a biomarker for a number of inflammatory conditions.[5,6,9]

In this study, we hypothesized that circulating levels of S100A8/A9 (calgranulin A and B, MRP8/14, calprotectin) are increased in acute appendicitis. S100A8/A9 is a demonized calcium-binding protein of the S100 family that has a proven and pivotal role in gastrointestinal inflammation and could potentially differentiate acute appendicitis from noninflammatory causes of acute abdominal pain. The aim of this pilot study was to provide a preliminary characterization of the clinical performance characteristics of a blood test for S100A8/A9.

METHODS

Study Design

This was a prospective pilot study of human serum samples investigating the potential relationship between plasma levels of S100A8/A9 and acute appendicitis. The study protocol was approved and monitored by the individual institutional review boards at the participating hospitals.

Study Setting and Population

This study was conducted in the emergency departments (EDs) of three separate community hospitals. The study population of 181 patients included both adults and children presenting with acute abdominal pain. Inclusion criteria were defined as follows: 1) chief complaint was abdominal pain, 2) chief complaint was a new complaint for the patient, 3) duration of symptoms was less than 2 weeks, and 4) pain was located primarily on the right side of the body (iliac fossa) and/or primarily below the umbilicus. Exclusion criteria were as follows: 1) not meeting the above inclusion criteria; 2) patient had dysuria with burning, stinging, or itching at the urethral meatus; 3) a history of end-stage or metastatic cancer; 4) a history of recent trauma; and 5) previous appendectomy.

Study Protocol

Between February 2008 and May 2008, patients complaining of right lower quadrant abdominal pain presenting to the study EDs were offered entry into our study. After informed consent was obtaining, patients were enrolled and received clinical care as directed by the treating physicians. As part of their clinical management, white blood cell (WBC) counts were determined for most patients, although the protocol did not require any specific diagnostic approach. In addition, all of these patients had a separate blood sample drawn for the determination of plasma S100A8/A9 levels. The S100A8/A9 concentration was determined using a sandwich enzyme-linked immunosorbent assay (Aspen-Bio Pharma, Inc., Castle Rock, CO), read by the Bio-Rad model 680 microplate reader (Bio-Rad Laboratories, Inc., Hercules, CA). A preliminary normal range was determined by measuring S100A8/A9. The cutoff for this analysis was determined retrospectively based on a receiver operating characteristic (ROC) analysis of this patient population. The results of these tests were collected, but not reported to the physicians. Clinical outcome was defined on the basis of two separate data sources obtained following the ED visit. The first data source was histologic evaluation of the appendix specimen in those patients who underwent an appendectomy. For those patients who did not have an appendectomy, the second source of information was a telephone interview performed 1–4 weeks following presentation to the ED. Patients who had not undergone appendectomy and those with normal appendix histology were defined as not having appendicitis, while those with pathologic confirmation of the disease were defined as having appendicitis.

Data Analysis

Performance characteristics (to predict acute appendicitis) were estimated and 95% confidence intervals (CIs) were calculated using SAS (SAS Institute, Cary, NC). Clinical performance characteristics of S100A8/A9 were evaluated using ROC data. Given that the goal of S100A8/A9 is to help identify true positives, the ROC curve generated indicated a potential cutoff value between 14–20 units (Figure 1). The WBC cutoff was $10 \times 10^3/\mu L$.

RESULTS

A total of 181 patients were evaluated in the study; 118 patients were female and 63 were male. The mean age was 31.5 years, with a range of 8–76 years. The prevalence of appendicitis was 22.7% (41/181) for the patient population. Idiopathic abdominal pain accounted for 84 patients (46%), and other significant diagnoses were found in 56 patients (31%). There were 31 patients who had an appendectomy, with a negative appendectomy rate of 9.7% (3/31). Table 1 presents performance characteristics of S100A8/A9 with a cutoff of 20 units.

DISCUSSION

This study shows a correlation between circulating S100A8/A9 and acute appendicitis, suggesting that the test has the potential to be sensitive for detecting the disease in patients with acute, right-lower-quadrant abdominal pain.

While there are no previous reports linking circulating S100A8/A9 and appendicitis, there is growing experience using this marker to detect other gastrointestinal inflammatory conditions. For example, fecal S100A8/A9 levels are elevated in patients with inflammatory bowel disease (IBD) and these levels are used clinically to quantify intestinal inflammation.[10–16] The increased S100A8/A9 in fecal samples of IBD patients is partially explained by an excellent correlation between fecal excretion of indium-111–labeled granulocytes and fecal S100A8/A9 in these patients, suggesting that fecal S100A8/A9 is a reflection of granulocyte migration through the gut wall into the feces.[17] Other studies further support the connection between S100A8/A9 and the gut. For example, global gene expression profiles of

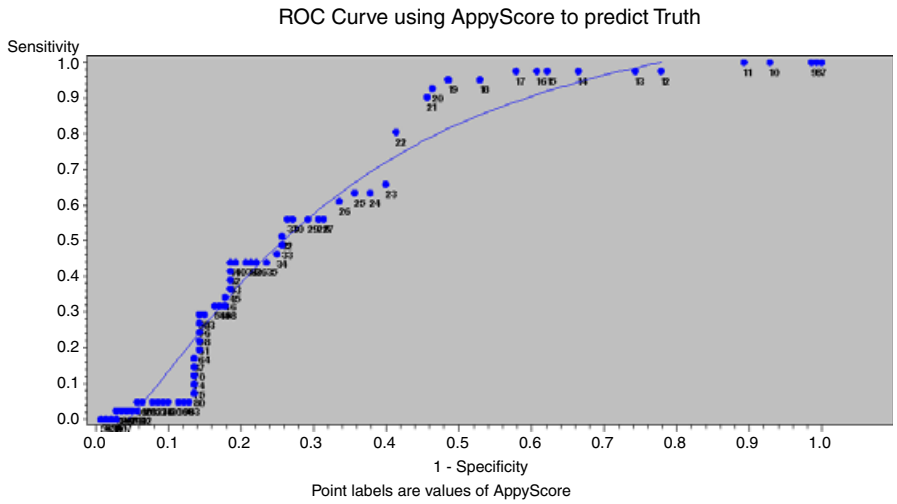


Figure 1. ROC curve. *Point labels are values of S100A8/A9. The area under the ROC curve is 0.71. ROC = receiver operating characteristic.

Table 1
Performance Characteristics Estimate

Diagnostic Method	Negative Likelihood Ratio	Positive Likelihood Ratio	NPV	PPV	Sensitivity	Specificity
WBC count (n = 178)*	0.56 (0.37–0.84)	1.92 (1.37–2.69)	86.1 (78.34–91.4)	35.7 (25.5–47.41)	62.5 (47.03–75.78)	67.4 (59.19–74.65)
S100A8/A9 (n = 181)	0.14 (0.04–0.39)	2.00 (1.64–2.43)	96.2 (89.29–98.68)	36.9 (28.2–46.53)	92.7 (80.57–97.48)	53.6 (45.33–61.63)

S100A8/A9 cut-off = 20 units; WBC cutoff = $10 \times 10^3/\mu L$; 95% CIs are in parentheses.
NPV = negative predictive value; PPV = positive predictive value; WBC = white blood cell.
*WBC count was not available in patient medical records for three subjects.

inflamed colonic tissue using DNA microarrays have shown S100A9 to be highly up-regulated compared to healthy tissue.[18] In addition, several immunohistologic studies of inflamed bowel have confirmed an elevated expression of S100A8 and S100A9 by infiltrating monocytes, neutrophils, and epithelial cells.[19] Similar to our findings in appendicitis, there is also a good correlation between S100A8 and S100A9 serum concentrations and IBD.[9] This evidence, including our findings between S100A8/A9 and appendicitis, supports a potentially powerful role for S100A8/A9 as a biomarker of gastrointestinal disease.

