SELECTED ARTICLES

Should emergency physicians use B-type natriuretic peptide testing in patients with unexplained dyspnea?

Clinical question

Does the measurement of B-type natriuretic peptide (BNP) in adult emergency department (ED) patients with dyspnea result in sufficient change in the clinical likelihood of congestive heart failure (CHF) to affect medical decision-making?

Articles chosen

Maisel AS, Krishnaswamy P, Nowak RM, McCord J, Hollander JE, Duc P, et al; Breathing Not Properly Multinational Study Investigators. Rapid measurement of B-type natriuretic peptide in the emergency diagnosis of heart failure. N Engl J Med 2002;347(3):161-7.

McCullough PA, Nowak RM, McCord J, Hollander JE, Herrmann HC, Steg PG, et al. B-type natriuretic peptide and clinical judgment in emergency diagnosis of heart failure: analysis from Breathing Not Properly (BNP) Multinational Study. Circulation 2002;106:416-22.

Objective

To determine whether measurement of BNP helps to establish or refute the diagnosis of CHF in adult patients presenting to the ED with dyspnea.

Background

BNP is a neurohormone released by ventricular myocytes in response to elevations of end-diastolic pressure and volume. Reports have suggested that plasma BNP levels correlate with CHF severity and prognosis.[1] A recent pilot study concluded that BNP measurements might help distinguish patients with CHF from those with noncardiac dyspnea when the diagnosis is clinically ambiguous.[2]

In the following appraisal we discuss 2 recent articles authored by the Breathing Not Properly Multinational Study Investigators, which present complementary results derived from the same dataset. By reviewing both articles, we can provide a more accurate and comprehensive assessment of the quality of the dataset and the ED utility of BNP assays.

Reviewed by: Corinne M. Hohl, MD;* Bernice Y. Mitelman, MD;* Peter Wyer, MD;† Eddy Lang, MD*‡
*McGill University, Montréal, Que.; †New York Presbyterian Hospital, New York; ‡Jewish General Hospital, Montréal
Date appraised: Feb. 28, 2003

Received: Mar. 6, 2003; accepted: Mar. 21, 2003

Can J Emerg Med 2003;5(3):162-5

Population studied

Between April 1999 and December 2000 all adults (>18 years of age) who presented to a participating ED with the chief complaint of shortness of breath were recruited into the study. Patients were excluded if they had acute myocardial infarction, renal failure, or if their dyspnea was clearly not caused by CHF (e.g., chest wall trauma or tamponade). Those with unstable angina were only included if shortness of breath was the most prominent presenting symptom.

Study design

This study was an international multicentre prospective diagnostic test evaluation conducted in 7 centres, including 5 in the United States, 1 in France and 1 in Norway. Eligible patients were enrolled while in the ED. Blood samples were obtained and evaluated using a bedside fluorescence immunoassay kit (Triage BNP Test Kit, Biosite Inc, San Diego). Research assistants collected key clinical data in addition to the results of blood tests, x-rays and ECGs. All patients were examined in the ED by an attending emergency physician or internist, both blinded to BNP results. At the time of disposition from the ED, treating physicians were asked to estimate the probability that the patient's dyspnea was due to CHF.

The final diagnosis (reference standard) was assigned 30 days after the ED visit by 2 cardiologists at each centre, who independently reviewed case report forms and patient charts but remained blinded to BNP results and treating physicians' diagnostic impressions. The cardiologists re-

viewed the hospital course of patients who were admitted and reviewed the results of tests obtained during and after the ED stay, including ECGs, chest x-ray interpretations, echocardiograms, radionuclide angiography, ventriculography and cardiac catheterizations. The cardiologists classified patients into 3 diagnostic categories: dyspnea due to acute CHF; noncardiac dyspnea in a patient with underlying CHF; and noncardiac dyspnea in a patient with no prior CHF. In all cases of noncardiac dyspnea, medical records were reviewed 30 days after study enrollment to ensure that the patient had not been admitted with CHF after the index visit. In cases where the cardiologists were unable to agree on CHF classification, the study end-points committee assigned prior final diagnosis.

