Clinical Trials

A Randomized Control Trial Using a Validated Prediction Model for Diagnosing Acute Heart Failure in Undifferentiated Dyspneic Emergency Department Patients—Results of the GASP4Ar Study

BRIAN D. STEINHART, MD,[1,2] PHILLIP LEVY, MD, MPH,[3] HILDE VANDENBERGHE, PhD,[4,5] GORDON MOE, MD, MSc,[6,7] ANDREW T. YAN, MD,[6,7] ASHLEY COHEN, MSc,[8] KEVIN E. THORPE, MMath,[8,9] MELISSA MCGOWAN, MHK,[1] AND C. DAVID MAZER, MD[10,11]

Toronto, Ontario, Canada; and Detroit, Michigan

ABSTRACT

Background: Diagnosing acute heart failure (AHF) in undifferentiated dyspneic emergency department (ED) patients can be challenging. We prospectively studied a validated diagnostic prediction model for AHF that uses patient age, clinician pretest probability for AHF, and N-terminal pro–B-type natriuretic peptide (NT-proBNP) as a continuous value to determine its utility and performance.

Methods and Results: This was a multicenter randomized controlled trial of undifferentiated dyspneic patients with an indeterminate pretest probability of AHF as assessed by the treating emergency physician (EP). After recording its components, the calculated model results with validated treatment threshold guidelines were provided to EPs for patients randomized to the intervention arm. Final diagnoses with the use of 60-day follow-up information were adjudicated by 2 independent cardiologists. The primary outcomes were accuracy of the model and of physician diagnosis comparing intervention and standard care arms. A total of 197 patients were randomized and had outcome data recorded; 41% were determined to have had heart failure. Final EP diagnostic accuracy was 76% (sensitivity 68.2%, specificity 83.9%) with no significant difference between exposed versus blinded arms (accuracy 77% vs 74%; $P = .77$). Area under the model receiver operating characteristic curve was 0.93. Using the model treatment thresholds would have redirected 48% of patients with 95% accuracy.

Conclusions: This study prospectively validated the diagnostic accuracy of our AHF model in a significant proportion of indeterminate dyspneic ED patients, but provision of this information did not improve

From the [1]Department of Emergency Medicine, St Michael's Hospital, Toronto, Ontario, Canada; [2]Division of Emergency Medicine, Department of Medicine, University of Toronto, Toronto, Ontario, Canada; [3]Department of Emergency Medicine and Cardiovascular Research Institute, Wayne State University, Detroit, Michigan; [4]Department of Laboratory Medicine, St Michael's Hospital, Toronto, Ontario, Canada; [5]Department of Laboratory Medicine and Pathobiology, University of Toronto, Toronto, Ontario, Canada; [6]Division of Cardiology, St Michael's Hospital, Division of Cardiology, Toronto, Ontario, Canada; [7]Department of Medicine, University of Toronto, Toronto, Ontario, Canada; [8]Applied Health Research Centre, St Michael's Hospital, Toronto, Ontario, Canada; [9]Dalla Lana School of Public Health, University of Toronto, Toronto, Ontario, Canada; [10]Departments of Anaesthesia and Critical Care, Keenan Research Centre for Biomedical Science and Li Ka Shing Knowledge Institute of St Michael's Hospital, Toronto, Ontario, Canada and [11]Department of Anesthesia, University of Toronto, Toronto, Ontario, Canada.

Manuscript received April 12, 2016; revised manuscript received August 10, 2016; revised manuscript accepted August 11, 2016.

Reprint requests: Brian D. Steinhart, MD, 30 Bond Street, 1-008 Shuter Wing, Toronto, Ontario M5B 1W8, Canada. Tel: (416) 864-5976; Fax: (416) 864-5341. E-mail: steinhartb@smh.ca.

Clinical Trial Registration Information: ClinicalTrials.gov Identifier NCT01193998.

Funding: This work was supported by St Michael's Hospital Medical Services Association; the Academic Health Sciences Centres' Alternate Funding Plan Innovation Fund, a joint Ontario Medical Association/Ontario Ministry of Health and Long Term Care initiative (SMHAIF-014); and an unrestricted grant from Roche Diagnostics Canada.

Author Contributions: BDS, HV, GM, KT and CDM conceived the study and designed the trial. BDS obtained research funding. BDS, CDM and MM supervised the conduct of the trial and data collection. BDS, PL, and MM undertook recruitment of participating centers and patients and managed the data, including quality control. AY provided data analysis, interpretation and manuscript revision. AC and KT provided statistical advice on study design and analyzed the data. BDS drafted the manuscript, and all authors contributed substantially to its revision. BDS takes responsibility for the paper as a whole.