Our study suggests that the known physiology of S100A8/A9 in gut inflammation also applies to appendicitis. While not a specific marker of appendicitis, S100A8/A9's dual-activation pathway fits well with the known pathogenesis of appendicitis. In classically described appendicitis, luminal obstruction leads to increased transmural pressure that ultimately leads to venous occlusion and ischemia of the appendix. It is possible that ischemia would affect and damage the mucosa of the appendix first, as it is the most oxygen-sensitive tissue, and expose infiltrating neutrophils to ECM proteins or TNF-stimulated endothelium. If true, this would suggest that S100A8/A9 is being secreted at the earliest stages of appendicitis and may help explain the sensitivity we demonstrated in our study. It would also suggest that secretion of S100A8/A9 may increase in a crescendo fashion as appendicitis progresses, because greater disease progression would lead to greater tissue damage with increasing amounts of ECM exposure and neutrophil infiltration. This could suggest a role for S100A8/A9 in quantifying the severity of appendicitis far more accurately than the current nonspecific descriptive terminology.

WBC and C-reactive protein (CRP) measurements have been described as having some utility in the diagnosis of acute appendicitis. However, the sensitivity of these two tests to identify patients with acute appendicitis is relatively low.[20] Past studies have found that increased levels of CRP show a greater relationship to the severity of appendicitis than that of increased WBC levels.[21] Both tests show a correlation to acute appendicitis, although ultimately, the sensitivity of these tests is insufficient to achieve reliable rule-out.

LIMITATIONS

There are limitations to this study that should be taken into consideration. First, this was a pilot study representing an initial investigation into the relationship between appendicitis and S100A8/A9. We did not

prospectively require any specific diagnostic work-up. In addition, our patient population was fairly homogenous, being representative of the private suburban hospitals where the study was conducted. This limits our ability to make comments about specific demographic subpopulations, such as women and children, or how the test would perform among other patient populations in different hospital settings.

CONCLUSIONS

We report the first use of circulating S100A8/A9 to aid in the diagnosis of appendicitis, with a sensitivity of 93%, a specificity of 54%, and a positive predictive value (PPV) and a negative predictive value (NPV) of 39 and 96%, respectively, in this pilot study. If proven safe and effective in a statistically sized study, S100A8/A9 could possibly be a new diagnostic to aid in the work-up of appendicitis. Our findings have encouraged us to design a larger study, sized sufficiently to provide the statistical power to make specific conclusions about the utility of the test.

References

1. Partrick DA, Janik JE, Janik JS, Bensard DD, Karrer FM. Increased CT scan utilization does not improve the diagnostic accuracy of appendicitis in children. J Pediatr Surg. 2003; 38:659–62.
2. Flum DR, McClure TD, Morris A, Koepsell T. Misdiagnosis of appendicitis and the use of diagnostic imaging. J Am Coll Surg. 2005; 201:933–9.
3. Flum DR, Morris A, Koepsell T, Dellinger EP. Has misdiagnosis of appendicitis decreased over time? A population-based analysis. JAMA. 2001; 286: 1748–53.
4. Morris KT, Kavanagh M, Hansen P, Whiteford MH, Deveney K, Standage B. The rational use of computed tomography scans in the diagnosis of appendicitis. Am J Surg. 2002; 183:547–50.
5. Passey RJ, Xu K, Hume DA, Geczy CL. S100A8: emerging functions and regulation. J Leukoc Biol. 1999; 66:549–56.
6. Hessian PA, Edgeworth J, Hogg N. MRP-8 and MRP-14, two abundant Ca(2+)-binding proteins of neutrophils and monocytes. J Leukoc Biol. 1993; 53:197–204.
7. Rammes A, Roth J, Goebeler M, Klempt M, Hartmann M, Sorg C. Myeloid-related protein (MRP) 8 and MRP14, calcium-binding proteins of the S100 family, are secreted by activated monocytes via a novel, tubulin-dependent pathway. J Biol Chem. 1997; 272:9496–502.
8. Frosch M, Strey A, Vogl T, et al. Myeloid-related proteins 8 and 14 are specifically secreted during interaction of phagocytes and activated endothelium and are useful markers for monitoring disease activity in pauciarticular-onset juvenile rheumatoid arthritis. Arthritis Rheum. 2000; 43:628–37.
9. Striz I, Trebichavsky I. Calprotectin - a pleiotropic molecule in acute and chronic inflammation. Physiol Res. 2004; 53:245–53.
10. Costa F, Mumolo MG, Bellini M, et al. Role of faecal calprotectin as non-invasive marker of intestinal inflammation. Dig Liver Dis. 2003; 35:642–7.
11. Roseth AG, Aadland E, Grzyb K. Normalization of faecal calprotectin: a predictor of mucosal healing in patients with inflammatory bowel disease. Scand J Gastroenterol. 2004; 39:1017–20.
12. Roseth AG, Aadland E, Jahnsen J, Raknerud N. Assessment of disease activity in ulcerative colitis by faecal calprotectin, a novel granulocyte marker protein. Digestion. 1997; 58:176–80.
13. Roseth AG, Fagerhol MK, Aadland E, Schjonsby H. Assessment of the neutrophil dominating protein calprotectin in feces. A methodologic study. Scand J Gastroenterol. 1992; 27:793–8.
14. Tibble JA, Bjarnason I. Fecal calprotectin as an index of intestinal inflammation. Drugs Today (Barc). 2001; 37:85–96.
15. Tibble JA, Sigthorsson G, Bridger S, Fagerhol MK, Bjarnason I. Surrogate markers of intestinal inflammation are predictive of relapse in patients with inflammatory bowel disease. Gastroenterology. 2000; 119:15–22.
16. Tibble JA, Sigthorsson G, Foster R, Forgacs I, Bjarnason I. Use of surrogate markers of inflammation and Rome criteria to distinguish organic from nonorganic intestinal disease. Gastroenterology. 2002; 123:450–60.
17. Roseth AG, Schmidt PN, Fagerhol MK. Correlation between faecal excretion of indium-111-labelled granulocytes and calprotectin, a granulocyte marker protein, in patients with inflammatory bowel disease. Scand J Gastroenterol. 1999; 34:50–4.
18. Lawrance IC, Fiocchi C, Chakravarti S. Ulcerative colitis and Crohn's disease: distinctive gene expression profiles and novel susceptibility candidate genes. Hum Mol Genet. 2001; 10:445–56.
19. Schmid KW, Lugering N, Stoll R, et al. Immunohistochemical demonstration of the calcium-binding proteins MRP8 and MRP14 and their heterodimer (27E10 antigen) in Crohn's disease. Hum Pathol. 1995; 26:334–7.
20. Beltran MA, Almonacid J, Vicencio A, Gutierrez J, Cruces K, Cumsille M. Predictive value of white blood cell count and C-reactive protein in children with appendicitis. J Pediatr Surg. 2007; 42:1208–14.
21. Okamoto T, Sano K, Ogasahara K. Receiver-operating characteristic analysis of leukocyte counts and serum C-reactive protein levels in children with advanced appendicitis. Surg Today. 2006; 36:515–8.

Critical Appraisal Skills Programme (CASP)

making sense of evidence

12 questions to help you make sense of a diagnostic test study

How to use this appraisal tool

Three broad issues need to be considered when appraising a diagnostic test:

- **Are the results of the study valid?**
- **What are the results?**
- **Will the results help me and my patients/population?**

The 12 questions on the following pages are designed to help you think through these issues systematically.