Outcomes measured

The outcomes measured included sensitivity, specificity, positive and negative predictive values, and accuracy of different BNP cut-off values for diagnosing congestive heart failure.

Results

Of 1666 patients screened, 1586 were eligible, signed informed consent and were enrolled. The mean age of study subjects was 64 + 17 years, 56% were male, 33% had prior CHF, 27% had a prior myocardial infarction, 41% had chronic obstructive pulmonary disease and 25% had diabetes mellitus. Treating physicians estimated the pretest (i.e., pre-BNP) probability of CHF in 1538 patients. Of these, 46.9% fell into the 0%–20% probability group, 27.9% fell into the 20%–80% (clinically uncertain) group, and 25.4% fell into the 80%–100% probability group.

A final diagnosis of acute CHF was made in 744 (47%) patients, and in 97% of cases this diagnosis was confirmed by objective tests like chest x-ray (79%), echocardiography (77%), radionuclide ejection fraction (15%) and cardiac catheterization (19%). Table 1 correlates BNP levels with outcome diagnoses, and Table 2 summarizes the sensitivity, specificity, and negative and positive predictive values for BNP cut-off levels ranging from 50–150 pg/cc. It is important to note that the cut-off value proposed by the authors to indicate an abnormal BNP value (i.e., acute CHF) is greater than 100 pg/cc.

Table 1. Mean B-type natriuretic peptide (BNP) values for each group

Causes of dyspnea	No. (and %)	Mean BNP level, pg/cc (± SD)
Group 1 CHF	744 (46.9)	675 (± 450)
Group 2 Non-cardiac dyspnea + underlying CHF	72 (4.5)	346 (± 390)
Group 3 Non-cardiac dyspnea; no prior CHF	770 (48.5)	110 (± 225)

SD = standard deviation; CHF = congestive heart failure

Conclusions

The authors concluded that the rapid measurement of BNP, using a cut-off value of greater than 100 pg/cc, will improve clinicians' ability to differentiate CHF from non-cardiac dyspnea in the emergency department.

Additional analysis

A diagnostic test is most useful in patients with indeterminate disease status and least useful in those with very high or very low pretest likelihood of disease. With this in mind, we used subgroup data reported in the *Circulation* article (McCullough et al) to derive BNP performance characteristics for clinically indeterminate patients with pretest probability between 20% and 80%. For this analysis, we considered BNP as a binary test with a cut-off value of 100 pg/cc. Table 3 suggests that, in this patient group, BNP assays had less than optimal discriminative value.

In response[3] to letters to the editor,[4–6] the authors acknowledged that interpreting BNP levels in a binary man-

Table 2. BNP diagnostic parameters for congestive heart failure using different cutoff thresholds

BNP cut-off	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)
50 pg/cc	97 (96–98)	62 (59–66)	71 (68–74)	96 (94–97)
80 pg/cc	93 (91–95)	74 (70–77)	77 (75–80)	92 (89–94)
100 pg/cc	90 (88–92)	76 (73–79)	79 (76–81)	89 (87–91)
125 pg/cc	87 (85–90)	79 (76–82)	80 (80–83)	87 (84–89)
150 pg/cc	85 (82–88)	83 (80–85)	83 (80–85)	85 (83–88)

BNP = B-type natriuretic peptide; CI = confidence interval; PPV = positive predictive values; NPV = negative predictive values

ner risks losing valuable discriminative information.[3] It is likely that marked elevations are of greater diagnostic strength than marginal elevations; therefore, we used the sensitivities and specificities reported for each BNP range (<50, 50–80, 80–100, 100–125, 125–150 and >150) to calculate corresponding likelihood ratios (Table 4).[7]

Commentary

Shortness of breath is one of the most common chief complaints in the ED, yet the cause of dyspnea is often unclear. A reliable and affordable test that rapidly differentiates CHF from non-cardiac dyspnea in patients with a clinically uncertain diagnosis would be an important advance.

