Differential diagnosis of wide QRS tachycardias: comparison of two electrocardiographic algorithms

Elisabeth Kaiser*, Francisco C.C. Darrieux, Silvio A. Barbosa, Rodrigo Grinberg, Andre Assis-Carmo, Julio C. Sousa, Denise Hachul, Cristiano F. Pisani, Eva Kosa, Carlos A. Pastore, and Mauricio I. Scanavacca

Arrhythmia Unit—Heart Institute (InCor), University of Sao Paulo Medical School, Av. Dr. Eneas de Carvalho Aguiar, 44, Sao Paulo, SP 04011-060, Brazil

Received 13 August 2014; accepted after revision 6 November 2014; online publish-ahead-of-print 19 January 2015

Aims

This study’s aim is to compare the ability of two ECG criteria to differentiate ventricular (VT) from supraventricular tachycardia (SVT): Brugada et al. (horizontal plane (HP) leads) and Vereckei et al. (frontal plane (FP), specifically aVR lead), having electrophysiological study (EPS) as gold standard. After comparing, suggestions for better diagnosis of wide QRS-complex tachycardia (WCT) in emergency situations were made.

Methods and results

Fifty-one consecutive patients with 12-lead ECG registered during EPS-induced regular WCT were selected. Each ECG was split into two parts: HP (V1–V6) and FP (D1–D3, aVR, aVL, and aVF), randomly distributed to three observers, blinded for EPS diagnosis and complementary ECG plane, resulting in total 306 ECG analyses. Observers followed the four steps of both algorithms, counting time-to-diagnosis. Global sensitivity, specificity, percentage of incorrect diagnoses, and step-by-step positive/negative likelihood ratios (+LR and −LR) were calculated. Kaplan–Meier curve was plotted for final time-to-diagnosis. Inter-observer agreement was assessed with kappa-statistic. Global sensitivity was similarly high in FP and HP algorithms (89.2 vs. 90.1%), and incorrect classifications were 27.4 vs. 24.7%. Forty-eight correct analyses by Vereckei criteria took 9.13 s to diagnose VT in the first step, showing that first step was fast, with high +LR, generating nearly conclusive pre- (72.6%) to post-test (98.0%) changes for VT probability.

Conclusion

Both algorithms as a whole are similar for diagnosis of WTC; however, the first step of Vereckei (initial R in aVR) is a simple, reproducible, accurate, and fast tool to use. The negativity of this step requires a ‘holistic’ approach to distinguish VT from SVT.

Keywords

Wide QRS tachycardia • Brugada algorithm • Vereckei algorithm • Ventricular tachycardia • Supraventricular tachycardia • Electrocardiographic diagnosis

Introduction

Ventricular arrhythmia, particularly ventricular tachycardia (VT), is the most common cause of sudden cardiac death. Therefore, establishing prompt diagnosis has a special importance. If tachycardia is present, the first exam to be performed in the emergency room is an electrocardiogram (ECG).

Despite of decades of meticulous evaluations of the ECG, the electrocardiographic differential diagnosis of tachycardias with regular wide QRS continues to be a great challenge. These tachycardias may be ventricular (VT) or supraventricular (SVT) of origin. The latter may have a presentation with aberrant conduction or ventricular pre-excitation. A variety of criteria has been proposed for the differentiation of these tachycardias using morphological and/or numerical aspects. One of the most common differential diagnostic methods is the algorithm proposed by Brugada et al., which uses the horizontal plane (HP) leads of the ECG.

Vereckei et al. proposed a novel algorithm using frontal plane leads (FP) and later using only aVR lead for this purpose.
The goal of this study was to compare the ability of two four-step algorithms (Brugada and Vereckei) to diagnose wide complex tachycardias, since the differential diagnosis of this condition is still a challenge. Additionally, we intended to observe whether our results might contribute to the improvement of the diagnosis of WCT at the emergency department and intensive care unit.