The first two questions are screening questions and can be answered quickly. If the answer to both is "yes", it is worth proceeding with the remaining questions.

You are asked to record a "yes", "no" or "can't tell" to most of the questions.

A number of italicised prompts are given after each question. These are designed to remind you why the question is important. Record your reasons for your answers in the spaces provided.

The 12 questions are adapted from Jaesche R, Guyatt GH, Sackett DL, Users' guides to the medical literature, VI. How to use an article about a diagnostic test. JAMA 1994; 271 (5): 389-391

A/ Are the results of the study valid?

Screening Questions

1. Was there a clear question for the study to address? ❑ Yes ❑ Can't tell ❑ No

A question should include information about:

- *the population*
- *the test*
- *the setting*
- *the outcomes*

2. Was there a comparison with an appropriate reference standard? ❑ Yes ❑ Can't tell ❑ No

HINT: Is this reference test(s) the best available indicator in the circumstances?

Is it worth continuing?

Detailed Questions

3. Did all patients get the diagnostic test and the reference standard? ❑ Yes ❑ Can't tell ❑ No

Consider:

- *Were both received regardless of the Results of the test of interest?*
- *Check the 2 x 2 table (Verification bias)*

4. Could the results of the test of interest have been influenced by the results of the reference standard? ❑ Yes ❑ Can't tell ❑ No

Consider:

- *Was there blinding?*
- *Were the tests performed independently? (Review bias)*

5. Is the disease status of the tested population clearly described? ❑ Yes ❑ Can't tell ❑ No

Consider:

- *Presenting symptoms*
- *Disease stage or severity*
- *Co-morbidity*
- *Differential diagnoses (Spectrum bias)*

6. Were the methods for performing the test Described in sufficient detail? ❑ Yes ❑ Can't tell ❑ No

HINT: Was a protocol followed?

Is it worth continuing?

B/If so, what are the results?

7. What are the results?

Consider:

- *Are the sensitivity and specificity and/or likelihood ratios presented?*
- *Are the results presented in such a way that we can work them out?*

8. How sure are we about these results?

Consider:

- *Could they have occurred by chance?*
- *Are there confidence limits?*
- *What are they?*

C/ Will the results help me and my patients/population?

(Consider whether you are primarily interested in the impact on a population or individual level)

9. Can the results be applied to your patients/ the population of interest? ❑ Yes ❑ Can't tell ❑ No

HINT: Do you think you patients / population are so different from those in the study that the results connot be applied? Such as age, sex, ethnicity and spectrum bias.

10. Can the test be applied to your patient or population of interest? ❑ Yes ❑ Can't tell ❑ No

Consider:

- *Think of resources and opportunity costs*
- *Level and availability of expertise required to interpret the tests*
- *Current practice and availability of services*

11. Were all outcomes important to the Individual or population considered? ❑ Yes ❑ Can't tell ❑ No

Consider:

- *Will the knowledge of the test result improve patient wellbeing*
- *Will the knowledge of the test result lead to a change in patient management?*

12. What would be the impact of using this test on your patients/population?

jfponline

Does Amoxicillin Improve Outcomes in Patients with Purulent Rhinorrhea?

A Pragmatic Randomized Double-Blind Controlled Trial in Family Practice

AN I. DE SUTTER, MD; MARC J. DE MEYERE, MD, PHD; THIERRY C. CHRISTIAENS, MD; MIEKE L. VAN DRIEL, MD, MSC; WIM PEERSMAN; AND JAN M. DE MAESENEER, MD, PHD
Ghent, Belgium

■ OBJECTIVE To compare the efficacy of amoxicillin vs placebo in patients with an acute upper respiratory tract infection and purulent rhinorrhea.
■ STUDY DESIGN Double-blind randomized placebo-controlled trial.
■ POPULATION The 416 patients included from 69 family practices were 12 years or older, presenting with acute upper respiratory complaints, and having a history of purulent rhinorrhea and no signs of complications of sinusitis.
■ OUTCOMES MEASURED Therapy success (disappearance of symptoms that most greatly affected the patient's health) at day 10 and duration of general illness, pain, and purulent rhinorrhea.
■ RESULTS Therapy was successful in 35% of patients with amoxicillin and in 29% of patients with placebo (relative risk [RR] 1.14, 95% confidence interval [CI], 0.92-1.42). There was no effect on duration of general illness or pain. Duration of purulent rhinorrhea was shortened by amoxicillin (9 days vs 14 for clearing of purulent rhinorrhea in 75% of patients; $P = .007$). Diarrhea was more frequent with amoxicillin (29% vs 19%, RR 1.28, 95% CI, 1.05-1.57). No complications were reported. One patient (0.5%) receiving amoxicillin and 7 (3.4%) receiving placebo discontinued trial therapy because of exacerbation of symptoms (RR 0.25, 95% CI 0.04-1.56, $P = .07$). All 8 patients recovered with antibiotic therapy.
■ CONCLUSIONS Amoxicillin has a beneficial effect on purulent rhinorrhea caused by an acute infection of the nose or sinuses but not on general recovery. The practical implication is that all such patients, whatever the suspected diagnosis, can be safely treated with symptomatic therapy and instructed to return if symptoms worsen.
■ KEY WORDS Respiratory tract infections; sinusitis; antibiotics; therapeutics; family practice. (*J Fam Pract 2002; 51:317-323*)

KEY POINTS FOR CLINICIANS

- In patients with an acute upper respiratory tract infection that includes purulent rhinorrhea, treatment with amoxicillin has no effect on general recovery and increases the frequency of diarrhea.
- In most patients, symptoms of acute respiratory tract infection last for more than 10 days.
- Treatment without antibiotics and with appropriate follow-up is safe.
- Patients with purulent rhinorrhea caused by an acute infection of the nose or sinuses can initially be treated with symptomatic therapy, whatever the suspected diagnosis, and instructed to return if symptoms worsen.

Infections of the nasal passages are very common[1] and among the most frequent reasons for the prescription of antibiotics.[2,3] Such infections comprise diagnoses that include upper respiratory tract infection (URTI), rhinitis, rhinopharyngitis, and rhinosinusitis, which are very difficult to distinguish because of the lack of specific clinical features or simple office-based diagnostic tests.[4-7] These diagnostic difficulties probably explain why it remains unclear whether and when antibiotics should be used for such patients in clinical practice.

Although evidence shows that a small minority of patients benefit from antibiotic therapy, these patients are extremely difficult to recognize or identify. Three meta-analyses[8-10] on the effect of antibiotics in rhinosinusitis and 5 of 6 recent trials investigating the effect of antibiotics in rhinosinusitis,[11-13]

From the Department of General Practice and Primary Health Care (A.I.DeS., M.J.DeM., T.C.C., M.L.v.D., J.M.DeM.) and the Department of Population Studies and Social Sciences Research Methods (W.P.P.), University of Ghent, Belgium. This study was presented at the 28th annual meeting of the North American Primary Care Research Group, Amelia Island, Fla., November 2000, and at the European meeting of the World Organization of National Colleges, Academies, and Academic Associations of General Practitioners/Family Physicians (WONCA), Tampere, Finland, June 2001. Competing interest: This trial was financed by a grant by Eurogenerics NV, Brussels. Reprint requests should be addressed to An De Sutter, MD, Department of General Practice UG.UZG–1K3, De Pintelaan, 185, B 9000 Ghent, Belgium. E-mail: an.desutter@rug.ac.be.

rhinitis,[14] and bacterial rhinopharyngitis[15] almost exclusively studied patients with a diagnosis established by laboratory or imaging investigation. As a result, implementing the findings is difficult in daily practice, where radiologic or laboratory tests are not obtained for most patients with respiratory infections. Only 1 of the 6 trials[16] included patients with a set of clinical symptoms indicating rhinosinusitis. Because inclusion criteria were rather stringent, however, findings are applicable only to a small group of patients.

