Acta Diabetologica (2019) 56:39–43
https://doi.org/10.1007/s00592-018-1215-y

ORIGINAL ARTICLE

Autonomic testing: which value for each cardiovascular test? An observational study

Federico Bellavere[1] · Eugenio Ragazzi[2] · Nino Cristiano Chilelli[3] · Annunziata Lapolla[3] · Giuseppe Bax[3]

Received: 14 February 2018 / Accepted: 13 August 2018 / Published online: 29 August 2018

Abstract

Aims Cardiovascular autonomic testing is time consuming when adopting the entire Ewing battery of tests, hence, clinicians usually adopt an empirically reduced number of tests which may give controversial results. Our purpose was to examine the reliability of the cardiovascular tests most commonly used in autonomic diagnoses.

Methods We tested 334 subjects, from an original group of 3745, who had shown an altered deep breathing test to both Lying to standing and Valsalva manoeuvre, assuming a value of postural hypotension of more than 15 mmHg as a sign of almost true dysautonomia.

Results VM showed the highest sensitivity (85%) and, when coupled to LS, highest specificity (83%).

Conclusions VM could be useful when screening for possible or early autonomic neuropathy, VM + LS is useful as a diagnostic tool for probable or advanced autonomic neuropathy, and VM + LS + PH is useful for certain diagnosis of definite or late stage autonomic neuropathy.

Keywords Autonomic neuropathy · Cardiovascular tests · Postural hypotension · Diabetic neuropathy

Introduction

Cardiac autonomic neuropathy (CAN) is now recognized as a serious complication of diabetes which leads to poor quality of life and reduced life expectancy for patients [1]. CAN is usually diagnosed using a battery of simple, non-invasive bedside cardiovascular autonomic tests which were first introduced in the late 1970s and have since been shown to be a reliable tool in clinical practice [2, 3]. The so-called Ewing battery of autonomic tests comprises analysis of heart rate variations during deep breathing (DB), lying to standing (LS), standardized Valsalva manoeuvre (VM), analysis of blood pressure variations on standing (PH) and sustained handgrip (SHG). It has, for a long while, been considered as a "gold standard" for the diagnosis of autonomic disorders. However, this battery, although originally described as taking only 15 min, in reality requires almost 30 min to administer especially, as often happens when dealing with complications of diabetes, when used for patients with motor illnesses. Hence, more recently, many physicians have adopted a "reduced" number of the tests included in the original battery resulting, at times, in controversial and/or uncertain diagnoses. Indeed, each test in the Ewing battery explores different pathways of cardiovascular reflexes and, so far, no evidence has been offered that could justify the choice of one rather than another parameter even though a scoring system has at times been adopted so as to give a semblance of grading and/or uniformity to the diagnosis [4].

This study offers an attempt to verify the diagnostic reliability of each test in the original Ewing battery with the purpose of indicating the degree of effort required to achieve different degrees of diagnostic accuracy.

Managed by Antonio Secchi.

✉ Giuseppe Bax
pino.bax@unipd.it

[1] Endocrinology and Diabetes section "Rizzola" Foundation Hospital, San Donà, Venezia, Italy

[2] Department of Pharmaceutical and Pharmacological Sciences, University of Padova, Padova, Italy

[3] Department of Medicine, University of Padova, Diabetology and Dietetics, ULSS 6 Euganea, via dei Colli, 35143 Padova, Italy

Subjects and methods

The study initially involved 3745 diabetics, who had undergone the DB test as previously described [3, 5], using a specially devised computerized system for the evaluation of RR beat-to-beat variations on an ECG trace. DB was determined as the mean ratio between the three longest and the three shortest RR intervals during breathing paced at 6 breaths per minute (IR). Of the 3745 patients in the study, 902 (24%) showed a DB below the 10th percentile, with a cut off value of IR 1.11. After a few months, 334 of these 902 repeated the DB test together with LS, VM and PH. However, SHG was not performed, as today it is considered of little interest in clinical practice. Values of LS and VM below the 10th percentile were also considered altered, while for PH a cut-off value of − 15 mmHg was taken as a reliable index of almost true dysautonomia. This cut-of point was chosen in the attempt of including all possible patients having overt dysautonomia, as a cut-off value of − 10 mmHg does not ensure a reliable diagnosis of CAN. On the other hand, patients showing a cut-off value of − 15 mmHg are almost considered to be affected, otherwise a − 20 mmHg cut-off could rise the number of false-negative patients.

The tests were performed by trained nurses, under supervision of well-experienced medical staff to ensure the test was correctly performed.

