Should natriuretic peptide testing be incorporated into emergency medicine practice?

Corinne M. Hohl, MD

SEE RELATED ARTICLE, PAGE 251.

Shortness of breath is a common complaint among adults presenting to the emergency department (ED). The most frequent diagnosis made in this context is heart failure (HF). The prognosis for patients with HF remains poor, with only 45% to 60% of patients surviving 5 years after the diagnosis is made, a rate comparable to that of colorectal cancer.[1–3] Given the high incidence of HF, its poor prognosis and the existence of therapies that improve the duration and quality of life, it is important that emergency physicians make the diagnosis as rapidly and accurately as possible.

Point-of-care assays for brain natriuretic peptide (BNP) and its *N*-terminal pro-peptide (NT-proBNP) are now available. There is literature to support the utility of these assays for the ED diagnosis of HF. However, most publications have been based on post hoc subgroup analyses of a few large industry-sponsored prospective studies.[4–6] In this issue of the *Canadian Journal of Emergency Medicine*, Murray and colleagues report the results of a small but innovative primary study that looked at the potential impact of incorporating NT-proBNP into emergency medicine practice.[7] The results raise questions regarding the assumptions of previous research in this area and shed light on the challenges emergency physicians may face if they choose to adopt natriuretic peptide testing.

Clinical uncertainty is the raison d'être for any new diagnostic test. If clinical uncertainty does not exist, additional diagnostic tests are not warranted. The degree of clinical certainty (or uncertainty) is determined by comparing the clinical impression with an independent validated criterion (gold) standard. Because there is no validated criterion standard for the diagnosis of HF, previous natriuretic peptide studies have compared the ED diagnosis with a retrospective diagnosis by 2 cardiologists who reviewed the medical records and were blinded to natriuretic peptide results and ED diagnoses. This unvalidated criterion standard has been assumed, although never proven, to be more accurate than the diagnostic impression of physicians in the ED.

Ironically, the studies that used a retrospective criterion standard call into question the assumption that emergency physicians face diagnostic uncertainty for all patients presenting with shortness of breath.[4,8] Although it is clear that emergency physicians are uncertain of the diagnosis in approximately 30% of dyspneic patients, they are rarely wrong when they rate the probability of HF as very high or very low.[8] In a study of 1586 patients, the prevalence of disease in patients who were rated by emergency physicians as very likely to have HF was 95%. Conversely, the prevalence in patients who were rated as very unlikely to have HF was only 8%.[8] At these extremes of pretest probability, it would be difficult for any diagnostic test to improve upon the clinical impression of emergency physicians. In fact, in a post hoc analysis[8] of the BNP [Breathing Not Properly] Multinational Study Group data,[4] the accuracy of BNP in very high and very low probability patients was found to be worse than clinical judgment alone.

How do BNP or NTpro-BNP perform in clinically uncertain cases? All studies to date, including the one in this issue of *CJEM*, have evaluated diagnostic test performance in the entire spectrum of acutely dyspneic patients, including clinically certain ones. The appropriateness of extrapolating the findings from such studies to the subgroup of clinically uncertain patients is highly questionable. Two independent post hoc analyses of the BNP Multinational Study Group data calculated diagnostic test performance

Division of Emergency Medicine, Department of Surgery, University of British Columbia, Vancouver, BC

Received: May 25, 2006; accepted: June 3, 2006

Can J Emerg Med 2006;8(4):259-61

characteristics in clinically uncertain patients and concluded that test performance decreases significantly; sensitivity dropped from 90% to 79%, and specificity dropped from 76% to 71%.[8–10] These findings suggest that natriuretic peptide testing may not be useful in the very subgroup of patients for whom a new diagnostic test would be most beneficial.

Although Murray and colleagues did not exclude clinically certain patients, they provided detailed outcomes of clinically uncertain and discordant cases, thus allowing an appreciation of what might happen in real life. In contrast to previous studies, Murray and colleagues compared the clinical impression to an equally unvalidated but more explicit criterion standard: a priori–defined HF-related end points. The objectivity and face validity of using clearly defined, transparent end points is appealing, particularly given the absence of any validated criterion standard. As would be anticipated, the rate of HF-related end points decreased with decreasing clinical certainty of disease. In comparison to clinical impression, NT-proBNP identified a few additional patients who suffered HF-related end points in each of the categories of clinical certainty, but also categorized 86% of *all* clinically uncertain patients and 75% of *all* clinically unlikely patients as positive.

Why did NT-proBNP identify so many patients as positive in the clinically uncertain or unlikely groups? Using manufacturer recommended cut-offs, NT-proBNP categorized 86% of the entire study sample as having HF. The recently proposed PRIDE Study[5] cut-offs resulted in a 71% positive rate. These striking findings mean that either the emergency physicians significantly underdiagnosed HF, or the assay incorrectly categorized many patients as positive. Murray and colleagues found that 51% of their study subjects developed at least 1 HF-related end point, a rate consistent with previous ED studies on adults with acute undifferentiated dyspnea. Therefore, it appears that the latter scenario regarding the test is more likely, where the assay incorrectly identified many patients as positive.