The purpose of this trial was to investigate the benefits of antibiotic therapy in a larger group of patients with nose or sinus infections, thereby making the results more widely applicable. Accordingly, we conducted a randomized, double-blind, placebo-controlled trial comparing the effect of amoxicillin with that of placebo in family practice patients with an acute upper respiratory tract infection and presenting with purulent rhinorrhea. Purulent rhinorrhea was chosen as the minimal criterion because it is the symptom most consistently associated with rhinosinusitis in diagnostic studies[5,17-21] and because its presence often leads family physicians (FPs) to prescribe antibiotics.[23-26] The trial was designed as a pragmatic effectiveness trial. Patient inclusion and evaluation were defined on a purely clinical basis to maximize relevance for routine daily practice.

METHODS

Study Population

Between October 1998 and December 1999, 69 FPs in Flanders, Belgium, agreed to enroll patients meeting the following inclusion criteria: age 12 years or older, presenting with a respiratory tract infection, and having purulent rhinorrhea. Exclusion criteria were allergy to penicillin or ampicillin; having received antibiotic therapy within the previous week; complaints lasting for more than 30 days; abnormality on clinical chest examination; complications of sinusitis (facial edema or cellulitis; orbital, visual, meningeal, or cerebral signs)[27]; pregnancy or lactation; comorbidity that might impair immune competence; and inability to follow the protocol because of language or mental problems. The Ethics Committee of the Ghent University Hospital (GUH) approved the study. All patients (or their guardians, for those younger than 16 years of age) gave written informed consent.

Treatment Assignment and Masking

In this double-blind trial, patients were assigned via a computer-generated random number list to receive 500 mg amoxicillin 3 times a day or placebo for 10 days. The trial medication was supplied in numbered uniform cardboard boxes, each containing 30 capsules of the same size, color, and shape for active and placebo treatment. The randomization list, kept at the pharmacy of GUH, was accessible to the participating FPs only in case of a serious adverse event.

To assess the effectiveness of masking, patients and their FPs guessed the treatment group at 10-day follow-up. Data were encoded and entered without knowledge of treatment allocation. Compliance was assessed by counting leftover medication. All patients were allowed to use xylometazoline 1% nose drops and paracetamol or ibuprofen to alleviate symptoms; these data were registered.

Assessment of Potential Recruitment Bias Caused by Exclusion

First, we compared the characteristics of patients enrolled by high-recruiting FPs (at least 14 patients recruited) with those of patients from low recruiters (at most 5 patients recruited). Second, we asked all participating FPs to complete a short questionnaire over a 6-week period on all patients eligible for the trial but not included in it (sex, age, body temperature, severity of nasal discharge and pain, reason for non-recruitment). Third, to estimate the proportion of sinusitis cases among included patients, all patients were invited for an optional radiologic examination of the maxillary sinuses (single Waters view).[28] Radiographs were taken in the nearest radiology unit, collected centrally, and evaluated by a radiologist of the GUH who specialized in the ear, nose, and throat.

Baseline Measurements

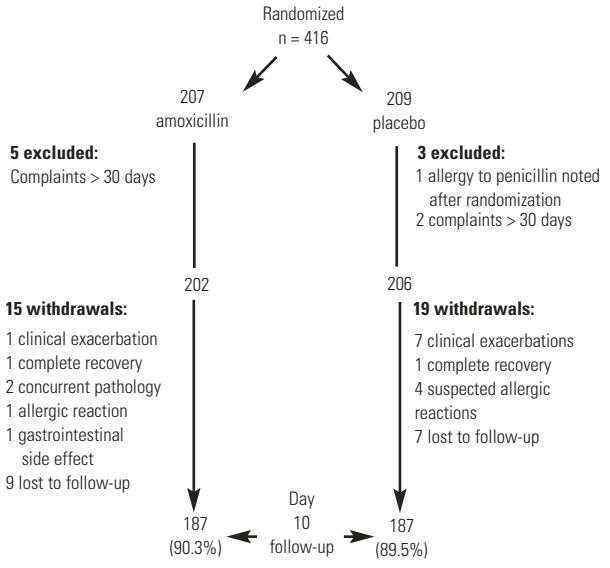
Randomized patients completed an extensive questionnaire and were physically examined by their FP. To evaluate the symptoms, we used the 20 items of the sinonasal outcome test (SNOT-20)[29,30] supplemented by 3 questions about pain. Symptoms were scored on a 6-category (0-5) Likert scale. Patients were also asked to indicate which of their symptoms (no more than 5) were most troublesome.

Follow-Up

During 10 days of treatment, all patients recorded their daily drug intake (trial medication and symptomatic medication); their general feeling of illness; the presence of nasal discharge, pain, and cough; body temperature; the occurrence of presumed adverse drug effects; and absence from work or school. On day 10 they underwent a second physical examination and completed the symptom questionnaire again. In case of insufficient recovery, the FP was then at liberty to prescribe an open antibiotic course (we recommended amoxicillin clavulanate) without revealing the previous treatment phase. Patients who

FIGURE 1

PATIENTS' PROGRESS THROUGH THE TRIAL



had recovered on day 10 did not have to return on day 15. Any patient with poor recovery on day 10 was asked, regardless of open antibiotic treatment, to continue writing in the diary and to come back on day 15 if complaints were still present.

The 2 primary endpoints were the therapy success rate on day 10 and the duration of general illness, pain, and purulent rhinorrhea as recorded in the diary. Treatment was considered successful when all symptoms that the patient had included in the list of "most important item affecting my health" scored 0 (absent) or 1 (very mildly present) after 10 days of treatment. Secondary endpoints were the mean change in severity score between day 1 and 10 on the various symptoms, incidence of unfavorable evolution, incidence of side effects, intake of analgesics, and duration of sick leave. The number of patients needed to demonstrate a difference in the therapy success rate of 15% at day 10 (α = 0.05, β = 0.20) was 168 per treatment group.[31] This determination assumed a success rate of 50% in the placebo group.[11,12]

Statistics

Data were analyzed with SPSS-7. Differences in proportions are presented as relative risks with 95% confidence intervals and tested by chi-square test. The duration of symptoms is presented by Kaplan–Meier survival plots. Differences in duration are tested by the log rank test. Other continuous variables are tested by Student's t test or the nonparametric Mann–Whitney U test.

RESULTS

Participant Flow and Follow-Up

Of 416 patients enrolled in the study, 8 were excluded after randomization. Of the 408 patients remaining, 202 received amoxicillin and 206 placebo; 34 patients (8%) withdrew from the trial. Their personal characteristics and clinical conditions at inclusion were not different from those of patients with follow-up. Figure 1 lists reasons for exclusion or withdrawal. The treatment code was broken once for a suspected allergic reaction and once because of an exacerbation of symptoms. In accordance with the intention-to-treat principle, all enrolled patients were included in the analyses in the groups to which they were originally randomized. Patients who had withdrawn because of side effects were also included in the analysis of side effects.

Complete or partial follow-up data were obtained for 374 patients (90%) after 10 days (mean 10.3 days, standard deviation 1.44): 334 patients completed the questionnaire, 348 returned the diary, and 338 underwent a physical examination. In 265 (71%) patients, data (questionnaire, diary, and physical examination) were complete; in 109 (29%), data at day 10 were partly missing. The two treatment groups were very similar in terms of sex, age, duration of preinclusion complaints, and frequency of various physical signs and symptoms (Table 1).*

Primary Outcomes

Of the 374 patients with follow-up data on day 10, 334 completed the symptom questionnaire twice. Treatment was successful—defined as a score of 0 (absent) or 1 (very mildly present) for all symptoms that had been included as "the most important item affecting my health"—in 35% of patients in the amoxicillin group (59/170) and 29% in the placebo group (47/164) (Table 2). Relative risk of success was 1.14 (95% CI, 0.92-1.42, P = .24): more patients were cured in the amoxicillin group, but this difference was not statistically significant.