Continuous values were expressed as means ± standard deviations (SD). Logistic regression was used to assess the ability of DB, LS and VM tests to identify individuals with "overt" CAN, based on PH positivity (≤ − 15 mmHg after 2 min standing). ROC analysis was then performed to find the most favorable discriminatory values by testing the relationship between the true positive rate (sensitivity) and the false positive rate (1—specificity) for each DB, LS and VM cut-off. The area below the ROC curve (AUC) was calculated to measure the ability of the test to discriminate between patients or, in other words, to measure its accuracy.

Results

Table 1 shows the clinical features of the 334 patients with previous DB impairment who underwent further tests, while in Table 2 the rate of different anti-hypertensive treatments is specified.

Figure 1 shows the percentage distribution of normal/abnormal values (identified as 0 and 1) of previously identified normal/abnormal DB in the context of the more comprehensive battery which now included LS, VM, PH as well as the repeated DB. A wide divergence between the diagnostic reliability of the tests emerged, as there was poor agreement between their positive/negative values. In the case of repeated DB, it should be noted that 79 (23.65%) of the patients with previously abnormal values, showed values in the upper 10th percentile, which suggests that, in the medium/long term, there is a low degree of reproducibility for this test.

As regards the specificity and sensitivity of each of the three heart rate-based tests (Table 3), in relation to the abnormal PH results, i.e. with systolic blood pressure fall > 15 mmHg was considered a reliable marker of true CAN, while VM offered the best sensitivity (84.6%), however, the best specificity was given by LS (72.6%).

In an attempt to achieve better specificity than that obtained with LS alone, which was not entirely satisfactory for clinical purposes, a second logistic regression analysis, which considered test pairs, was then performed (Table 4). It was shown that LS linked to VM gave a diagnostic specificity of 82.8%, while VM, when linked to DB, failed to

Table 1 Individual characteristics 334 patients with type 2 diabetes previously recognized with DBIR impairment

	Subjects with DBIR impairment ($N = 334$)
Gender (m/f)	209/124
Age (years)	58.9 ± 10.1
Duration of disease (years)	12.9 + 8.5
BMI (kg/m^2)	30.7 ± 5.0
Deep Breathing (DBIR)	1.12 ± 0.05
Lying to Standing (LS ratio)	1.11 ± 0.09
Valsalva manoeuvre (Valsalva ratio)	1.37 ± 0.26
BP variations on standing (PH mmHg)	0.17 ± 10.47

Continuous variables are expressed as mean ± standard deviation

Table 2 Summary of anti-hypertensive drugs used by patients (monotherapy or combination)

Drugs	No. of treated patients	% of treated patients
Beta-blockers	10	2.99
ACEi	150	44.91
α1-antagonists	50	14.97
ARBs	40	11.98
Diuretics	100	29.94
Calcium channel blockers (SA)	5	1.50
Calcium channel blockers (LA)	90	26.95
Antiarrhythmics	10	2.99

ACEi, inhibitors of angiotensin-converting enzyme; ARBs, Angiotensin II receptor blockers

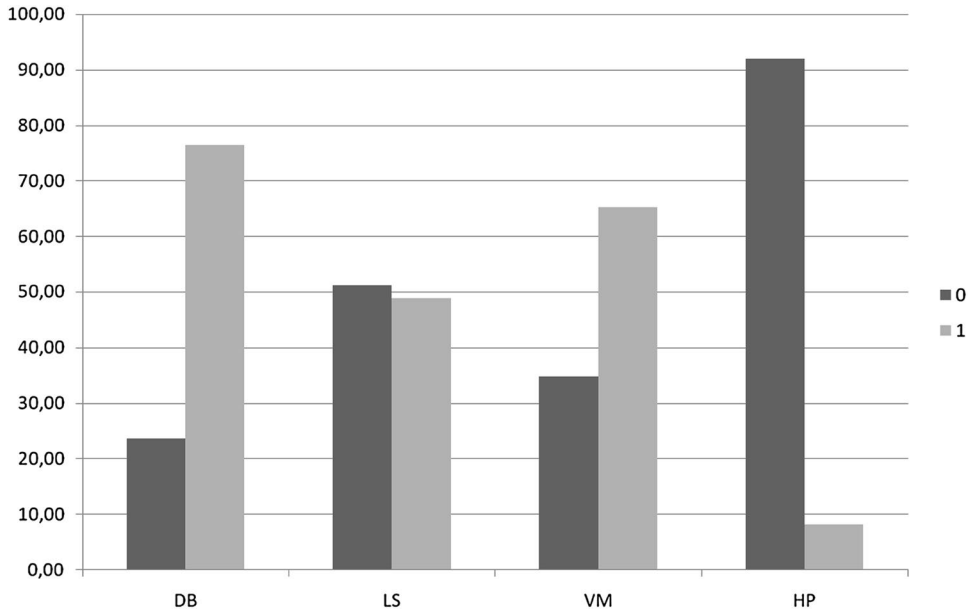

Fig. 1 Percentage distribution of normal/abnormal values (identified as 0 and 1) for deep breathing (DB), lying-to-standing (LS), Valsalva manoeuvre (VM), postural hypotension (PH) tests performed on 334 patients

