**Popularity and prevalence of gas exchange data processing**

**methods in peer-reviewed literature: a scoping review**

**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, and it is common to collect one data point per breath, that is, breath-by-breath. Notably, the variability in respiratory measures is much higher than that of the underlying metabolism, so some data processing is required before later analyses. There is no universal method to process breath-by-breath data, and previous research shows that different approaches can affect later results. Moreover, there is little research on the popularity of different data processing methods and how often they are described in peer-reviewed literature. This scoping review aims to document the popularity and prevalence of CPET data processing methods.

**Methods:** We followed current scoping review methodology and searched three databases for papers whose methods included CPET with gas exchange. We screened those results using regular expressions and machine learning to identify papers where authors collected gas data breath-by-breath. Using regular expressions, we documented which, if any, gas data processing procedures were described in the methods section regarding outlier removal and interpolation.

**Results:** Of the 7833 articles analyzed, 330 (4.2%) described their outlier removal procedures, and 472 (6.0%) described some interpolation procedures. The most popular outlier cutoffs are mean ± 3 or 4 SD (40.9% and 50.3%, respectively). When documented, the dominating interpolation time frame and procedure were one second (94.7%) and linear interpolation (92.8%), respectively.

**Discussion:** Few papers explicitly describe gas data processing regarding outlier removal and data interpolation. The importance of these steps depends on the type of subsequent analysis, but it may be helpful to increase such documentation for reproducibility and study comparisons.