In 82 (19.7%) of the 416 randomized patients (37 amoxicillin, 45 placebo), data on this main outcome are missing. In 40 of these 82 patients, follow-up data are available from the diary (n = 38) or physical examination (n = 2). According to these data, in 13/17 of the amoxicillin group and 11/23 of the

* For an expanded version of this table, see Table W1 at http://www.jfponline.com.

TABLE 1

BASELINE CHARACTERISTICS

General (placebo = 205, amoxicillin = 204)	**Placebo**	**Amoxicillin**
Mean age (SD)	39 (15)	37 (14)
Mean days of complaint before contact (SD	7.2 (5.5)	7.6 (5.4)
Women (%)	54	55
Mean Score on SNOT-20 (placebo = 196, amoxicillin = 192)	40.8 (SD 15.9)	38.4 (SD 16.1)
History (placebo = 196, amoxicillin = 192)		
Generally ill to very ill (%)	46	53
Unilateral facial pain (%)	56	53
Pain on bending forward (%)	70	66
Pain in upper teeth or when chewing (%)	44	41
Examination (placebo = 209, amoxicillin = 207)		
Sinus tenderness (%)	61	67
Pain on bending forward (%)	60	60
Postnasal discharge on throat inspection (%)	55	50
Purulent rhinorrhea on rhinoscopy (%)	47	40
Body temperature > 37°C (%)	38	41

SD denotes standard deviation; SNOT, Sino-Nasal Outcome Test.

placebo group the outcome was favorable: in the diary, the patient reports feeling "well" again at day 10 or sooner, or on physical examination, all signs of respiratory infection have cleared). Eight patients withdrew for clinical exacerbation and 2 patients after full recovery. Adding the 50 patients with a known course of illness to those in the treatment and result groups does not alter the overall result (RR 1.20, 95% CI, 0.98-1.47, P = .08). Furthermore, when considering the 24 nonexcluded patients (13 amoxicillin, 11 placebo) with total lack of follow-up in their allocated treatment group, first as treatment failures (RR 1.18, 95% CI, 0.97-1.44, P = .11) and then as successes (1.20, 95% CI, 0.99-1.46, P = .07), the result also remains the same. Regarding the success rate from the complete diary data (n = 348) and the results of physical examinations (n = 338) (Table 3), we find no significant difference between treatment groups.

Duration of purulent rhinorrhea was significantly shorter in the amoxicillin group than in the placebo group (75% of patients were free of purulent rhinorrhea after 9 days versus after 14 days in the placebo group, log rank P = .007). There is no difference between treatment groups in the duration of general illness or pain (Figure 2).

Secondary Outcomes

The mean score reduction on the symptom "thick nasal discharge" between day 1 and day 10 is significantly larger in the amoxicillin group than in the placebo group (2.2 vs 1.5, Student's t test: P <.0001) (Table 3). There is no significant difference in change for any other symptom. Seven patients in the placebo group (3.4%) withdrew before day 10 because of exacerbation of symptoms versus 1 patient (0.5%) in the amoxicillin group (RR 0.25, 95% CI, 0.04-1.56, P = .07). All 8 patients recovered after starting open antibiotic therapy and had no complications or referrals.

The chance of receiving open antibiotic treatment at day 10 follow-up (n = 34: 19 placebo, 15 amoxicillin) or of having to return because of persistent complaints at day 15 (n = 73: 41 placebo, 32 amoxicillin) was not significantly different between the treatment groups (chi-squared test: P = .46 and P = .26, respectively). Diarrhea was more frequent in the amoxicillin group (29% vs. 19%, RR 1.28, CI 1.05-1.57, P = .02). There was no difference in incidence of skin rash, abdominal pain, or vomiting. Absence from work or school was comparable in both treatment groups (RR 0.95, 95% CI, 0.86-1.05, P = .34). Patients in the amoxicillin group took an analgesic an average of 5 times, mainly in the first days of treatment,

TABLE 2

MAIN OUTCOME: RATE OF TREATMENT SUCCESS AT 10-DAY FOLLOW-UP

Outcome Measure	**N***	**Number with Successful Therapy (%)**		**Relative Risk of Success (95% CI)**	***P***
		Amoxicillin	Placebo		
Survey†	334	59/170 (35)	47/164 (28)	1.14 (0.92-1.42)	.24
Diary ‡	348	92/174 (52)	97/174 (55)	0.94 (0.77-1.16)	.59
Physical signs §	338	97/170 (57)	86/168 (51)	1.13 (0.91-1.40)	.28
All ‖	384	73/189 (39)	59/195 (30)	1.2 (0.98-1.47)	.08
Sensitivity analysis¶					
Best case	408	86/202	70/206	1.2 (0.99-1.46)	.07
Worst case	408	73/202	59/206	1.18 (0.97-1.44)	.11

* Data on at least one of these outcome measures were obtained in 374 patients (90% of the total population).
† All symptoms indicated by the patients at inclusion as "most important item affecting my health" score 0 (absent) or 1 (very mildly present) on day 10.
‡ Patient states in diary that he or she feels generally "well" again on day 10 or sooner.
§ All physical signs have disappeared at day 10 (pain on bending, sinus tenderness, postnasal drip, purulent rhinorrhea on rhinoscopy, elevated body temperature).
‖ Incorporating all available information from the questionnaire, diary, physical examination, and dropouts.
Patients without data are considered, respectively, as treatment success (best case) or treatment failures (worst case).

TABLE 3

MEAN SYMPTOM CHANGE BETWEEN BASELINE AND 10-DAY FOLLOW-UP

Symptom	Mean Score Reduction: Amoxicillin n = 170	Mean Score Reduction: Placebo n = 164	*P* *
Unilateral facial pain	1	1.1	.56
Pain on bending forward	1.21	1.32	.55
Pain in upper teeth or when chewing	0.7	0.93	.17
Need to blow nose	1.73	1.70	.85
Sneezing	1.13	1.05	.63
Runny nose	1.47	1.55	.33
Cough	1.0	1.11	.46
Thick nasal discharge	2.2	1.5	< .0001
Postnasal discharge	1.29	1.09	.26
Ear fullness	1.13	1.31	.32
Dizziness	0.95	0.87	.63
Ear pain	0.64	0.77	.36
Facial pain or pressure	1.54	1.61	.69
Difficulty falling asleep	1.14	1.26	.54
Wake up at night	1.39	1.44	.79
Lack of a good night's sleep	1.24	1.44	.28
Wake up tired	1.34	1.65	.09
Fatigue	1.46	1.61	.38
Reduced productivity	1.45	1.63	.29
Reduced concentration	1.24	1.46	.19
Frustrated, restless, irritable	0.87	1.41	.91
Sad	0.38	0.52	.18
Embarrassed	0.36	0.76	.36

* Student's t test.

compared with 4 for the placebo group (Mann–Whitney U test, $P = .24$).

Other Results

The lack of correlation between the estimated and actual treatment demonstrates that masking was maintained. Compliance was good in both groups: 89% of patients in the amoxicillin group and 91% of those in the placebo group took at least 25 of 30 capsules.