Table 3 Logistic regression and ROC analysis of the single tests *versus* HP's predictive value < − 15 mmHg, considered as reference

Test	AUC	Cut-off value	Sensitivity	Specificity
DB	0.698	1.09	0.667	0.654
LS	0.747	1.05	0.704	0.726
VM	0.674	1.33	0.846	0.482

DB, deep breathing; LS, lying to standing; VM, Valsalva manoeuvre

Table 4 Logistic regression and ROC analysis of the various tests selected as pairs (total negative: 00 or total positive: 11) versus HP's predictive value < − 15 mmHg, considered as reference

Test	AUC	Sensitivity	Specificity	*p* value
VM + LS	0.615	0.420	0.828	0.001
DB + LS	0.601	0.331	0.881	0.011
VM + DB	0.564	0.284	0.895	0.103

DB, deep breathing; LS, lying to standing; VM, Valsalva manoeuvre

adequately predict CAN as the ROC analysis did not achieve statistical significance. However, LS linked to DB also gave a high specificity of 88.1%, although with lower significance when compared to LS linked to VM ($p = 0.011$ versus $p = 0.0012$).

Discussion

CAN is a dreadful complication in diabetes, as well as in other illnesses [6]. It leads to poor quality of life because its symptoms make patients unsteady on standing, weak on walking, unable to tackle physical and environmental stress and, it often leads to the patient becoming bedridden which, in its turn, can lead to the onset of other illnesses. This just to mention CAN's contribution to pathophysiology of cardiovascular system; however, it is well known that it is a marker of a widespread autonomic neuropathy which involves many other body areas and functions giving rise to a myriad of other frustrating symptoms.

The widely adopted Ewing battery of cardiovascular autonomic tests is a well-established tool for the diagnosis of CAN, as well as a powerful prognostic marker for early death from any cause, not only from cardiovascular morbidity and mortality [1, 7–10].

Today, despite this mostly confirmed evidence, little attention is being paid to the clinical value of adequate autonomic testing and to suggest that physicians detect autonomic dysfunctions in a clinical setting, especially in the early stages of the disease, when the therapeutic intervention can be useful for combatting symptoms and prolonging life. This seeming reluctance to spend time diagnosing CAN is probably related to an old idea that little can be done, except offer sympathy and "commiseration", for patients affected by diabetic neuropathy [11]. However, it is still hard to explain exactly why these reliable, low cost, non-invasive and easy to perform bedside tests are so often neglected in clinical practice. One reason for not using them could be the time required to perform them; the original Ewing battery requires up to 30 min and we are well aware that this is a relatively long time when examining patients.

Although other diagnostic tools based on heart rate variability (HRV) analysis on resting have been proposed since the last 30 years, in the attempt of saving time with comparable, if not better, sensitivity [12–14], the Ewing's

battery still remains the most useful tool in autonomic testing. This is probably due to the technical simplicity of performing these tests, which do not require advanced mathematical or computer knowledge to be interpreted—as generally needed in HRV analysis [12]. Anyway, as Ewing's battery is undoubtedly the most used tool in autonomic testing, we analyzed it in the attempt to obtain a reliable clinical performance of reducing time.

Hence, to find a more acceptable balance between the time and the effort required and reliability of the Ewing's battery of cardiovascular tests, we analyzed the single diagnostic power of each test, taking the PH as a reliable marker of CAN.

An initial analysis of our data, based on a cohort of almost 4000 diabetics, revealed a modest degree of medium-term (6 months) reproducibility for DB which, however, has been used in many studies as a marker for CAN [15].

Moreover, a clear lack of agreement between test results shows that when used to diagnose autonomic dysfunctions the tests are very differently oriented, i.e. they probe different anatomical sites and functions of the autonomic system which have little common basis. This raises the question of which test or tests are most reliable when diagnosing CAN. We adopted an arbitrary, but clinically justified, method of comparing the results of each test to values of PH < -15 mmHg as an expression of true dysautonomia.

Our data showed that the best sensitivity was derived from VM and the worst from DB, with LS falling somewhere in between. However, VM alone showed poor specificity which was better, although not entirely satisfactory, for LS (Table 3).

Analyzing the degree of specificity that can be achieved when considering pairs of tests, we found that high specificity was achieved with VM + LS (83%) and DB + LS (88%) even though the latter showed lower statistical significance. However, MV + DB did not even achieve statistical significance in ROC analysis.