**Methods**

All consecutive patients with regular WCT and indication for EPS were enrolled in an observational and prospective study at our institution. A 12-lead ECG was registered during the EPS in the presence of induced regular wide QRS tachycardia. Each ECG was cut into two parts according to the two electrocardiographic planes: frontal plane (FP) and horizontal plane (HP). The FP, comprising leads D1, D2, D3, aVR, aVL, and aVF, was used for the diagnostic method suggested by Vereckei (Figure 1A). The HP, with leads V1–V6, was used for the algorithm suggested by Brugada (Figure 1B).

Three cardiologists, specialized in arrhythmias, participated as independent observers. Each of them analysed all of the ECG cuts at different time and place. They were blinded for the results of the EPS, for all the clinical and demographic information about the patients and for the complementary part of the ECG plane, which they analysed at a particular moment.

Each observer followed the four sequential steps of both algorithms in an open book fashion. They recorded in a spreadsheet the diagnosis established by the applied method (HP or FP), the step used for making the definitive conclusion, and the time spent, in seconds, for the conclusion. If the diagnosis was inconclusive even after the fourth step of any of the algorithms, the result was considered VT. The decisions were made without conferring between the observers. To compare the algorithms analysis, the EPS results were considered as the gold standard for the final diagnosis.

**Statistical analysis**

For statistical analysis, the Statistical Package for Social Science (SPSS), a statistical software version 15 was used. Global sensitivity, specificity, misdiagnosis rate, and the time spent on diagnosis were calculated for each diagnostic method.

The Kaplan–Meier curve (Breslow test) was plotted for the analysis of the time spent until the establishment of the diagnosis. Mann–Whitney test was used to compare the correct diagnosis rates in both algorithms and to compare the time spent in step I of the FP and the corresponding diagnostic time of the HP. The inter-observer concordance was established by kappa-test (k-statistic). The receiver operating characteristics (ROC) curve was obtained for both algorithms and the comparison between them was done by t-test.

**Results**

According to the 51 EPS, there were 37 cases of VT (7 fascicular tachycardias) and 14 cases of SVT (72.55 and 27.45%, respectively), resulting in a total of 306 ECG analyses performed by the three observers, 153 in each electrocardiographic plane (Figures 2 and 3).

The global sensitivity was high and similar in both methods (90.1% in HP vs. 89.2% in FP). The global specificity was similar in both methods (35.7% in HP vs. 28.6% in FP). The area under the ROC curve was 0.62 in HP and 0.59 in FP (P = 0.07). The percentage of incorrect classification was on average 24.9% in HP and 27.5% in FP (P = 0.14).

The Kaplan–Meier curve showed that the longer was the time until conclusion the higher was the rate of misdiagnosis. When the time until the diagnosis exceeded 25–30 s in both algorithms, the number of incorrect classifications significantly increased for all observers, especially in FP (Figure 4).

**What’s new?**

- The aim of this study was to compare the ability of two four-step algorithms (Brugada and Vereckei) to diagnose wide complex tachycardias, since the differential diagnosis of this condition is still a challenge.
- The first step of Vereckei criteria was fast, easy to use, and obtained high score for correct diagnosis, with high inter-observer concordance.

The sample ECG in (A) frontal plane and (B) horizontal plane, as received by the observers.
The kappa index showed a moderate to good inter-observer concordance in the FP (k = 0.612 95% CI 0.328–0.896) and it was negative in the HP (k = –0.088 95% CI −0.170 to −0.005).

Among the 153 ECG analysed by the FP criteria, 49 were diagnosed as VT in the first step, with 48 confirmed cases by EPS (31.8%) in a mean time of 9.13 s. Using HP algorithm, 26 cases of the 153 were diagnosed as VT in the first step (19 confirmed as correct by EPS) in a mean time of 9.9 s. According to step I analysis in both algorithms, there were less incorrect diagnoses using FP than HP (2.1 vs. 26.9%; P = 0.002).

