NITIN SHARMA, PATRICK THUMM, DIETMAR SAUPE

A simple algorithm for ventilatory threshold estimation

Measures of oxygen consumption (VO_2) are used in basic assessments specifically of endurance capacity since many years. The most common practice is to conduct these tests as ramp or incremental step tests on a treadmill or a bicycle ergometer. Within these protocols a ventilatory threshold (VT) value can be found which is separating the aerobic from the partly anaerobic metabolism. The VT is mainly used to guarantee a functional training of athletes or to compare this value in pre-post design studies (Hansen and Sue, 2012).

To extract the VT from VO_2 tests the preferred way in many clinics, universities or laboratories is a visual inspection based on given criteria. These subjective results highly depend on the experts. For this reason a number of algorithms were proposed until 1992 to assess the VT objectively. However, the results of the algorithms differ widely and, thus, there still is no generally accepted method to algorithmically compute the VT from spiroergometric data. For a definition and critical review of all methods see Ekkekakis, Lind, Hall, and Petruzzello (2008). After over 20 years we contribute a new algorithm with the goal to correctly predict the VT also in cases of 'abnormal' VO_2 curve progressions where previous algorithms often fail.

Materials and methods

The bike ergometer ramp tests were individually designed to achieve a duration of 10 to 12 min with a power increase between 19 and 55 W per minute, starting from 40 to 110 W. All respiration values where recorded breath-by-breath (Ergostik, Geratherm Respiratory, Germany). Each test yielded a sequence of n values of VE, VO_2, VCO_2, power, and heart rate (n varies). Our algorithm is based on VCO_2 vs. VO_2 data. It searches for the maximal set of k consecutive data points that are close to the line through the origin with slope m. The VT estimate is defined as the mean of the k VO_2 values, see Fig. 1 (right). If there are no k consecutive points close to the line, then k is reset to 1 and the search is restarted. We recommend the parameters $m = 1.02$ and $k = 4$ when $n \approx 400$. For smaller data sets, k should be reduced accordingly. We analyzed the tests of 45 males and 3 females (40,8±17y, 177,5±7,5cm, 77,4±9,8kg, sedentary to amateur level) and acquired 'ground truth' by two experts that independently estimated the VT by visual inspection and, in case of disagreement, after discussing, arrived at a common VT. We compared results with three objective algorithms: Brute-Force, Break-Point, and V-Slope.

Results

For all data sets and each method we computed the differences between VT estimates and corresponding subjective expert ratings in terms of VO_2 and power. The figure and the table below show the differences of algorithmic and expert ratings.

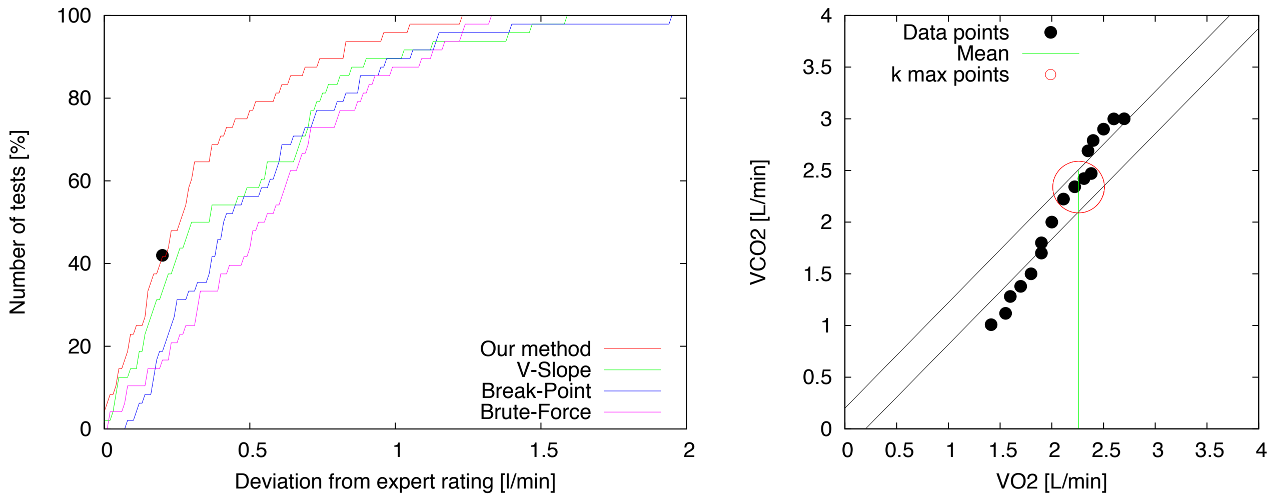

Fig. 1: Left: Novel performance graph showing the percentage of computed ventilatory thresholds (N=48) as a function of the accuracy given by the *maximal* deviation from the 'ground truth' VT. E.g., with our method 20 of 48 computed VTs (42 %) yielded VO_2 values differing from the 'ground truth' by at most 0.2 l/min (see the black dot). The areas under the graphs correspond to quality of the algorithms. Right: Illustration of our algorithm. The two lines indicate the neighborhood of the line of slope *m*.

Tab. 1: Mean differences and median absolute differences of VT estimates and expert ratings. *The mean differences for our method differ significantly from those of all other methods (p<0.001).

	Our method	V-Slope	Break-Point	Brute-Force
VO_2 mean diff. (l/min)	0.21±0.39*	-0.30±0.54	-0.51±0.41	-0.30±0.60
VO_2 median abs. diff. (l/min)	0.26	0.30	0.41	0.53
Power mean diff. (W)	15±31*	-34±47	-46±37	-34±52

Discussion

We have introduced a new objective method for estimation of the ventilatory threshold. It is simple yet provides better matches with subjective expert ratings than previous state-of-the-art methods, especially in 'difficult' cases. The limitations of our study are as follows. We compared results only with three of about ten previous algorithms, and only two expert opinions were considered. Therefore, a larger study is required to confirm our preliminary findings. We plan to apply crowdsourcing to gather a wide spectrum of expert ratings. The resulting mean opinion scores and variances will be used to train and validate our algorithm and combinations of algorithms for VT estimation.

A problem of conventional algorithms is that each one fixes the data type used (like VCO_2 vs. VO_2) while in subjective estimation all types are considered. Thus, we also propose to apply machine learning techniques to provide a method for selecting the best algorithm and most suitable data category for a given data set.

References

Ekkekakis, P., Lind, E., Hall, E. E., Petruzzello, S. J. (2008). Do regression-based computer algorithms for determining the ventilatory threshold agree? *Journal of Sports Sciences*, *26*(9), 967-976.

Hansen, J. E., Sue, D. Y. (2012). Principles of exercise testing and interpretation: including pathophysiology and clinical applications. Philadelphia: Lippincott Williams & Wilkins.