Patients from low recruiters were not significantly different from patients enrolled by high recruiters. Included patients had slightly more complaints of pain (58% vs 50%, RR 1.20, CI 1.02-1.42, $P = .03$) than the 332 eligible but excluded patients registered during the 6-week period. The most frequent reasons for exclusion were the presence of an exclusion criterion (22%), the patient's refusal to participate (16%), the patient's request for antibiotic therapy (14%), and lack of time by the FP (10%). Of the 292 patients who agreed to undergo a radiologic examination, about two thirds had abnormalities of the maxillary sinuses.

DISCUSSION

This study produced 3 important findings. First, we found that patients consulting their FP for acute URTI with purulent rhinorrhea do not experience any important benefit from amoxicillin therapy. With treatment, the purulent rhinorrhea disappears more quickly, but this seems to be of little importance in relation to a general recovery. Moreover, amoxicillin therapy increases the risk of diarrhea. We further found that with or without amoxicillin, complaints last long: after 10 days, two thirds of patients still had complaints and about half of the patients still felt ill. The natural course to recovery takes a long time and is not influenced by taking amoxicillin. Finally, we observed that failure to prescribe antibiotics is safe. The placebo group had no complications. A small number of exacerbations occurred, but these responded swiftly to a course of amoxicillin–clavulanate.

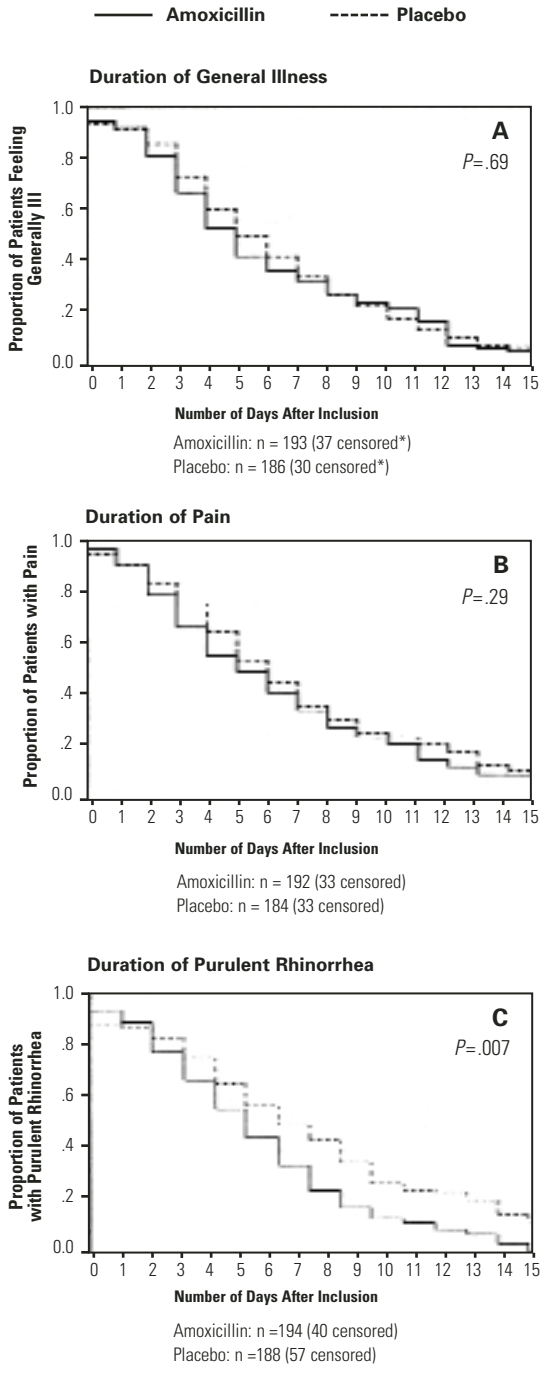
To our knowledge, this is the first time that the effect of an antibiotic in adult patients presenting with acute purulent rhinorrhea (but with an otherwise unspecified diagnosis) has been investigated in a randomized, placebo-controlled trial. This trial is in line with a number of other family practice–based pragmatic trials in which patients were included on the basis of respiratory symptoms instead of by diagnosis[16,32-37] and in which the emphasis was on practical relevance rather than on diagnostic accuracy.

Since 1995, 6 randomized clinical trials of high methodologic quality[11-16] have studied the efficacy of antibiotics in general practice patients suffering from various acute infections of the nasal passages and usually presenting with purulent rhinorrhea. In 3 of these trials, no beneficial effect of antibiotics was found. Study populations consisted, respectively, of patients with a set of clinical symptoms (including purulent rhinorrhea) indicating rhinosinusitis[16]; patients with clinical suspicion of rhinosinusitis plus sinus abnormalities on conventional radiology[11]; and patients with clinical suspicion of sinusitis but without the radiologic signs.[14] In the 3 other trials, treatment was (more or less) effective. Included were patients with clinical suspicion of sinusitis and abnormalities on CT scan,[12] patients with unilateral facial pain and elevated C-reactive protein levels or erythrocyte sedimentation rate,[13] and patients with rhinopharyngitis and positive bacteriologic cultures of nasopharyngeal secretions.[15] These trials show that antibiotics are efficacious in some patients. In our trial, which probably included a mix of all these populations, we also found more patients in the amoxicillin group to be symptom free after 10 days. Despite a fairly large sample size, however, this difference was too small (less than 15%) to be statistically significant.

In this trial, as in daily practice, we did not know the precise diagnosis of included patients. Moreover, despite our frequent requests, participating FPs

FIGURE 2

DURATION OF ILLNESS, PAIN, AND PURULENT RHINORRHEA BETWEEN TREATMENT GROUPS



"Censored" denotes that patient is not described as "well," "pain free," or "free of purulent rhinorrhea" at end of follow-up.

included only a minority of eligible patients. Concern might arise that only patients with mild disease were studied. We made 3 efforts to verify that the population was truly representative. First, we determined that the personal characteristics and severity of symptoms of patients of low-recruiting FPs (who tend to include patients with worse symptoms[38]) were no different from those of patients included by high recruiters. Second, an analysis of questionnaires from all eligible but excluded patients over a 6-week period showed that included and excluded patients were very much alike. The analysis also showed that in only 3% of patients did the FP consider the subject too ill to be included. Third, the results obtained on plain radiography of the maxillary sinuses were in line with the imaging results of other family practice populations with clinical suspicion of rhinosinusitis.[11,19-21]

With regard to the methodology, we wish to clarify certain choices. Amoxicillin was selected because it is recommended as the first-line drug for rhinosinusitis in several practice guidelines[39-41] and the sensitivity of respiratory pathogens to it was sufficient in our geographic area at the start of the trial.[42]* To evaluate symptoms, we chose the 20 items of the SNOT-20 questionnaire (Table 1), an abbreviated version of the RSOM-31,[29] a disease-specific quality-of-life test for sinusitis. These 20 items include not only all classic rhinosinusitis symptoms but also a number of more subjective symptoms, such as sleep disturbances and reduced productivity, which may also severely inconvenience patients. Any beneficial effect of amoxicillin on these symptoms would be just as important as an effect on the classic sinusitis symptoms.

Outcome measures were mainly self-assessed by patients, since in this kind of pathology, for which subjective inconvenience is often greater than objective signs might indicate, the patient is in our view the best and only judge of symptom improvement. The main outcome measure, disappearance of perceived worst symptoms, was designed to take into account the heterogeneity of clinical presentations.

