

# The Accuracy of Physical Examination to Detect Abdominal Aortic Aneurysm

Howard A. Fink, MD; Frank A. Lederle, MD; Craig S. Roth, MD; Carolyn A. Bowles, MD; David B. Nelson, PhD; Michele A. Haas, BA

**Background:** Abdominal palpation during physical examination is an important means of detecting abdominal aortic aneurysm (AAA), but limited information is available on its accuracy.

**Methods:** Two hundred subjects (aged 51-88 years), 99 with and 101 without AAA as determined by previous ultrasound, each underwent physical examination of the abdomen by 2 internists who were blinded to each other's findings and to the ultrasound diagnosis.

**Results:** The overall accuracy of abdominal palpation for detecting AAA was as follows: sensitivity, 68% (95% confidence interval [CI], 60%-76%); specificity, 75% (95% CI, 68%-82%); positive likelihood ratio, 2.7 (95% CI, 2.0-3.6); negative likelihood ratio 0.43 (95% CI, 0.33-0.56). Interobserver pair agreement for AAA vs no AAA between the first and second examinations was 77% ( $\kappa = 0.53$ ). Sensitivity increased with AAA diameter, from 61% for AAAs of 3.0 to 3.9 cm, to 69% for AAAs of 4.0 to 4.9 cm, 72% for AAAs of 4.0 cm or larger, and 82% for AAAs of 5.0 cm or larger. Sensitivity in subjects with

an abdominal girth less than 100 cm (40-in waistline) was 91% vs 53% for girth of 100 cm or greater ( $P < .001$ ). When girth was 100 cm or greater and the aorta was palpable, sensitivity was 82%. When girth was less than 100 cm and the AAA was 5.0 cm or larger, sensitivity was 100% (12 examinations). Factors independently associated with correct examination findings included AAA diameter (odds ratio [OR], 1.95 per centimeter increase; 95% CI, 1.06-3.58); abdominal girth (OR, 0.90 per centimeter increase; 95% CI, 0.87-0.94); and the examiner's assessment that the abdomen was not tight (OR, 2.68; 95% CI, 1.17-6.13).

**Conclusions:** Abdominal palpation has only moderate overall sensitivity for detecting AAA, but appears to be highly sensitive for diagnosis of AAAs large enough to warrant elective intervention in patients who do not have a large girth. Abdominal palpation has good sensitivity even in patients with a large girth if the aorta is palpable.

*Arch Intern Med.* 2000;160:833-836

**A**BDOMINAL AORTIC aneurysms (AAAs) cause more than 10 000 deaths each year in the United States,<sup>1</sup> and many of these deaths should be preventable by early detection and elective repair of the aneurysm.<sup>2</sup> Abdominal palpation during physical examination is an important means of AAA detection, accounting for about one third of new diagnoses.<sup>3</sup> Palpation for AAA is one of the few physical examination maneuvers recommended for the periodic health examination of older men,<sup>4</sup> and may be more cost-effective for screening than ultrasound.<sup>5</sup> The literature on the accuracy of abdominal palpation for AAA has recently been reviewed,<sup>6</sup> but important questions remain. Of the 15 studies of abdominal palpation for AAA in which the diagnosis was not known in advance by the examiner, only 1 included more than

10 patients with AAAs 4.0 cm in diameter or larger.<sup>6</sup> Furthermore, the effect of patient factors, such as abdominal obesity or tightness, on the accuracy of AAA palpation has received little attention. We are also aware of no studies reporting interobserver agreement on palpation findings; nor are we aware of any data addressing whether negative findings are more reliable for excluding AAA when the aorta is palpable than when it is not. The purpose of the present study is to provide more detailed information on the accuracy of abdominal palpation for AAA.

## RESULTS

There were 200 subjects included in the study, for a total of 400 abdominal examinations. The mean age was 73 years (range, 51-88 years), and 196 were men. Ninety-nine subjects had AAA as determined by

From the Department of Medicine and Center for Chronic Disease Outcomes Research, Minneapolis Veterans Affairs Medical Center, University of Minnesota, Minneapolis.