See page 150 for disclosure information.

1071-9164/$ - see front matter

http://dx.doi.org/10.1016/j.cardfail.2016.08.007

EP diagnostic accuracy. Future studies should determine how such a clinical prediction tool could be effectively integrated into routine practice and improve early management of suspected AHF patients in the ED. (*J Cardiac Fail 2017;23:145–152*)
Key Words: AHF, Diagnosis, Prediction Model.

Patients presenting to the emergency department (ED) with shortness of breath are a challenge, and diagnostic uncertainty after initial assessment is common for many clinicians.[1–4] Possible etiologies which are important to differentiate include acute heart failure (AHF), chronic obstructive lung disease (COPD), pneumonia, atrial fibrillation, and pulmonary embolus. We previously used randomized control trial (RCT) databases that included ED patients with undifferentiated dyspnea (ie, shortness of breath that, before work-up, was due to an uncertain cause)[5,6] to retrospectively derive and validate a simple AHF diagnostic model that incorporates patient age, N-terminal pro–B-type natriuretic peptide (NT-proBNP) as a continuous value, and gestalt-derived clinician pretest probability of AHF[7] (Equation 1). We found that binary NT-proBNP testing has only modest diagnostic test rule-in performance, and in a significant cohort of patients (44%) with undifferentiated dyspnea the use of model treatment probability thresholds ≤0.20 and ≥0.80 should quickly and accurately redirect the uncertain clinician to rule out or rule in AHF with minimal misdirection.[7] The model reinforced the fundamental premise that a decision tool integrating gestalt and a biomarker value in a formal structured fashion would yield superior diagnostic capabilities over use of its individual components in an unstructured manner. Herein, we report prospective performance of the model and its impact in a prospective randomized controlled trial of its utility as a clinical decision tool.

$$\begin{aligned}\text{Mathematical model equation.}: \Pr(\text{ahf}) &= 1/1+\exp(8+0.011\,\text{age}-5.9\,\text{ptprob}\\ &\quad -2.3\,\text{lntbnp}+0.82\,\text{ptprob}\times\text{lntbnp})\end{aligned}$$

$\Pr(\text{ahf})$ = post-test probability for acute heart failure;
ptprob = patient's pretest probability;
lntbnp = log (base 10) of N-terminal pro−B-type natriuretic peptide value.

Importance and Goals of This Investigation

Demonstration of the superior test characteristics and outcomes using the model in a cohort of ED patients with undifferentiated dyspnea and an indeterminate probability of an AHF diagnosis after initial assessment would support its use in clinical practice, enabling more accurate and timely ED diagnosis of AHF while potentially contributing to better outcomes via improved treatment in both confirmed AHF and non-AHF cohorts. Accordingly, the aims of the present investigation were to compare the diagnostic accuracy of the model and the accuracy of emergency physician (EP) diagnosis (EPDx) between intervention (exposure to model) and standard care (blinded to model) study arms in a similar cohort of ED patients. The secondary aim was to analyze the clinical impact of using the model.

Methods

Study Design and Setting

This was a multicenter randomized controlled trial of undifferentiated dyspneic patients with indeterminate probability of AHF presenting to the ED at 4 sites—Saint Michael's Hospital, Toronto, Ontario, Canada; Detroit Receiving Hospital, Detroit, Michigan, USA; Waikato District Hospital, Hamilton, New Zealand; and Saint Boniface General Hospital, Winnipeg, Manitoba, Canada—from October 2010 to October 2013. Institutional research ethics approval was obtained at all sites prior to commencement of the study. Clinical Trial Registration Information: ClinicalTrials.gov Identifier NCT01193998.

Selection of Participants

Certified staff EPs carried out a bedside assessment as part of routine clinical care to determine probability for AHF among patients with undifferentiated dyspnea. Adult patients with a probability of AHF within our predefined indeterminate range of 21%–79% were eligible for inclusion. After initial EP evaluation including review of history and current chest x-ray (CXR) and electrocardiogram (ECG), but before further testing, patients who met either of the following criteria were eligible for inclusion: i) planned treatment for AHF but requiring further investigation for other causes; or ii) planned treatment for other causes but requiring further investigation for AHF while awaiting further testing and/or observation. Exclusion criteria were age <18 years, suspected acute coronary syndrome, and renal insufficiency (serum creatinine >2.8 mg/dL or >250 μmol/L). Once deemed to be eligible, the patient was approached by the research coordinator or clinical personnel to obtain written informed consent.