Considering the 48 correct ECG analyses in the first step of FP criteria (mean time of 9.13 s), 47 were also diagnosed on corresponding HP, although the majority of them were achieved in the subsequent steps, in a mean time of 26.53 s (P < 0.001) (Figure 5).

The systematic analysis of the two algorithms, according to sensitivity, specificity, accuracy, and LRs, demonstrated a high specificity in both algorithms, with low sensitivity (Table 1).

Another relevant aspect of the first step of FP algorithm was the inter-observer agreement. As observed in Table 2, there was concordance between three observers in 10 cases (30 analyses) and concordance between two observers in 5 cases (10 analyses).
Among the seven fascicular tachycardia (21 analyses per plane), 15 were correct in the HP method, while 11 were correct in the FP method ($P = 0.20$).

Six cases in the FP method and two cases in the HP method were classified as inconclusive ($P = 0.15$).

### Discussion

The reliability or validity of a diagnostic test has to include two aspects: the reproducibility and the exactness. The reproducibility relates to the concordance of the results when the test is repeated in similar circumstances by the same observer or when other observers perform the test. Low reproducibility may result from problems with the test itself or when the test requires interpretation. The more the test is subjective, higher is the possibility of disagreement, even among experts. The reproducibility of a test is measured by the kappa-statistic ($\kappa$). The possibility for different interpretations (subjectivity) of one or more steps of the HP algorithm resulted in a disappointing negative $\kappa$, showing more disagreement than inter-observer agreement. The FP algorithm revealed a moderate to good concordance index, which is indicative of acceptable reproducibility. The FP criteria, however, produced slightly higher percentage
of incorrect classification compared with the gold standard than the HP criteria, although this difference is statistically non-significant.

The exactness is measured by sensitivity, specificity, area under ROC curve, predictive values, and the positive and negative likelihood ratio (+LR and −LR, respectively). High levels of exactness and reproducibility lead to the recognition of a good diagnostic test. Failure to fulfill any of these characteristics threatens the clinical usefulness of the diagnostic test in focus.4,5

Among regular WCT, VT is the most prevalent, between 70 and 75%, as shown in previous studies.1,3,6,7 Our data showed a 72.55% prevalence of VT. Therefore, it would be ideal to have a high-sensitivity diagnostic test, even if its specificity was somewhat lower.8 In our study both algorithms showed high global sensitivity but moderate to low global specificity. However, sensitivity and specificity are useful by themselves only when both have high values.

Concerning the average time spent to complete the whole algorithms, we observed that HP was shorter (31 s) than FP (54 s). Probably this difference was related to the fact that HP algorithm was more widely used in our service, in addition to the difficulty presented by steps III and IV of the FP algorithm. However, as shown in Figure 4, the FP algorithm yielded better results than HP in the first 25–30 s, in terms of correct diagnosis. Beyond this, the first step of FP criteria showed to be very fast.

The intriguing observation mentioned above aroused our interest to calculate the sensitivity, specificity, accuracy, and positive and negative LRs step-by-step. In this analysis, different from the global results, each step of both algorithms showed lower sensitivity and higher specificity. This apparent contradiction is due to the fact that as the cases of VT are eliminated from the contingent, a relatively high number of ‘true-negative’ cases remains, i.e. cases of SVT.

As stated earlier, the sensitivity and the specificity need to be high in order to be useful. For this reason, we analysed our data according to the newfangled and powerful LRs, which constitute one of the best ways to measure and express diagnostic accuracy.6,9 The main advantage of LR over other measures of diagnostic accuracy, such as sensitivity and specificity, is its usefulness to compare different diagnostic strategies. As seen in Table 1, the greatest value of +LR is the first step of the FP algorithm (+LR = 18.16). This level of +LR increased the pre-test probability for VT from 72.55 to 98% post-test probability, nearly to a conclusive change. On the other hand, when the first step of the FP criteria is not diagnostic, the VT diagnosis does not rule out, because its −LR change the pre-test probability of VT from 72.55% only to 60% post-test probability.

