Comparison Between Smartwatch-Derived and CPET-Measured VO$_2$max

Alexandra Jamieson$^1$, Siana Jones$^1$, Claire Steves$^2$, Nicholas Timpson$^3$, Nishi Chaturvedi$^1$, Alun D Hughes$^1$, Michele Orini$^1$

**Background.** Maximal oxygen consumption (VO$_2$max) during exercise and the percentage of predicted VO$_2$max (%pVO$_2$max) are standard measures of cardiorespiratory fitness with established clinical predictive value. They are usually measured during a cardiopulmonary exercise test (CPET), which requires expensive equipment and experienced personnel. Novel smartwatches estimate VO$_2$max using proprietary algorithms, but their accuracy remains undetermined.

**Aims.** To determine smartwatch device accuracy in estimating VO$_2$max and %pVO$_2$max using data from CPET as a reference.

**Methods.** 215 adults (44 (21%) male; median [interquartile range; IQR] 56 [32, 62] years old) were recruited from two population-based cohorts, the Avon Longitudinal Study of Parents and Children (ALSPAC) and TwinsUK. Participants performed a maximal CPET on a semi-recumbent bike to measure VO$_2$max and wore a Garmin Vivoactive 4s (GV4) smartwatch for 60 days following CPET. The first and last VO$_2$max estimates provided by GV4 were compared to CPET measures. %pVO$_2$max was measured using predicted VO$_2$max from Wasserman and Whipp’s anthropometric-based equations. In a subgroup wearing the smartwatch during CPET, peak heart rate (HR) from GV4 and CPET was also compared. Agreement was assessed using Bland-Altman analysis (bias and limits of agreement (LoA)), absolute percentage error (APE), reported as median [interquartile range], and Pearson’s correlation coefficient (cc).

**Results.** VO$_2$max and %pVO$_2$max measured during CPET was 22.4 [17.5, 27.4] ml/kg/min and 90.9% [78.1%, 101.3%], respectively. VO$_2$max estimates from GV4 were moderately correlated with CPET measures (cc=0.62 and 0.66 for first and last estimates) and showed a large positive bias ~14 ml/kg/min with LoA from 0 – 27 ml/kg/min. Correlation between VO$_2$max from GV4 and anthropometric-based prediction of VO$_2$max was high (cc>0.90). Agreement between %pVO$_2$max from GV4 and CPET was poor (cc≈0.15, bias ~52%, LoA 7-98 %).

**Conclusions.** GV4 provides estimates of VO$_2$max that overestimates but moderately correlates with CPET measured VO$_2$max. The agreement for %pVO$_2$max is poor.