What was found in this study, could be useful in clinical practice as it can be argued that performing VM alone can be considered sufficient when screening for CAN: it allows a diagnosis of "possible" CAN. On the other hand, adding LS to VM would offer enough specificity to permit a diagnosis of "probable" CAN and, when their altered values are associated to an altered PH test, the diagnosis of "confirmed" CAN should be made. Using only three tests saves time and overturns previous interpretations of autonomic neuropathy grading [3, 4].

Lastly, our data do not suggest using the DB test, except as a means of confirmation, either for screening or for diagnostic purposes. Furthermore, recent evidence from a cohort controlled study, which revealed its low prognostic worth, would suggest removing its role as a routine bedside test, something which could save precious time in clinical practice [16, 17].

What this paper adds:

1. It is already known that the most reliable standard in autonomic testing, the classical Ewing battery of cardiovascular tests, is time consuming for routine clinical purposes. But, most physicians arbitrarily adopt only some tests taken from the battery with the result that they are unlikely to achieve comparable conclusions in a clinical setting
2. Our study offers a balanced evaluation of reliability of each test based on its sensitivity and specificity. Our result suggest using the Valsalva manoeuvre for screening of "possible" autonomic neuropathy while, if this latter is linked first to Lying to Standing and then to Postural Hypotension it could be useful to improve the accuracy of diagnosis from probable to definite autonomic neuropathy.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval All procedures followed were in accordance with the institutional guidelines and therefore the ethical standards as laid down in the 1964 Declaration of Helsinki and its later amendments.

Informed consent Written informed consent was obtained from all the study participants prior to their inclusion in the study.

References

1. Vinik AI, Maser RE, Mitchell BD, Freeman R (2003) Diabetic autonomic neuropathy. Diabetes Care 26:1553–1579
2. Wheeler T, Watkins PJ (1973) Cardiac denervation in diabetes. Br Med J 8:584–586.
3. Ewing DJ, Clarke BF (1982) Diagnosis and management of diabetic autonomic neuropahy. Br Med J. 285:916–918
4. Bellavere F, Bosello G, Fedele D, Cardone C, Ferri M (1983) Diagnosis and management of diabetic autonomic neuropathy. Br Med J 287:61
5. Spallone V, Bellavere F, Scionti L et al (2010) Recommendations for the use of cardiovascular tests in diagnosing diabetic autonomic neuropathy. Nutr Metab Cardiovasc Dis 21:69–78
6. Bellavere F (1995) Heart rate variability in patients with diabetes and other non cardiological diseases. In: Malik M, Camm AJ (eds) Heart rate variability. Futura Publishing Co., Armonk, pp 507–516
7. Ewing DJ, Campbell IW, Clarke BF (1976) Mortality in diabetic autonomic neuropathy. Lancet 1:601–603
8. Pop-Busui R (2010) Cardiac autonomic neuropathy; a clinical perspective. Diabetes Care 33:434–441
9. Lykke JA, Tarnow L, Parving H-H, Hilsted J (2008) A combined abnormality in heart rate variation and QT corrected interval is a

strong predictor of cardiovascular death in type 1 diabetes. Scan J Clin Lab Invest 68:654–659
10. Young LH, Wackers FJ, Chyun DA et al (2009) Cardiac outcomes after screening for asymptomatic coronary artery disease in patients with type 2 diabetes. The DIAD study: a randomized controlled trial. J Am Med Assoc 301:1547–1555
11. (1983) Editorial: diabetic neuropathy: where are we now? Lancet:1366–1367
12. Bellavere F, Balzani I, De Masi G et al (1991) Power spectral analysis of heart-rate variations improves assessment of diabetic cardiac autonomic neuropathy. Diabetes 41:633–640
13. Task force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology (1996) Heart rate variability: standards of measurement, physiological interpretation and clinical use. Circulation 93:1043–1065
14. Tang ZH, Wang L, Zeng F et al (2014) Bayesian estimation of cardiovascular autonomic neuropathy diagnostic test based on short term heart rate variability without a gold standard. BMJ Open 4(9):e005096
15. Vinik AI, Ziegler D (2007) Diabetic cardiovascular autonomic neuropathy. Circulation 115:387–397
16. Chyun DA, Wackers FJ, Inzucchi SE et al (2015) Autonomic dysfunction independently predicts poor cardiovascular outcomes in asymptomatic individuals with type 2 diabetes in the DIAD study. SAGE Open Med 3:2050312114568476
17. Cha SA, Yun JS, Lim TS et al (2016) Diabetic cardiovascular autonomic neuropathy predicts recurrent cardiovascular diseases in patients with type 2 diabetes. PLoS One 11(10):e0164807