CONCLUSIONS

Patients with an acute upper respiratory tract infection with purulent rhinorrhea (and without signs of complications of sinusitis) represent a large, clearly defined, clinically recognizable group. Our results show that amoxicillin pro-

* *Streptococcus pneumoniae* 97% sensitive and *Haemophilus influenzae* 87% sensitive: data from Ghent University Hospital, Laboratory of Bacteriology, De Pintelaan 185, B-9000 Ghent, Belgium. Director: Prof. G. Verschraege. Personal communication.

vides no clinically important benefits for this population. The implication for practice is that whatever diagnosis is suspected, all these patients can safely be treated with symptomatic therapy only. Patients should, however, be informed that whichever treatment is chosen, symptoms can last for a long time. In the rare event that symptoms worsen, they should consult their FP for antibiotic therapy. If patients are clearly distressed by the purulent rhinorrhea itself, this trial suggests reasons for considering the use of amoxicillin, but potential patient benefits still probably do not outweigh the disadvantages. JFP

· ACKNOWLEDGMENTS ·

The authors wish to thank all participating family physicians and patients and Erna Eeckhout, Adrienne Dubron, Anselme Derese, MD, PhD, and John Marshall for their invaluable help.

REFERENCES

1. Okkes IM, Oskam SK, Lamberts H. Van klacht naar diagnose. Episodegegevens uit de huisartspraktijk. Coutinho, Bussum, Nl; 1998.
2. De Maeseneer J. Het voorschrijven van antibiotica bij luchtwegproblemen. Een explorerend onderzoek. Huisarts Wet 1990; 33:223-6.
3. De Melker RA, Kuyvenhove MM. Management of upper respiratory tract infection in Dutch general practice. Br J Gen Pract 1991; 41:504-7.
4. van Buchem L, Peeters M, Beaumont J, Knottnerus JA. Acute maxillary sinusitis in general practice: the relation between clinical picture and objective findings. Eur J Gen Pract 1995; 1:155-60.
5. Hansen JC, Schmidt H, Rosborg J, Lund E. Predicting acute maxillary sinusitis in a general practice population. BMJ 1995; 311:233-6.
6. Stalman W, Van Essen GA, Gubbels JW, De Melker RA. Difficulties in diagnosing acute sinusitis in a Dutch group practice. Relative value of history, radiography and ultrasound. Eur J Gen Pract 1997; 3:12-5.
7. Hueston WJ, Mainous AG, Dacus EN, Hopper JE. Does acute bronchitis really exist? A reconceptualization of acute viral respiratory infections. J Fam Pract 2000; 49:401-6.
8. Ferranti de SD, Ioannidis JPA, Lau J, Anninger WV, Barza M. Are amoxicillin and folate inhibitors as effective as other antibiotics for acute sinusitis? A meta-analysis. BMJ 1998; 317:632-7.
9. Williams JW Jr, Aguilar C, Makela M, et al. Antibiotic therapy for acute sinusitis: a systematic literature review. In: Douglas R, Bridges-Webb C, Glasziou P, Lozano J, Steinhoff M, Wang E, eds. Acute respiratory infections module of the Cochrane Database of Systematic Reviews. The Cochrane Library. Oxford, England: Updated Software; 1997.
10. Zucker DR, Balk E, Engels E, Barza M, Lau J. Agency for Health Care Policy and Research Publication No. 99-E016. Evidence report/technology assessment number 9. Diagnosis and treatment of acute bacterial rhinosinusitis. Available at www.ahrq.gov/clinic/sinussum.htm.
11. van Buchem FL, Knottnerus JA, Scrhrijnemaekers VJJ, Peeters MF. Primary-care-based randomised placebo-controlled trial of antibiotic treatment in acute maxillary sinusitis. Lancet 1997; 349:683-7.
12. Lindbaek M, Hjortdahl P, Johnsen UL-H. Randomised, double blind, placebo controlled trial of penicillin V and amoxycillin in treatment of acute sinus infections in adults. BMJ 1996; 313:325-9.
13. Hansen JG, Schmidt H, Grinsted P. Randomised, double blind, placebo controlled trial of penicillin V in the treatment of acute maxillary sinusitis in adults in general practice. Scand J Prim Health Care 2000; 18:44-7.
14. Haye R, Lingaas E, Hoivik HO, Odegard T. Azithromycin versus placebo in acute infectious rhinitis with clinical symptoms but without radiological signs of maxillary sinusitis. Eur J Clin Microbiol Infect Dis 1998; 17:309-12.
15. Kaiser L, Lew D, Hirschel B, et al. Effects of antibiotic treatment in the subset of common-cold patients who have bacteria in nasopharyngeal secretions. Lancet 1996; 347:1507-10.
16. Stalman W, van Essen GA, van der Graaf Y, de Melker RA. The end of antibiotic treatment in adults with acute sinusitis-like complaints in general practice? A placebo-controlled double-blind randomized doxycycline trial. Br J Gen Pract 1997; 47:794-9.
17. Axelsson A, Runze U. Symptoms and signs af acute maxillary sinusitis. ORL J Otorhinolaryngol Relat Spec 1976; 38:298-308.
18. Berg O, Carenfelt C. Analysis of symptoms and clinical signs in the maxillary sinus empyema. Acta Otolaryngol (Stockholm) 1988; 105:343-9.
19. Lindbaek M, Hjortdahl HR, Johnsen UL-H. Use of symptoms, signs and bloodtests to diagnose acute sinus infections in primary care: comparison with computed tomography. Fam Med 1996; 28:181-6.
20. Van Duyn NP, Brouwer HJ, Lamberts H. Use of symptoms and signs to diagnose maxillary sinusitis in general practice: comparison with ultrasonography. BMJ 1992; 305:684-7.
21. Williams JW, Simel DL, Leroy R, Samsa GP. Clinical evaluation for sinusitis. Making the diagnosis by history and physical examination. Ann Int Med 1992; 117:705-10.
22. Axelsson A, Runze U. Comparison of subjective and radiological findings during the course of acute maxillary sinusitis. Ann Otol Rhinol Laryngol 1983; 92:75-7.
23. Gonzales R, Barrett PH, Steiner JF. The relation between purulent manifestations and antibiotic treatment of upper respiratory tract infections. J Gen Intern Med 1999; 14:151-6.
24. Little DR, Mann BL, Sherk DW. Factors influencing the clinical diagnosis of sinusitis. J Fam Pract 1998; 46:147-52.
25. Mainous AG, Hueston WJ, Eberlein C. Colour of respiratory discharge and antibiotic use. Lancet 1997; 350:1077.
26. De Sutter AI, De Meyere MJ, De Maeseneer JM, Peersman WP. Antibiotic prescribing in acute infections of the nose or sinuses: a matter of personal habit? Fam Pract 2001; 18:209-13.
27. Gray WC, Blanchard CL. Sinusitis and its complications. Am Fam Physician 1987; 35:232-43.
28. Williams JW, Roberts L, Distell B, Simel DL. Diagnosing sinusitis by X-ray: is a single Waters view adequate? J Gen Intern Med 1992; 7:481-5.
29. Piccirillo JF, Edwards D, Haiduk A, et al. Psychometric and clinimetric validity of the 31-item rhinosinusitis outcome measure (RSOM-31). Am J Rhinol 1995; 9:297-306.
30. Bhattacharyya T, Piccirillo J, Wippold FJ II. Relationships between patient-based descriptions of sinusitis and paranasal sinus computed tomographic findings. Arch Otolaryngol Head Neck Surg 1997; 123:1189-92.
31. Altman DG. Practical statistics for medical research. London, England: Chapman & Hall; 1991: 455-60.
32. De Meyere M. Acute keelpijn in de eerste lijn. [dissertatie]. 1990 Rijksuniversiteit Gent, Faculteit Geneeskunde.
33. Verheij TJM, Hermans J, Mulder JD. Effects of doxycycline in patients with acute cough and purulent sputum: a double blind placebo controlled trial. Br J Gen Pract 1994; 44:400-4.
34. Burke P, Bain J, Robinson D, Dunleavy J. Acute red ear in children: controlled trial of non-antibiotic treatment in general practice. BMJ 1991; 303:558-62.
35. Damoiseaux RAMJ, Balen van FAM, Hoes AW, Melker de RA. Primary care based randomised, double blind trial of amoxicillin versus placebo for acute otitis media in children aged under 2 years. BMJ 2000; 320:350-4.
36. Little P, Could C, Williamson I, Moore M, Warner G, Dunleavy J. Pragmatic randomised controlled trial of two prescribing strategies for childhood acute otitis media. BMJ 2001; 322:336-42.
37. Little P, Williamson I, Warner G, Gould C, Gantley M, Kinmonth AL. Open randomised trial of prescribing strategies in managing sore throat. BMJ 1997; 314:722-7.
38. Wilson S, Delaney BC, Roalfe A, et al. Randomised controlled trials in primary care: case study. BMJ 2000; 321:24-7.
39. De Bock GH, Van Duijn NP, Dagnelie CF, et al. NHG-Standaard Sinusitis. Huisarts Wet 1993;36:255-7.
40. Low DE, Desrosiers M, McSherry J, et al. A practical guide for the diagnosis and treatment of acute sinusitis. CMAJ 1997; 156(suppl 6):S1-14.
41. Snow V, Mottur-Pilson C, Hickner JM. Principle of appropriate antibiotic use for acute sinusitis in adults. Ann Intern Med 2001; 134:495-7.
42. Pierard D, De Meyer A, Vanzeebroeck A, Lauwers S. In vitro evaluatie van de gevoeligheid van 205 recente klinische isolaten van streptococcus pneumoniae voor minocycline en andere antibiotica. Tijdschr Geneesk 1996; 52:281-5.