## SUBJECTS AND METHODS

The study was carried out at the Veterans Affairs Medical Center, Minneapolis, Minn, from August 1997 to June 1998. Subjects older than 50 years with known AAAs, defined as 3.0 cm in diameter or greater, or known absence of AAAs (as determined by ultrasound, within the past 6 months for AAAs and within 18 months for absence of AAAs) were invited to attend the study clinic. The medical center's institutional review board approved the study and informed consent was obtained from all participants. Three board-certified internists served as examiners for the study. Each subject was examined by 2 examiners on the same day, who were blinded to each other's findings and to the ultrasound results. Subjects with and without AAAs were mixed in the schedule, and subjects were asked not to reveal their ultrasound findings to the examiners. After each examination, the examiner was asked whether the subject had revealed the ultrasound findings. The examiners were given a brief (10-minute) training session in AAA palpation at the beginning of the study. The ultrasound findings were provided as feedback (on receipt of the completed data form) to each examiner after his or her first 25 examinations to ensure moderate expertise.

The examination was conducted with the patient supine with knees raised and the abdomen relaxed. Abdominal girth was measured supine at the umbilicus once for each subject. Auscultation was not performed because bruits were not found to contribute to the diagnosis of AAA in our previous study.<sup>7</sup> A transmitted epigastric pulse was considered present if the examiner felt pulsation when placing the palm of the hand over the epigastrium. The examiner made subjective judgments as to whether the abdomen was obese and/or tight (ie, did not yield to deep palpation), and attempted to identify the aortic pulsation by deep palpation of the upper abdomen. Both hands were then placed on the abdomen with palms down and an index finger on either side of the pulsating area to measure the aortic width. An AAA was considered to be present when the aorta was judged to be 3.0 cm or greater in maximum diameter after accounting for skinfold thickness. The examiner could rate the aorta "suggestive" or "definite" for AAA or as no AAA. Suggestive and definite were considered

together as positive in the analysis because either should prompt a follow-up imaging study. This decision and the decisions to use girth of 100 cm or greater (40-in waistline) and AAA diameter by centimeter as cutoff points were made a priori based on our previous study.<sup>7</sup>

Sensitivity was calculated as the proportion of subjects with disease with positive findings, and specificity as the proportion of subjects without disease with negative findings. In clinical practice, ultrasound serves as a safe, accurate, and inexpensive method of identifying false-positive findings from palpation for AAA. Therefore, the outcome of primary interest in this study was the sensitivity of palpation.

Likelihood ratios were also calculated. The positive likelihood ratio (LR+) is defined as sensitivity/(1 - specificity) and is the increase in the odds of having the disease when the finding is positive (LR+ values are  $\geq 1$ ). The negative likelihood ratio (LR-) is defined as (1 - sensitivity)/specificity and is the decrease in the odds of having the disease when the finding is negative (LR- values are 0-1).

Interobserver agreement was assessed using the  $\kappa$  statistic, for which 1.0 represents perfect agreement and 0 represents the agreement expected due to chance. On this scale, values above 0.75 are considered excellent agreement, values from 0.40 to 0.75 indicate fair to good agreement, and values below 0.40 indicate poor agreement.<sup>8</sup>

We considered an examination as the unit of analysis, recognizing that the 2 examinations per subject were not independent. To calculate confidence intervals in this setting, the covariance of the within-subject examinations was estimated and incorporated into the overall variance.<sup>9</sup> Repeated measures logistic regression models assuming that the 2 within-subject examinations had the same variance and some correlation (ie, an exchangeable correlation structure<sup>10</sup>) were used to examine how the odds of correct examination findings (the ratio of true positives to false negatives) changed with AAA diameter, abdominal girth, and abdominal tightness. Abdominal obesity was not included in this model because it was highly correlated with abdominal girth and less strongly associated with correct examination findings. *P* values were calculated from univariable repeated measures logistic models in which the examination finding was the dependent variable.