Intervention

After patient enrollment, managing clinicians recorded their estimate of the probability of AHF (pretest AHF) as a percentage on a standardized study data collection form and on the central lab requisition for immediate NT-proBNP analysis of the study participant's plasma EDTA sample. As a standard of care, earlier hospital records were available to the EP as well as the current CXR and ECG, to incorporate into the bedside evaluation. Typically, no new laboratory values

would be available at the time of initial evaluation. A blinded secondary physician (usually an emergency medicine postgraduate trainee) also examined the patient, interpreted their ECG and CXR, and rendered an independent pretest probability. Their role was only to allow assessment of interrelater reliability of initial pretext probability for AHF among clinicians of varying experience; they were not involved in the management of the patient, did not receive other clinical testing data or model results, and did not render a final diagnosis.

In the hospital clinical laboratory, NT-proBNP was assayed immediately and the model results for probability of AHF were calculated. Randomization stratified by site was performed with the use of varying block sizes. For those patients randomized to the intervention arm, model results were reported immediately to the managing EP along with recommended AHF treatment for results ≥0.80 threshold and nontreatment for results ≤0.20. EPs remained blinded to model results for patients in the blinded arm. Standard EP management then continued. Follow-up 60-day data were collected on a standardized form. All completed study data forms were sent by Teleform to the Applied Health Research Centre in Toronto. Data forms underwent review by a trained clinical research specialist, and sites were queried for completeness and accuracy before upload into a Microsoft Access database. At study closure, quality assurance was conducted on primary outcome measures.

Methods and Measurements

Model results were calculated with the use of a computerized program of the formula in Equation (1). Before RCT commencement, plasma samples (n = 99) for NT-proBNP were analyzed in real time at the sites with the use of Roche Cobas 6000 platform on the Cobas e601 side (Toronto, Waikato, and Winnipeg sites) or Siemens Vista 1500 platform (Detroit). Internal quality assurance identified a small bias between the platforms in Toronto and Detroit over a range from 20 to 40,000 pg/mL. To prospectively standardize NT-proBNP values across all sites, a Deming regression analysis was applied to the Detroit NT-proBNP results (x) used in the model (y), where $y = 0.821x + 268.1$. Further validation was conducted comparing adjusted and unadjusted (straight) NT-proBNP values with a predetermined data set (n = 27) of patients' age, EP gestalt value (pretest AHF), and NT-proBNP values (pg/mL) to derive the model post-test AHF probability; no significant difference was found, so a sensitivity analysis excluding Detroit data was deemed to be unnecessary.

At the end of ED management, the EP committed to a final diagnosis for the cause of dyspnea, which was recorded as either AHF or other. This was defined as the EPDx. If >1 etiology was contributory to the presenting dyspnea, only the primary etiology was recorded. Study participants were followed for 60 days after the ED index visit with 100% retention and all relevant tests performed (eg, computerized tomography of the thorax, echocardiography, lung scan, pulmonary function and methacholine challenge tests, etc), readmissions, consultations, and death records were collated, deidentified, and sent to the Applied Health Research Centre in Toronto.

Records were reviewed by 2 staff cardiologists blinded to the model result and the EPDx and adjudicated for the determination of the index ED visit diagnosis as either AHF or other. Because several etiologies might have contributed to the presentation, only the primary cause of dyspnea was determined. The process of adjudication was conducted in a sequential fashion: first with ED records to the point of EP disposition but blinded to NT-proBNP value, then exposure to the NT-proBNP value, and finally with 60-day post–ED index visit records. The latter was deemed to be the criterion standard diagnosis (AdjDx). Agreement between cardiology reviewers was good (Cohen kappa 0.82), because there was initial disagreement in only 7% of cases, all of which were reviewed collectively with achievement of consensus without need for 3rd-party arbitration.