In the studies reviewed, the investigators assessed the performance of a bedside BNP assay as a means of confirming or refuting the diagnosis of CHF. These studies had several strengths. All treating physicians, including the cardiologists who adjudicated the final diagnoses, were blinded to BNP results. BNP results were compared to an independent gold standard in all patients and these results did not influence the decision to perform the gold standard. In addition, the adjudicating cardiologists established the final diagnoses retrospectively, which allowed them to incorporate information collected after the ED visit, such as diagnostic test results, the patients' clinical course, and response to treatment. Although this approach is comprehensive, it has not been validated for the diagnosis of CHF, and there is the potential disadvantage that the cardiologists did not evaluate the patients themselves, but instead relied on chart documentation. The cardiologists agreed on the final diagnosis in 90% of cases; however, given that the treating physicians placed only 27.9% of study subjects in the clinically uncertain category, this high rate of agreement may reflect the fact that the diagnosis was clinically obvious in many cases.

We disagree with the authors' conclusion that BNP, used as a binary test with a cut-off of 100 pg/cc, is useful. At this cut-off, neither the positive nor negative likelihood ratios are strong enough to be clinically useful. Given that the average BNP value for patients with noncardiac dyspnea is 110 pg/cc — above the recommended diagnostic threshold for CHF — it is likely there would be many false-positive results that would increase rather than decrease the need for further diagnostic testing. These shortfalls relate, in part, to using the BNP assay as a binary test and "lumping" all values above or below the cut-off. Clearly, values of 101 pg/cc and 990 pg/cc have different diagnostic and prognostic implications: the former is likely to be a false-positive while the latter is likely to be true-positive.

Likelihood ratios

Likelihood ratios are the most useful single indicator of a test's diagnostic strength, therefore the degree to which it can modify pretest probability and facilitate clinical decision-making.[8] As the positive likelihood ratio (LR+) increases, the test becomes a stronger positive predictor,

Table 3. BNP diagnostic parameters (cut-off = 100 pg/cc) in clinically indeterminate patients*

Sensitivity	Specificity	PPV	NPV	LR+	LR–
79 (72–86)†	71 (66–76)	58 (51–65)	87 (83–91)	2.7 (2.2–3.3)	0.3 (0.2–0.4)

*Clinically indeterminate patients are those with clinical pretest probability of 20%–80%.
†Values in parentheses are 95% confidence intervals.
BNP = B-type natriuretic peptide; PPV = positive predictive value; NPV = negative predictive value; LR+ = positive likelihood ratio; LR– = negative likelihood ratio

Table 4. Likelihood ratios for BNP ranges

BNP range, pg/cc	% of CHF patients with [BNP] in the specified range	% of non-CHF patients with [BNP] in the specified range	Likelihood ratio* (95% CI)
<50	0.03	0.62	0.05 (0.03–0.07)
50–79	0.04	0.12	0.34 (0.23–0.50)
80–99	0.03	0.02	1.46 (0.78–2.74)
100–124	0.03	0.03	1.04 (0.60–1.82)
125–150	0.02	0.04	0.50 (0.27–0.91)
>150	0.85	0.17	5.0 (4.29–5.82)

BNP = B–type natriuretic peptide; CHF = congestive heart failure; CI = confidence interval
*Likelihood ratio = proportion of patients with CHF who had BNP levels within range / proportion of patients without CHF who had BNP levels within range

and as the negative likelihood ratio (LR–) decreases, it becomes a stronger negative predictor. Tests with LR+ greater than 10 can substantially increase post-test probability of disease while those with LR– less than 0.1 can substantially decrease post-test probability. Different BNP values do not modify pretest probability to the same extent; therefore we calculated the likelihood ratios associated with BNP ranges (Table 4). Table 4 shows that BNP values below 50 pg/cc or above 150 pg/cc have likelihood ratios sufficient to alter the post-test probability of CHF, while those between 80 pg/cc and 125 pg/cc are non-discriminating (likelihood ratios near 1.0). For the reasons discussed, we feel it is more meaningful to consider BNP results as negative, indeterminate or positive, somewhat like perfusion scans for the diagnosis of pulmonary embolism.