What about the performance of NT-proBNP compared with that of emergency physicians? The physicians correctly identified 65% of patients with HF-related end points. The NT-proBNP assay identified an additional 21% of patients with HF-related end points (10% of the entire patient sample). At first glance this appears promising; however, the "price" of the increased sensitivity of NT-proBNP was an extremely high false-positive rate. Based on HF-related end points, NT-proBNP misclassified 30% of the *entire* patient sample as positive. To put this into context, the use of NT-proBNP would mislabel an additional 30% of patients as suffering from HF in order to correctly identify an additional 10% with the disease. In the real world, such test performance could misguide physicians to inappropriately administer HF treatment, abandon diagnostic work-ups for other serious conditions such as pulmonary embolism or sepsis, and initiate further HF work-ups in one-third of acutely dyspneic patients who do not have the disease!

Why was the performance of NT-proBNP in Murray and colleague's study so much worse than in previous reports? The majority of previous studies on BNP or NT-proBNP were designed to find optimal diagnostic test cut-offs. In such a scenario, cut-off selection is under the control of the researchers and repeated analyses are typically performed until a cut-off yielding maximal sensitivity and specificity is identified. Although such methodology is valid for hypothesis-generating purposes, the performance characteristics of any cut-off derived from such analyses typically overestimate test performance in the real world. Any cut-off derived in this manner requires prospective validation. Often when this is done, diagnostic test performance drops.[11]

Murray and colleagues sought to externally validate 2 NT-proBNP cut-offs in a prospective manner. Although their study was small and conducted at only 1 centre, the results are provocative and underscore what can happen when a promising diagnostic test is applied in real world circumstances. As such, this study provides compelling evidence for the importance of further rigorous prospective studies using a priori–defined cut-offs in clinically uncertain patients. Until this is done, it would be prudent for emergency physicians to be cautious in embracing natriuretic peptide testing — particularly in clinically uncertain patients. For these patients (in whom the misclassification rate is high), any benefit from identifying a few additional HF patients would likely be outweighed by the risk to those whose diagnostic work-ups for other, more immediately life-threatening conditions would be terminated prematurely. Also to consider would be the cost and discomfort experienced by those who would undergo additional and unnecessary investigation for HF. Routine implementation of natriuretic peptide testing for acute dyspnea in the ED should be postponed until future prospective studies are conducted that refute the high false-positive rate suggested by Murray and colleague's study and demonstrate that any benefits from such testing outweigh the risks.

Acknowledgement: Dr. Hohl is supported by a Mentored Clinician Scientist Award from the Vancouver Coastal Health Research Institute.

Competing interests: None declared.

References

1. American Heart Association. 2002 Heart and Stroke Statistical Update. Dallas: American Heart Association; 2001.
2. Levy D, Kenchaiah S, Larson MG, et al. Long-term trends in the incidence of and survival with heart failure. N Engl J Med 2002; 347:1397-402.
3. National Cancer Institute of Canada. Five year relative cancer survival in Canada. Available at: www.ncic.cancer.ca/ncic/internet /standard/0,84658243_85787780_910357 96_langId-en,00.html #fig11_1 (accessed 2006 May 16).
4. Maisel AS, Krishnaswamy P, Nowak RM, et al. Rapid measurement of B-type natriuretic peptide in the emergency diagnosis of heart failure. N Engl J Med 2002;347:161-7.
5. Januzzi JL Jr, Camargo CA, Anwaruddin S, et al. The N-terminal Pro-BNP investigation of dyspnea in the emergency department (PRIDE) study. Am J Cardiol 2005;95:948-54.
6. Maisel A, Hollander JE, Guss D, et al. Primary results of the Rapid Emergency Department Heart Failure Outpatient Trial (REDHOT). J Am Coll Cardiol 2004;44:1328-33.
7. Murray H, Cload B, CollierP C, et al. Potential impact of *N*-terminal proBNP testing on the emergency department evaluation of acute dyspnea. Can J Emerg Med 2006;8(4):251-8.
8. Schwam E. B-type natriuretic peptide for diagnosis of heart failure in emergency department patients: a critical appraisal. Acad Emerg Med 2004;11:686-91.
9. Hohl CM, Mitelman BY, Wyer P, et al. Should emergency physicians use B-type natriuretic peptide testing in patients with unexplained dyspnea? Can J Emerg Med 2003;5(3):162-5
10. Worster A, Innes G, Abu-Laban RB. Diagnostic testing: an emergency medicine perspective. Can J Emerg Med 2002;4(5): 348-54.
11. Sackett DL, Haynes RB. Evidence base of clinical diagnosis: the architecture of diagnostic research. BMJ 2002;324:539-41.

Correspondence to: Dr. Corinne Hohl, Department of Emergency Medicine, Vancouver General Hospital, 855 W 12th Ave., Vancouver BC V5Z 1M9