Outcome Measures and Analyses

Baseline characteristics were summarized with the use of appropriate descriptive statistics (eg, means and standard deviations for quantitative data and counts and percentages for categoric data). The primary outcomes were diagnostic accuracy compared with AdjDx for the model itself, and EPDx between the intervention (exposure to model) and standard care (blinded to model) study arms. Diagnostic accuracy of the model was determined by calculating the area under the receiver operating characteristic curve (AUC). EPDx accuracy was based on the proportion with agreed diagnoses in each group and compared by means of a chi-square test, with treatment effect expressed as an odds ratio with 95% confidence interval. In addition, secondary outcomes based on clinical impact of the model, including time from physician assessment to EPDx, time to readiness for disposition, total ED length of stay (LOS), intensive care unit (ICU) admission rate, and mortality at 60 days, were compiled and compared between treatment arms with the use of Kaplan-Meier survival analysis for time measures and logistic regression for categoric data.

Accuracy for EPs diagnosing AHF in past studies of undifferentiated dyspneic patients with pretest probabilities of 1%–99% averaged 85% accuracy[5,6]; because this study recruited the more undifferentiated, more challenging cohort with a of pretest probability of 21%–79%, we estimated a control group accuracy of 70%. With a sample size of 75 subjects per group, the model accuracy would need to be ≥92% to have 80% power of detecting significance (alpha 0.05). As such, total enrolment was calculated to be 150 patients. All analyses were performed with the use of R version 3.0.3.

Results

Characteristics of Study Subjects

A total of 201 patients were enrolled, 4 of whom had blood samples misplaced before lab analysis and were excluded, leaving 197 patients randomized with outcome data recorded (see Appendix 1 for the CONSORT flow diagram).

Table 1. Patient Characteristics and Clinical Findings

Characteristic		Blinded Arm	Exposed Arm	Overall
Demographics	Age (y)	63.9 ± 14.9	64.4 ± 14.1	64.2 ± 14.5
	Sex female	53% (54)	60% (60)	57% (114)
	Race African American/Black	55% (51)	53% (50)	54% (101)
	Caucasian/White	31% (29)	41% (39)	36% (68)
History	Congestive heart failure	56% (53)	51% (49)	53% (102)
	Previous coronary artery disease	27% (26)	26% (25)	27% (51)
	Atrial fibrillation	20% (19)	21% (20)	21% (39)
	COPD	39% (38)	58% (56)	49% (94)
	Pulmonary embolism	4% (4)	7% (7)	6% (11)
	Pneumonia	15% (14)	27% (26)	21% (40)
	Renal disease	10% (9)	15% (14)	12% (23)
	Recent weight gain	17% (16)	16% (15)	17% (31)
	Other significant disease	60% (54)	67% (62)	63% (116)
Vital signs	% O_2 saturation	95 ± 7.8	94.4 ± 6.9	94.7 ± 7.3
	Respiratory rate (breaths/min)	21 ± 4	21 ± 4	21 ± 4
	Systolic blood pressure (mm Hg)	151 ± 31.2	145.6 ± 28	148.4 ± 29.7
	Diastolic blood pressure (mm Hg)	81.5 ± 16.9	83.0 ± 18.1	82.2 ± 17.6
	Resting pulse (beats/min)	89.8 ± 19.5	92.6 ± 19.9	91.2 ± 19.7
Clinical signs	Diaphoresis	5% (5)	3% (3)	4% (8)
	Jugular venous distention	18% (17)	20% (19)	19% (36)
	Hepatojugular reflux	9% (8)	13% (12)	11% (20)
	Crackles	42% (40)	43% (40)	42% (80)
	Wheeze	31% (29)	43% (40)	37% (69)
	Cardiac murmurs	16% (15)	11% (10)	13% (25)
	S3 gallop	5% (5)	4% (4)	5% (9)
	New/increased leg edema	53% (49)	47% (44)	50% (93)
ECG findings	New ischemia	4% (4)	3% (3)	4% (7)
	Normal sinus rhythm	68% (63)	67% (62)	68% (125)
	Atrial fibrillation	14% (13)	13% (12)	14% (25)
	Other significant abnormality	46% (42)	45% (41)	46% (83)
Chest x-ray results	Cardiomegaly	58% (53)	62% (58)	60% (111)
	Redistribution vasculature	43% (40)	46% (43)	45% (83)
	Pleural effusion	16% (15)	21% (20)	19% (35)
	Alveolar edema	16% (15)	16% (15)	16% (30)
	Hyperinflation	20% (18)	26% (24)	23% (42)
	New infiltrate	7% (6)	13% (12)	10% (18)

Values are presented as mean ± SD or % (n).
COPD, chronic obstructive pulmonary disease; ECG, electrocardiography.

Patient characteristics included a mean age of 64 years, 43% male, 36% white, 53% with a history of AHF, and 49% with a history of COPD (Table 1). Demographics were matched between investigation arms.

