School of Kinesiology
University of Minnesota
Driven to Discover®

Popularity and prevalence of gas exchange data processing methods in peer-reviewed literature: a scoping review

Anton Hesse, MS
College of Education + Human Development

Introduction

- Cardiopulmonary exercise testing (CPET) is the gold standard for non-invasively measuring cardiorespiratory fitness and is popular in many fields.
- CPET analyzes expired gases, usually generating one data point per breath (breath-by-breath). Consequentially, this data requires processing before later analysis because respiratory variability is much higher than the underlying metabolism it reflects.
- There is no universal method to process breath-by-breath data, and previous research shows that different approaches can affect downstream results.
- There is little research on the popularity of different data processing methods and how often they are described in peer-reviewed literature.

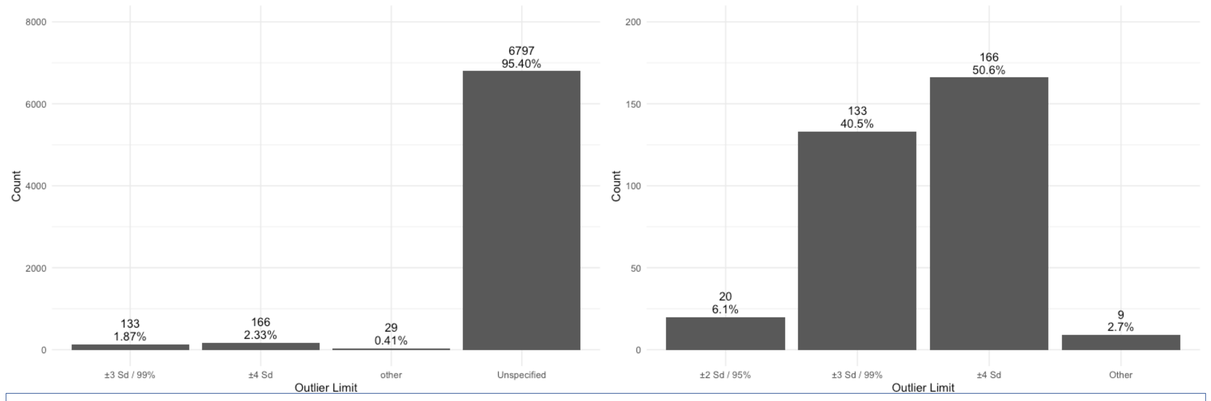
Methods

- This scoping review follows the Preferred Reporting Items for Systematic reviews and Meta-Analyses extension for Scoping Reviews
- Download a list of original, peer-reviewed articles from 3 databases
- Download full-text articles
- Convert articles to txt format
- Normalize article text for easier analysis
- Screen articles not satisfying search criteria
- Use regular expressions to find studies that collected data breath-by-breath
- Use regular expressions to extract text regarding outliers, interpolation, and data averaging

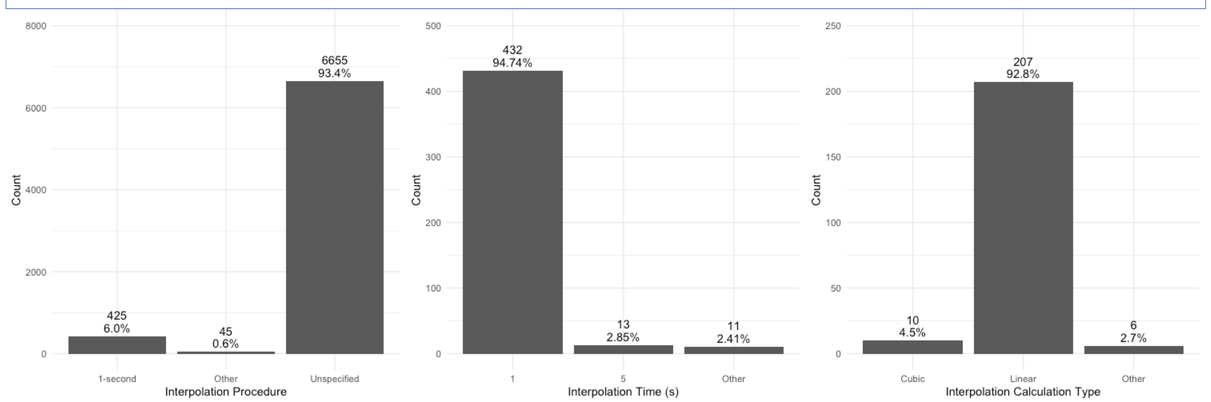
Results

- Of the 7119 articles analyzed, 328 (4.6%) described outlier removal, 470 (6.6%) described interpolation, and an estimated 4366 (61.3%) described averaging methods.
- The most popular outlier cutoffs are mean ± 3 or 4 SD (40.5% and 50.6%, respectively).
- When documented, the dominating interpolation time frame and procedure were one second (94.7%) and linear interpolation (92.8%), respectively.
- Time-based bin averages (84.6%) were the most popular averaging methods.
- 30-second bin averages are the most common averaging methods

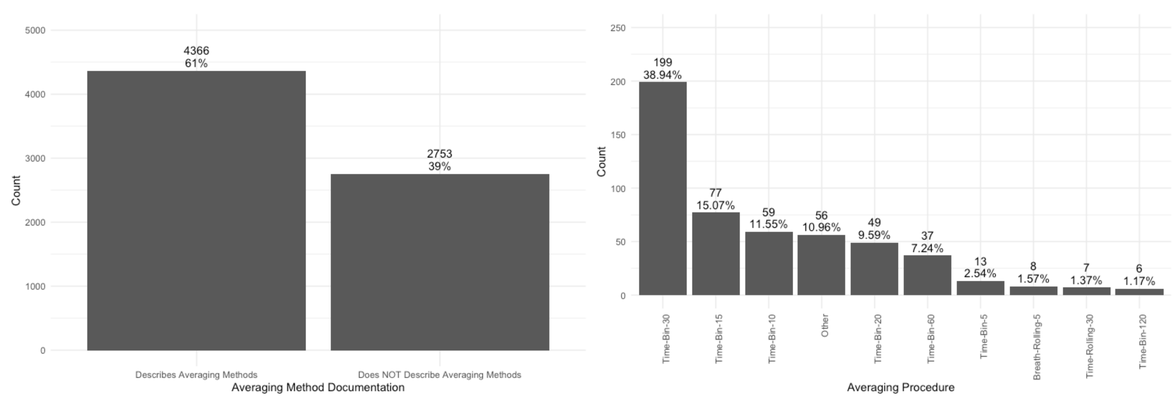
The prevalence of data processing reporting for variable breath-by-breath data is low



The most common outlier limit is ± 3 or 4 standard deviations



1-second and linear interpolation are the most prevalent



Time-based bin averages, especially every 30 seconds, are the most common breath-by-breath averaging method

Discussion

- Very few papers describe their outlier removal and interpolation methods
- Only 60% of papers describe their averaging methods, despite known implications of averaging on VO_2max
- The lack of interpolation methods is less concerning because this applies to mainly VO_2 kinetics studies
- More research should include descriptions of breath-by-breath data processing to improve reproducibility and comparability between studies. This is especially important for data averaging.