ultrasound. Of these AAAs, 41 measured 3.0 to 3.9 cm, 44 measured 4.0 to 4.9 cm, 12 were 5.0 to 5.9 cm, and 2 were 6.0 to 6.9 cm. Abdominal girth was 100 cm or greater in 59% of subjects. The abdomen was rated obese in 45% of examinations (interobserver pair agreement, 88%;  $\kappa = 0.76$ ), and tight in 28% of examinations (interobserver pair agreement, 70%;  $\kappa = 0.24$ ). Pair agreement between obesity and abdominal girth 100 cm or larger was 85% ( $\kappa = 0.71$ ).

As shown in the **Table**, the overall accuracy of abdominal palpation for detecting AAA was as follows: sensitivity, 68% (95% confidence interval [CI], 60%-76%); specificity, 75% (95% CI, 68%-82%); LR+, 2.7 (95% CI, 2.0-3.6); LR-, 0.43 (95% CI, 0.33-0.56). Interobserver pair agreement for AAA vs no AAA between the first and second examinations was 77% ( $\kappa = 0.53$ ). If definite, suggestive, and no AAAs are considered as 3 different cat-

egories, interobserver pair agreement was 65% ( $\kappa = 0.42$ ). There were no significant differences in sensitivity, specificity, or likelihood ratios between the first and second examinations, or among the 3 examiners (Table). The Table also shows that accuracy did not increase after the first 10 or 25 examinations by each examiner, suggesting the absence of an important training effect. The subject revealed the ultrasound findings to the examiner during 12 of the 400 examinations. Exclusion of these examinations (ie, limiting the analysis to those that remained blind) had no clinically or statistically significant effect on accuracy (Table).

The abdominal aorta was palpable in 67% of all examinations and in 77% of examinations on subjects with AAA. Interobserver agreement for palpable aorta was 85% ( $\kappa = 0.66$ ). Sensitivity was 88% when the aorta was palpable, and of course was 0% when it was not. Sensitivity

**The Accuracy of Abdominal Palpation to Detect Aortic Aneurysm in 99 Subjects With Disease and 101 Subjects Without Disease, Each Examined by 2 Internists\***

Examinations	No. of Examinations	Sensitivity	Specificity	Likelihood Ratios	
				Positive	Negative
All	400	68	75	2.7	0.43
First examination per subject	200	69	75	2.8	0.42
Second examination per subject	200	67	74	2.6	0.45
Examiner A	156	66	74	2.5	0.46
Examiner B	145	68	70	2.3	0.46
Examiner C	99	70	83	4.1	0.36
First 10 per examiner	30	76	67	2.3	0.36
First 25 per examiner	75	73	80	3.7	0.33
Examiner remained blind	388	67	76	2.7	0.44
Aorta was palpable	268	88	56	2.0	0.22
Positive = "definite" only	400	44	92	5.3	0.61
Girth					
<100 cm†	164	91	64	2.5	0.14
≥100 cm†	236	53	83	3.2	0.56
Abdomen "obese"	180	46	84	2.9	0.64
Abdomen "not obese"	220	89	66	2.6	0.17
Abdomen "tight"	111	52	89	4.7	0.54
Abdomen "not tight"	289	74	68	2.3	0.38
Abdomen ≥100 cm† and not tight	158	61	78	2.8	0.50
Abdomen ≥100 cm† and aorta palpable	125	82	59	2.0	0.30

\* See text for details and P values.

† A 100-cm girth is approximately a 40-in waistline.

decreased to 44% and specificity increased to 92% if only examination findings rated definite for AAA were considered positive. The finding of a positive transmitted epigastric pulse added 1 true-positive and 12 false-positive results to the group of positive examination results defined by suggestive or definite pulsatile mass. This is a lower ratio of true positives to false positives than would be expected by selecting patients through chance alone, indicating that this finding did not add useful information.