Main Results

Overall, 41% of patients had a final AdjDx diagnosis of AHF. The model had an AUC value of 0.93 (Fig. 1). There was no significant difference in EPDx accuracy for the correct diagnosis in the exposed (n = 101) versus blinded (n = 96) arms (77% vs 74%; $P = .77$; OR 1.17, 95% CI 0.75–1.82), with an overall accuracy of 76% (sensitivity 68.2%, specificity 83.9%). If the validated model diagnostic probability thresholds of ≤0.20 and ≥0.80 had been used to define the absence and presence of AHF, respectively, the model result would have redirected clinicians in 48% of study patients with 95% accuracy. There were no significant differences between arms for any other clinical outcomes or processes of care,

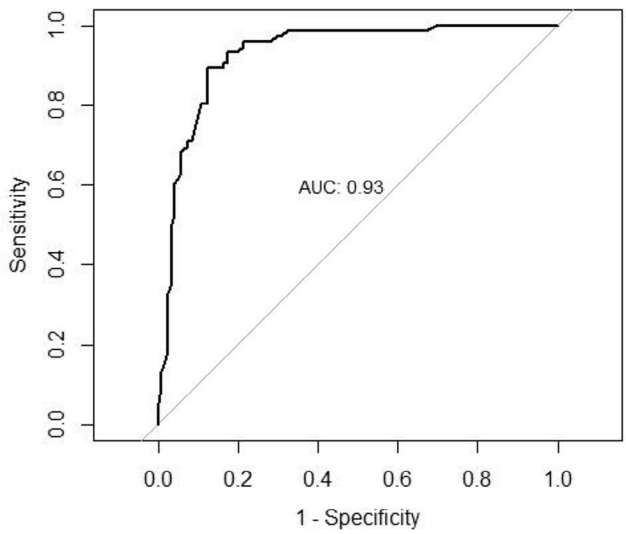


Fig. 1. Model receiver operating characteristic curve (area under the curve [AUC] = 0.93).

Table 2. Secondary Outcomes: Health Benefits

	Blinded Arm	Exposed Arm	*P* Value
Time of EP Care* (h)	2.59 (3.3 ± 2.8) (n = 96)	2.70 (3.2 ± 1.9) (n = 101)	.884
Time to readiness for discharge (h)	3.79 (4.77 ± 3.93) (n = 78)	3.87 (4.70 ± 2.88) (n = 98)	.89
ED LOS (h)	6.35 (8.99 ± 12.16) (n = 78)	5.38 (7.67 ± 7.94) (n = 98)	.366
ICU admission	17%	13%	.499
60-day mortality	97% survival	92% survival	.495

Values in parenthesis are median ± SD.
ICU, intensive care unit; LOS, length of stay.
*Time from initial emergency physician (EP) assessment to end of EP care (ie, consultation or emergency department [ED] discharge).

including time from physician assessment to EPDx, time to readiness for discharge, total ED LOS, ICU admission rate, and 60-day survival (Table 2).

Discussion

This study demonstrates that performance of a diagnostic model for AHF based on patient age, clinician impression, and NT-proBNP as a continuous variable is more accurate than experienced EPs for a large segment of the truly undifferentiated dyspneic ED patient population. The mathematical formula for our model had been derived from a 500-patient database[5] and then validated in a separate 600-patient database[6] from another country. At that time, we concluded that any post-test probability value of ≤0.20 or ≥0.80 would redirect the uncertain physician to respectively rule out or rule in AHF with minimal misdirection, which applied to 44% of patients.[7] With the present prospective trial we duplicated these results: the treatment thresholds would apply to 48% of patients with 95% accuracy. We think that a diagnostic test that yields such accuracy in any challenging (ie, clinically indeterminate) patient population bodes well for its future clinical potential.