The other steps of both algorithms have +LR greater than 1, arguing for the diagnosis of VT, but some steps have +LR or −LR very close to 1, meaning they almost lack diagnostic value. As we discussed previously, the step I of FP increases the probability of VT from 72.55 to 98%, and the subsequent steps show little contribution to increase this percentage. The step I of HP does not alter the pre-test probability, with steps II and III showing a ‘probabilistic leap’ (Figure 6). In the same figure, we can observe that the two methods at the end of the four steps arrive at a practically same level of probability of diagnosis of VT. However, the two algorithms were designed in a way that the positivity of any step makes the diagnosis of VT and not proceed the remaining steps. Because of this, the accuracy and the rapidity of each step is important. As shown in Table 2, of the total diagnoses performed by step I of FP, concordance was observed among three observers in 10 cases (30 analyses, corresponding to 61.2%) and between two observers in five cases (10 analyses or 20.4%). The correct diagnosis made only by one observer and the wrong diagnosis represent 16.3 and 2%, respectively. In this context, to calculate the k-value was not possible, since it can only be performed at the end of the diagnostic strategy in all agreements, regardless of whether the diagnosis is right or wrong.

The first step of the FP criteria is particularly useful. In 48 of our cases (see again Figure 5), the correct diagnosis was made within 9.13 s, with the use of only this step of the FP. Otherwise, the same cases took 26.53 s for the diagnoses using the HP. This can be exemplified returning to Figure 1A and B, where the tachycardia could be easily diagnosed as VT using only aVR lead, which depicted a pure R inscription (first step of FP). On the other hand, the same case could have the diagnosis done only on the morphologic analysis (fourth step) of HP, using the R/S relationship in leads V5 and V6. Despite the step I of HP being also quick, no further detailed analysis has been made, due to the considerable number of false-positive diagnoses.

It was a consensus among the observers that steps I and II of the new algorithm (FP) were easy to use, but steps III and IV showed to be more difficult, often requiring a magnifying glass, as observed also in another study,10 limiting their use in the real world, especially in the emergency room. Beyond that, the use of the third and the fourth steps prolonged the time to reach a conclusion and increased the possibility of incorrect diagnosis (Figure 4). For this reason, the authors suggest that the first step of FP should be the first approach to every WCT. Only when this approach is not diagnostic, other approaches or criteria should be used, including clinical data.11 Our findings, although relevant in the diagnosis of VT, do not obviate the clinical decision. This ‘holistic approach’ implies to analyse the WCT in all aspects, using other ECG criteria, in addition to clinical data, such as haemodynamic status, previous structural cardiac disease history, and other relevant features for each case.

![Figure 6](https://example.com/figure6.png) **Figure 6** Probabilistic increment of correct VT diagnosis step by step of both algorithms. Considering that the prevalence of VT in our material was 72.55% (pre-test probability), observe that the first step of FP increases the probability to 98%; in the subsequent steps the probability remains virtually stable. The graphic of the steps of HP exhibits a slow increase of the probability.
Conclusion

The Brugada and Vereckei algorithms as a whole are similar for the diagnosis of WCT. However, the step I of the Vereckei algorithm (FP) is a simple and helpful tool for making a differential diagnosis of wide QRS regular tachycardia. This step obtained high score for correct diagnosis, with short time, and high inter-observer concordance that justifies its use in the emergency room. If concern persists after step I of the Vereckei algorithm, we recommend the careful analysis of the whole ECG using other criteria, like the Brugada algorithm (HP), the subsequent steps of Vereckei algorithm, and others, altogether with the patient’s clinical data.

Acknowledgements

We thank Márcia Dancini and Rogério Ruscitto do Prado for their respective technical and statistical support.

Conflict of interest: none declared.

References