Sensitivity increased with AAA diameter, from 61% for AAAs of 3.0 to 3.9 cm, to 69% for AAAs of 4.0 to 4.9 cm, 72% for AAAs of 4.0 cm or larger, and 82% for AAAs of 5.0 cm or larger. When AAA diameter was assessed as a continuous variable, the increase in sensitivity with increasing diameter approached statistical significance ( $P < .07$ ).

Sensitivity of palpation in subjects with an abdominal girth of less than 100 cm was 91% compared with 53% when abdominal girth was 100 cm or greater (Table,  $P < .001$ ), and sensitivity also decreased with increasing abdominal girth when girth was assessed as a continuous variable ( $P < .001$ ). When abdominal girth was less than 100 cm and AAA diameter was 5.0 cm or larger, sensitivity was 100% (12 examinations). Similarly, the sensitivity of palpation when the abdomen was rated not obese was 89%, compared with 46% when rated obese ( $P < .001$ ), and was 74% when rated not tight, compared with 52% when rated tight ( $P = .005$ ). One goal of the study was to determine whether either a not tight abdomen or the finding of a palpable aorta would define a subset of subjects with large girth for whom palpation had high sensitivity for AAA. As shown in the Table, a palpable aorta substantially improved sensitivity in sub-

jects with large girth (to 82%), whereas a not tight abdomen had only a small effect.

In the multiple logistic regression analysis, all 3 factors (AAA diameter and abdominal girth as continuous variables and abdominal tightness as a dichotomous variable) were independently associated with correct examination findings. A 1.0-cm increase in AAA diameter approximately doubled the odds of correct examination findings (OR, 1.95; 95% CI, 1.06-3.58). A 1.0-cm increase in abdominal girth decreased these odds by approximately 10% (OR, 0.90; 95% CI, 0.87-0.94), and a not tight rating was associated with a 2½-fold increase (OR, 2.68; 95% CI, 1.17-6.13). There were no statistically significant interactions between these independent variables.

#### COMMENT

In this study, abdominal palpation had moderate sensitivity for detecting AAA. Sensitivity increased with AAA diameter, consistent with previous screening studies<sup>6</sup> and with the previously described relationship between sensitivity and disease severity.<sup>11</sup> The sensitivity of abdominal palpation was higher in the present study than in most previous screening studies. For AAAs measuring 3.0 to 3.9 cm, 4.0 to 4.9 cm, and 5.0 cm and greater, sensitivities were 61%, 69%, and 82%, respectively, in the present study, compared with 29%, 50%, and 76% in pooled estimates of previous screening studies.<sup>6</sup>

One likely reason for the increased sensitivity in the present study was increased diagnostic vigilance owing to the high prevalence of AAA.<sup>12</sup> Although the examiners in our study were not informed of the AAA prevalence in advance, their experience with early examina-

tions and ultrasound feedback clearly suggested a high AAA prevalence, and this likely increased their expectation and vigilance for finding AAA. Our findings could therefore be considered more representative of a consultation (as when a senior physician seeks to confirm a finding by a junior physician) than of a screening examination. (By this analogy, we do not mean to suggest that formal consultation should be obtained to confirm physical examination findings, as ultrasound would usually be a better alternative.)

Another possible reason for the high sensitivity we observed was that our examiners were completely focused on abdominal palpation for AAA without distraction by other physician-patient concerns. However, this was probably also true in all of the previously reported screening studies. Finally, although our examiners had no prior special expertise in AAA palpation, it is possible that they became very highly skilled during the course of the study. However, the results of the first 10 and 25 examinations (after which feedback ended) show that sensitivity did not increase after the early examinations.

Our study was the first to involve sufficient numbers of AAA to examine the effect of patient factors such as abdominal obesity, girth, and tightness, and the effect of a palpable aorta. Abdominal girth of 100 cm or more was closely correlated with the examiners' rating of the abdomen as obese. Sensitivity was significantly higher (91%) when girth was less than 100 cm or the abdomen was rated not obese (85%), and was 100% when girth was less than 100 cm and the AAA measured 5.0 cm or greater. A recent clinical trial found that elective repair of AAAs smaller than 5.5 cm did not improve survival<sup>13</sup>; so, our data suggest that abdominal palpation has high sensitivity to detect AAA large enough to require elective intervention in patients with girth less than 100 cm.