Antithetically, we found no appreciable clinical impact as a result of EPs being exposed to the model. Because no significant difference in EPDx accuracy was found between the model-exposed and blinded arms, this would explain why there were no subsequent outcome or process of care benefits found. It appears that physicians had difficulty being directed by a validated post-test probability figure despite educational sessions before commencement of the trial and inclusion of treatment threshold values in the model report that they received reminding them of these guidelines. The reasons for this difficulty remain conjectural. When undecided, it may be that clinicians intuitively preferred a "black and white" binary result to direct them.[8] Another theoretic cause would be the perceived unacceptable delays in NT-proBNP assay testing (a review of study cases at one site revealed a model turnaround time of 1 hour in 73% of cases and 85 minutes in 100%). Solutions to this could either be to measure NT-proBNP as part of a medically delegated "dyspnea panel" on presentation of these patients to the ED or to use NT-proBNP quantitative point of care testing at the bedside. The model could then be calculated online[9] or with a clinical app and the result made available at the time of initial clinical assessment. Despite the model having been derived and validated on 1100 dyspneic patients with known criterion standards, clinicians may still have been skeptical of its ability to accurately direct them. Applying standard knowledge translation strategies[10,11] would further facilitate clinician acceptance of and adherence to model use.

The criterion standard for AHF diagnosis used in this study was the AdjDx rendered by 2 blinded expert cardiologists. As is convention for diagnostic randomized controlled trials on this subject, their final AdjDx was based on guidelines,[12,13] NT-proBNP value, and their expert clinical impression. To address this potential for incorporation bias, a sequential adjudication was undertaken in most (87%) cases, first with standard ED records available but blinded to NT-proBNP value, followed by exposure to the NT-proBNP value, and finally with 60-day post-visit records (100% capture). Analysis showed that 19.7% of initial diagnoses were altered after exposure to the NT-proBNP result but that 27.7% of these reverted back to original diagnoses after further exposure to 60-day records. Ultimately only 14.2% of AdjDx cases could possibly have been influenced by exposure to NT-proBNP. As well, blinded to each other, adjudicators had AdjDx agreement on 93% of cases on 1st review. Therefore we feel that any incorporation bias would have minimum impact and we have confidence in the validity of the criterion standard adjudication that was used.

We chose to select sites from several countries with varying health care systems and ethnicities. Patient numbers (33% above the estimated sample size) and demographic diversities give further confidence in the credibility of our findings. Only the challenging dyspneic patients for whom the clinician was uncertain for AHF diagnosis were studied, ie, the cohort with a pretest probability AHF of 21%–79%[5,6] where diagnostic assistance would intuitively be of real value, making the results of this study more practical. The average staff EP pretest AHF value of 49% supports the appropriateness of the selected study group. In 42% of cases a blinded secondary physician (typically an EM trainee) also examined the patient, interpreted their ECG and CXR, and rendered an independent pretest probability. Average values for this determination were nearly identical to those of the primary physicians (49% vs 48%).[14] This suggests that our findings are stable and independent from variability in pretest inter-rater reliability, regardless of experience. Therefore we think that the design of the study is pragmatic and yields robust and generalizable conclusions.

Bedside ultrasound for assessment of interstitial lung fluid, cardiac function and inferior vena cava size was not routinely used by EPs to influence gestalt, because they required an advanced skill set. Interest in the utility of bedside lung and cardiac ultrasound[15] for undifferentiated dyspnea is growing, but most studies to date have not specifically focused on those ED patients with an indeterminate pretest probability of AHF. Although there is potential for ultrasound to augment pretest probability in such patients, it needs to be more rigorously studied, ideally in concert with the use of our model.

Study Limitations

Because NT-proBNP is dependent on renal excretion, patients with moderate to severe azotemia can have falsely elevated levels and so were excluded from the study, preventing applicability of our study findings to that patient population. The study was carried out with the use of a specific NT-proBNP assay on one platform at 3 sites and a different platform at the 4th site, which required a minor adjustment factor to standardize the model over all sites. All calculations were based on pg/mL units (identical to μg/L), and on subanalysis with the use of adjusted and unadjusted NT-proBNP values from the 4th site, no difference was noted, suggesting that the specific assay platform itself had no impact on study findings. The model has not been studied using other units or other NT-proBNP assays or platforms and should not be extrapolated to them without further study. Although we did not analyze BNP, its operating characteristics are similar to NT-proBNP and would benefit from a similar study.

In this study, the EPDx diagnostic accuracy was only 76%. This is lower than in other AHF studies,[6,16] and the inference could be made that our study results were skewed because managing EPs seemingly were not as experienced as others in diagnosing AHF. However, the other studies enrolled patients with the full range of pretest probabilities (0–100%), not limiting inclusion solely to the indeterminate cohort as in our study. Therefore, it is not surprising to find a lower rate of accuracy in this more challenging population. In effect, our study suffers from "reverse" selection bias. Importantly, all study clinicians who provided pretest and EPDx data were staff physicians board certified in emergency medicine. Therefore, we think that our study EPs had the same skills and experience as those in similar studies.