Spectrum bias

Diagnostic tests do not perform equally well across the spectrum of disease. They are more likely to detect "severe" cases, where the diagnosis is obvious and a test may not, in fact, be necessary, and they are more likely to miss "subtle" cases, where clinicians actually need a test.[8] Spectrum bias exists when test performance parameters are misrepresented for a given situation because the tests were evaluated in a different spectrum of patients than they are being applied to in real life. Consequently, in order to determine how useful BNP assays are in patients with dyspnea of uncertain etiology, the assays should be studied in a population of patients in whom the diagnosis is unclear.[9] The authors attempted to do this by excluding patients whose dyspnea was clearly not related to CHF, such as those with chest trauma or tamponade. They did not, however, exclude patients in whom the treating physicians were clinically certain about the etiology of the dyspnea. In fact, the treating physicians placed only 28% of study patients in the "clinically uncertain" category, and this may have introduced spectrum bias inflating the apparent performance characteristics of the BNP assay.

In order to test this hypothesis, we assessed the performance of BNP using the authors' proposed cut-off of 100 pg/cc in the clinically uncertain group of patients. Table 3 shows that, in these patients, sensitivity and specificity were 79% and 71% — substantially lower than the 90% and 76% reported for the study group as a whole (Table 2). Further, the likelihood ratios shown in Table 3 suggest that, in clinically indeterminate patients, BNP assays used in a binary fashion have insufficient diagnostic strength to influence clinical decision-making. It is likely that very high or very low BNP levels would be more helpful in this sub-group; unfortunately, the authors did not publish the data necessary to determine likelihood ratios for different result ranges.

Summary

The published data do not support the use of BNP as a diagnostic adjunct for CHF in the ED, particularly when it is used as a binary test with a cut-off value of 100 pg/cc. Very low and very high BNP levels may have better discriminative power, but because of the spectrum of patients studied, it is unclear whether these tests will enhance clinical decision-making. Further studies in patients with truly undifferentiated shortness of breath are required.

Competing interests: None declared.

References

1. Harrison A, Morrison LK, Krishnaswamy P, Kazanegra R, Clopton P, Dao Q, et al. B-type natriuretic peptide predicts future cardiac events in patients presenting to the emergency department with dyspnea. Ann Emerg Med 2002;39(2):131-8.
2. Morrison LK, Harrison A, Krishnaswamy P, Kazanegra R, Clopton P, Maisel A. Utility of a rapid B-natriuretic peptide assay in differentiating congestive heart failure from lung disease in patients presenting with dyspnea. J Am Coll Cardiol 2002;39: 202-9.
3. Maisel AS, Clopton P, McCullough PA. B-type natriuretic peptide in heart failure [letter]. N Engl J Med 2002;347:1977-8.
4. Colli A, Fraquelli M, Conte D. B-type natriuretic peptide in heart failure [letter]. N Engl J Med 2002;347:1976.
5. Hassan Y, Shapira AR, Hassan S. B-type natriuretic peptide in heart failure [letter]. N Engl J Med 2002;347:1976-7.
6. Foote RS, Pearlman JD. B-type natriuretic peptide in heart failure [letter]. N Engl J Med 2002;347:1977.
7. Carey M, Brennan J, Ciconelli R, Dayan P, Eskin B, Fabris L, et al. Likelihood ratios when the results of diagnostic tests are continuous variables. Am J Emerg Med 1998;16(7):722-3.
8. Worster A, Innes G, Abu-Laban RB. Diagnostic testing: an emergency medicine perspective. CJEM 2002;4:348-54.
9. Jaeschke R, Guyatt G, Lijmer J. Diagnostic tests. In: Guyatt G, Rennie D, editors. Users' guides to the medical literature: a manual for evidence-based clinical practice / the Evidence-Based Medicine Working Group. 1st ed. Chicago: AMA Press; 2002. p. 121.

Correspondence to: Dr. Corinne Hohl, 4221, ave. de l'Esplanade, Apt. 201, Montréal QC H2W 1T1; 514 204-9497, corinnehohl@primus.ca