Because our previous work indicated that abdominal palpation was insensitive when girth was 100 cm or greater,<sup>7</sup> we sought to determine whether subgroups of patients with large girth could be identified in whom abdominal palpation might be reliable. Tightness of the abdomen is sometimes mentioned as an important factor by examiners. While this finding was negatively associated with sensitivity, it did not define a subset of subjects with large girth for whom sensitivity was high. The finding of a palpable aorta was more successful, defining a subset of subjects with large girth for whom sensitivity was 82% (Table).

Our study was also the first to look at interobserver variability in abdominal palpation for AAA. The mean

pair agreement (77%) and  $\kappa$  (53%) for AAA vs no AAA are considered fair to good.<sup>8</sup> The interobserver agreement for abdominal tightness was poor, casting further doubt on the value of this finding. The transmitted epigastric pulse, a maneuver suggested by a senior clinician, was not useful.

Abdominal palpation to measure the width of the aortic pulsation thus remains the only physical finding of demonstrated value for AAA detection.<sup>6</sup> It appears to be highly sensitive for diagnosis of AAA large enough to warrant elective intervention in patients who do not have a large girth, and it has good sensitivity even in patients with a large girth if the aorta is palpable. Patients with definite or suggestive findings for AAA on physical examination should have the diagnosis confirmed by an appropriate imaging study such as ultrasound.

Accepted for publication July 31, 1999.

Reprints: Frank A. Lederle, MD, Department of Medicine (III-0), Minneapolis Veterans Affairs Medical Center, 1 Veterans Dr, Minneapolis, MN 55417.

## REFERENCES

1. Gillum RF. Epidemiology of aortic aneurysm in the United States. *J Clin Epidemiol.* 1995;48:1289-1298.
2. Scott RAP, Wilson NM, Ashton HA, Kay DN. Influence of screening on the incidence of ruptured abdominal aortic aneurysm: 5-year results of a randomized controlled study. *Br J Surg.* 1995;82:1066-1070.
3. Kiev J, Eckhardt A, Kerstein MD. Reliability and accuracy of physical examination in detection of abdominal aortic aneurysms. *Vasc Surg.* 1997;31:143-146.
4. Oboler SK, LaForce FM. The periodic physical examination in asymptomatic adults. *Ann Intern Med.* 1989;110:214-226.
5. Frame PS, Fryback DG, Patterson C. Screening for abdominal aortic aneurysm in men ages 60 to 80 years: a cost-effectiveness analysis. *Ann Intern Med.* 1993; 119:411-416.
6. Lederle FA, Simel DL. The rational clinical examination: does this patient have abdominal aortic aneurysm? *JAMA.* 1998;281:77-82.
7. Lederle FA, Walker JM, Reinke DB. Selective screening for abdominal aortic aneurysms with physical examination and ultrasound. *Arch Intern Med.* 1988;148: 1753-1756.
8. Fleiss JL. *Statistical Methods for Rates and Proportions.* 2nd ed. New York, NY: John Wiley & Sons Inc; 1981:212-236.
9. Berry DA, Lindgren BW. *Statistics: Theory and Methods.* 2nd ed. Belmont, Calif: Wadsworth Publishing Co; 1996:205-208, 339-347.
10. Liang KY, Zeger SL. Longitudinal data analysis using generalized linear models. *Biometrika.* 1986;73:13-22.
11. Lofgren RP. The dynamic nature of sensitivity and specificity. *J Gen Intern Med.* 1987;2:452-453.
12. Fowkes FGR. Diagnostic vigilance. *Lancet.* 1986;1:493-494.
13. The UK Small Aneurysm Trial Participants. Mortality results for randomised controlled trial of early elective surgery or ultrasonographic surveillance for small abdominal aortic aneurysms. *Lancet.* 1998;352:1649-1655.