Conclusion

This study prospectively validates the accuracy of this diagnostic model for AHF in a significant proportion of indeterminate dyspneic ED patients in whom the clinical diagnosis of AHF is uncertain, yet provision of this information did not improve EP accuracy. Future studies of this patient cohort should determine how such a clinical prediction tool can be effectively integrated into routine practice and the extent to which it improves early management in this challenging patient population.

Disclosures

P. Levy: research support, Roche Diagnostics; consultant, Roche Diagnostics, Beckman Coulter, and Siemens; honoraria, Thermofisher.

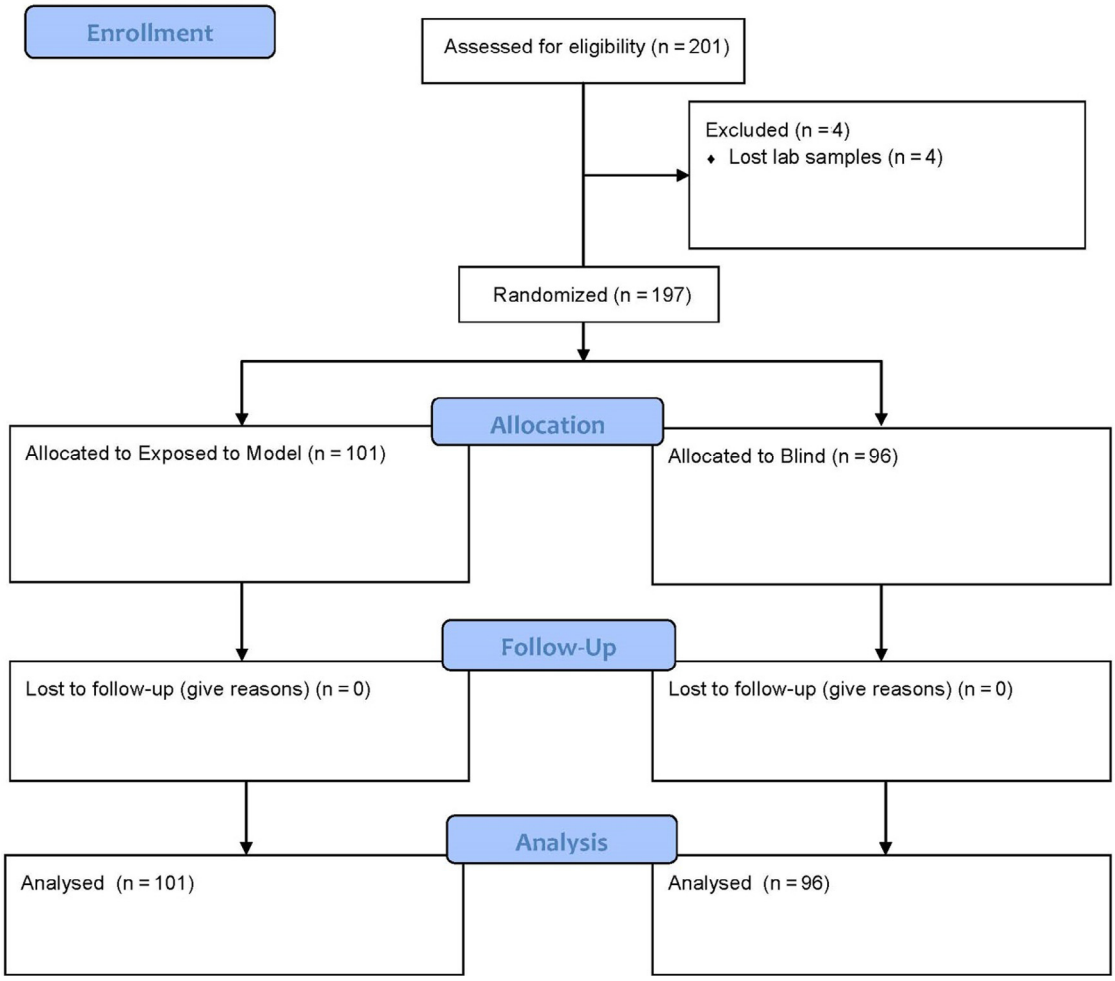
Acknowledgments

The authors graciously acknowledge and thank the following, because this study would not have been possible without their commitment and support: study site personnel including Dr. Donna Clark (Winnipeg); Dr. Gerard Devlin (Waikato) and respective laboratories; Judi Hall and Alice Dang at the Applied Health Research Centre, Saint Michael's Hospital; Carolyn Ziegler, Information Specialist Health Sciences Library, Saint Michael's Hospital; adjudicators Dr Howard Leong-Poi and Dr Akshay Bagai from the Department of Cardiology, St Michael's Hospital, and the participating emergency physicians, nurses, patients, and families.