Near patient testing for influenza in children in primary care: comparison with laboratory test

Anthony Harnden, Angela Brueggemann, Sasha Shepperd, Judy White, Andrew C Hayward, Maria Zambon, Derrick Crook, David Mant

Department of Primary Health Care, Institute of Health Sciences, University of Oxford, Oxford OX3 7LF
Anthony Harnden *university lecturer*
Sasha Shepperd *university research lecturer*
Judy White *research nurse*
David Mant *professor and head of department*
Academic Department of Microbiology and Infectious Disease, University of Oxford, Oxford OX3 9DU
Angela Brueggemann *senior research fellow*
Derrick Crook *consultant microbiologist*
University College London Centre for Infectious Disease Epidemiology, London NW3 2PF
Andrew C Hayward *senior lecturer*
Virus Reference Division, Central Public Health Laboratory, Colindale, London NW9 5HT
Maria Zambon *consultant virologist*
Correspondence to: A Harnden anthony.harnden@dphpc.ox.ac.uk

BMJ 2003;326:480

Influenza is an important cause of acute respiratory illness in young children. Common complications include febrile convulsions, otitis media, bronchiolitis, and croup. In epidemic years attack rates among preschool children often exceed 40%. During these years children with influenza may account for up to 30% of the increase in antibiotic prescribing.[1] Symptoms and signs of influenza in children are not specific and can mimic a range of other common respiratory viral pathogens. One quick way of reaching a precise diagnosis in primary care is to use a near patient test. Near patient testing for many conditions has expanded widely in primary care, though many tests have not been rigorously evaluated.[2]

Previous studies in children have compared near patient influenza tests with viral culture analysis using throat or nasal swabs.[3] However, a nasopharyngeal aspirate is the best specimen for detecting influenza viruses, and polymerase chain reaction (PCR) is more sensitive than tissue culture when serology is the reference standard.[4 5] We compared a near patient influenza test in children in primary care with laboratory based reverse transcription PCR (RT-PCR) testing of nasopharyngeal aspirates.

Participants, methods, and results

From January to March 2001 and October to March 2002 we asked general practitioners in Oxfordshire to identify children with cough and fever who they thought had more than a simple cold. Using a nasal swab we performed a near patient test for influenza (QuickVue; Quidel, San Diego, CA). A research nurse did the test, which took 12 minutes.

We collected a nasopharyngeal aspirate from the other nostril and transported the sample to the laboratory within four hours. The laboratory staff were blind to the result of the near patient test. After adding phosphate buffered saline to the aspirate we added the emulsified sample to viral lysis buffer before freezing it at −80°C. We used RT-PCR to convert the extracted nucleic acids from RNA to complementary DNA. We performed a multiplex, nested PCR assay, using primer sets specific to influenza A and B, on all the samples. To validate our results we included quantified tissue culture specimens of influenza A and B as positive controls and water as negative control with every batch of samples tested.

A nasal swab and a nasopharyngeal aspirate were taken from 157 children. The children's median age was 3 years (range 6 months to 12 years), and 100 were boys. We detected influenza by RT-PCR in 61 children (39%). The near patient test was positive in 27 of these 61 children, giving a sensitivity of 44% (95% confidence interval 32% to 58%) and a specificity of 97% (91% to 99%) (table). The likelihood ratio for a positive test result was 14.2 (4.5 to 44.7) and for a negative result 0.58 (0.46 to 0.72).

Comparison of near patient testing with reverse transcription polymerase chain reaction (RT-PCR) testing for influenza in children

	RT-PCR test		
	Positive	Negative	Total
Near patient test:			
Positive	27	3	30
Negative	34	93	127
Total	61	96	157

Comment

The high specificity of this near patient test, combined with its ease of use, makes it suitable to "rule in" diagnosis of influenza in children in primary care, although its low sensitivity means it cannot "rule out" influenza. A sensitivity lower than has been described previously can be explained by our choice of RT-PCR as our reference standard, on a nasopharyngeal aspirate, rather than tissue culture testing on a nasal swab.[3] Future evaluations of near patient tests should use molecular reference standards rather than traditional culture based techniques. A secure diagnosis of influenza in children in primary care may be important in guiding the general practitioner's optimal management, improving the surveillance of influenza, and satisfying parents, rather than telling them, "It's just a virus."

We thank the 40 Oxfordshire general practitioners who took part and Tim Lancaster for comments on an earlier draft of the paper.

Contributors: AH, ACH, DC, MZ, and DM designed the study. AH and JW took part in the fieldwork. AB, DC, and MZ were responsible for the laboratory work. AH and SS did the analysis. AH drafted the manuscript, and all authors commented on the text. AH is guarantor for the study.

Funding: The study was funded by the Medical Research Council as part of a programme grant in childhood infection in primary care (G0000340). The guarantor accepts full responsibility for the conduct of the study, had access to the data, and controlled the decision to publish.

Competing interests: None declared.

Ethical approval: Oxford clinical research ethics committee (C00.180).

1 Neuzil KM, Mellen BG, Wright PF, Mitchel EF Jr, Griffin MR. The effect of influenza on hospitalizations, outpatient visits, and courses of antibiotics in children. *N Engl J Med* 2000;342:225-31.
2 Delaney BC, Hyde CJ, McManus RJ, Wilson S, Fitzmaurice DA, Jowett S, et al. Systematic review of near patient test evaluations in primary care. *BMJ* 1999;319:824-7.
3 Rodriguez WJ, Schwartz RH, Thorne MM. Evaluation of diagnostic tests for influenza in a pediatric practice. *Pediatr Infect Dis J* 2002;21:193-6.
4 Carman WF, Wallace LA, Walker J, McIntyre S, Noone A, Christie P, et al. Rapid virological surveillance of community influenza infection in general practice. *BMJ* 2000;321:736-7.
5 Zambon M. Laboratory diagnosis of influenza. In Nicholson KG, Webster RG, Hay AJ, eds. *Textbook of influenza*. Oxford: Blackwell Science, 1998: 291-313.

(Accepted 17 January 2003)