Appendix

CONSORT
TRANSPARENT REPORTING of TRIALS

CONSORT 2010 Flow Diagram



References

1. Parissis JT, Andreoli C, Kadoglou N, Ikonomidis I, Farmakis D, Dimopoulou I, et al. Differences in clinical characteristics, management and short-term outcome between acute heart failure patients chronic obstructive pulmonary disease and those without this co-morbidity. Clin Res Cardiol 2014;103:733–41.
2. Mueller C, Frana B, Rodriguez D, Laule-Kilian K, Perruchoud AP. Emergency diagnosis of congestive heart failure: impact of signs and symptoms. Can J Cardiol 2005;21:921–4. PMID: 16239975.
3. Wang CS, FitzGerald JM, Schulzer M, Mak E, Ayas NT. Does this dyspneic patient in the emergency department have congestive heart failure? JAMA 2005;294:1944–56.
4. Worster A, Balion CM, Hill SA, Santaguida P, Ismaila A, McKelvie R, et al. Diagnostic accuracy of BNP and NT-proBNP in patients presenting to acute care settings with dyspnea: a systematic review. Clin Biochem 2008;41:250–9. PMID: 17915204.
5. Moe GW, Howlett J, Januzzi JL, Zowall H. N-Terminal pro–B-type natriuretic peptide testing improves the management of patients with suspected acute heart failure: primary results of the Canadian prospective randomized multicenter IMPROVE-CHF study. Circulation 2007;115:3103–10.
6. Januzzi JL Jr, Camargo CA, Anwaruddin S, Baggish AL, Chen AA, Krauser DG, et al. The N-Terminal Pro-BNP Investigation of Dyspnea in the Emergency Department (PRIDE) study. Am J Cardiol 2005;95:948–54.
7. Steinhart B, Thorpe KE, Bayoumi AM, Moe G, Januzzi JL Jr, Mazer CD. Improving the diagnosis of acute heart failure using a validated prediction model. J Am Coll Cardiol 2009;54:1515–21.

8. Croskerry P. From mindless to mindful practice–cognitive bias and clinical decision making. N Engl J Med 2013;368:2445–8.
9. Coyle J. Diagnosing acute heart failure (AHF): a validated prediction model [online calculator]. Available at: http://www.zunis.org/Risk_Calculators/Acute%20Heart%20Failure%20Likelihood%20Calculator.htm. Accessed February 5, 2016, By permission Dr. J Coyle.
10. Straus SE, Tetroe J, Graham ID. Knowledge translation in health care: moving from evidence to practice. 2nd ed. Chichester, West Sussex: Wiley and Sons/BMJ Books; 2013. p 150–62. ISBN: 978-1-118-41354-8.
11. Carpenter CR, Sherbino J. How does an "opinion leader" influence my practice? CJEM 2010;12:431–4. PMID: 20880436.
12. McKee PA, Castelli WP, McNamara PM, Kannel WB. The natural history of congestive heart failure: the Framingham study. N Engl J Med 1971;285:1441–6.
13. Schocken DD, Arrieta MI, Leaverton PE, Ross EA. Prevalence and mortality rate of congestive heart failure in the United States. J Am Coll Cardiol 1992;20:301–6. PMID: 1634664.
14. Steinhart B, Levy P, Vandenberghe H, Moe G, Cohen A, Thorpe K, et al. A simple clinical decision tool for diagnosing acute heart failure in the undifferentiated dyspneic patient using a Bayesian approach (abstr). CJEM 2015;17:S6.
15. al Deeb M, Barbic S, Featherstone R, Dankoff J, Barbic D. Point-of-care ultrasonography for the diagnosis of acute cardiogenic pulmonary edema in patients presenting with acute dyspnea: a systematic review and meta-analysis. Acad Emerg Med 2014;21:843–52.
16. McKelvie RS, Moe GW, Ezekowitz JA, Heckman GA, Costigan J, Ducharme A, et al. The 2012 Canadian Cardiovascular Society heart failure management guidelines update: focus on acute and chronic heart failure. Can J Cardiol 2013;29:168–81.