**VALIDITY OF COMMERCIALLY AVAILABLE WEARABLE DEVICES FOR MEASURING STEPS, ENERGY EXPENDITURE, AND HEART RATE BY GENDER, AGE, WEAR LOCATION AND COUNTRY**

**Introduction**

Level of physical activity varies according to age groups, gender, country and health status, whether healthy or not healthy.

In the past decade, there has been an increasing number of studies (Wahl et al, 2017) assessing the validity of wearables that measure energy expenditure by comparison to a criterion standard such as indirect calorimetry or accelerometry. Many of these validation studies were conducted among healthy adult participants (Feehan et al, 2018), with studies reporting conflicting findings, both overestimation (Feehan, 2018) and underestimation (Imboden, 2017) with Fitbit devices were reported. In a recent systematic review (Fuller, 2019) investigating the validity of Fitbit devices, it was reported that 49% (43 of 88 comparisons) overestimated energy expenditure, particularly during physical activity. In an earlier systematic review of the field, Evenson et al (2015) reported a high validity of different brands of wearable activity tracking devices regarding step count when compared to various criterion standards made in laboratory settings (Bunn et al, 2018). Regarding the validity of heart rate estimations from by wrist-worn activity tracking devices, one study (Boudreaux, 2018) have shown that the agreement between true rate and the estimated rate made by a wrist-worn device is higher during rest than during MVPA in healthy subjects.

There is little to no knowledge about the validity of these wearables in measuring heart rate, step count and energy expenditure based on gender, age group, non-healthy population, location where the device is worn and the country where it is being used. Therefore, the purpose of this study was to provide information about the difference in the validity of wearables in measuring step count, heart rate, and energy expenditure in the male and female populations, healthy and non-healthy populations, children, adults and older adults population, by different countries, and by the different parts of the body where the device was worn.

**METHODS**

**Design**

This study is a descriptive quantitative design which seeks to describe and compare the differences in the criterion validity of the wearables based on gender, age, health status, part of the body and the country location.

The research question to be answered will be “Does variability in gender, age, health status, country and part of body on which the device is worn affect the validity of wearable devices in measuring step count, heart rate and energy expenditure?”.

**Null Hypothesis**

There will be no difference in the validity of wearable devices due to gender, age, health status, country and part of body on which the device is worn in measuring step count, heart rate and energy expenditure.

**Data**

The publicly available dataset used in this study was the resulting data of a systematic review carried out by Fuller (2019) on the validity of consumer wearables, and was downloaded from the BeapLab Dataverse website (Harvard Dataverse, 2019).

**Analysis Plan**

All quantitative analyses and plots will be done using RStudio version 1.4.1106 (RStudio Inc) and R version 4.0.4 (The R Foundation). The dataset will be imported into RStudio, cleaned and analyzed.

The variables that are necessary for the analysis will be checked for missing values. This includes population of male, population of female, mean age, the location where the study was carried out, the part of the body the device was worn, the mean percentage error, the study setting, and the outcomes measured. Correlation comparisons and group percentage difference will be used as metrics for criterion validity. Where group percentage differences were not reported, group percentage error was calculated as ([wearablemean – criterionmean]/criterionmean × 100) to allow for comparison across studies.

The mean age of the participants in each study will be used to categorized them into groups as categorized by the CDC for age-recommended guidelines for Physical Activity:

* Children (5 – 11)
* Adolescents (12 – 17)
* Adults (18 – 64)
* Older adults (65 and above)

The percentage of males and females in a particular study will be determined, and the gender with the higher percentage will be the labelled as the study’s main gender. Healthy and non-healthy populations, and the part of the body where the device is worn on for the activity will remain the same.

The objective of this study is to check the differences based on the country in which the study was carried out. The locations of the studies will be recoded as a single country and not by states. All states in USA will be recoded as USA only and the same process will be done for other countries too.

The groups will be visualized using bar charts and all groups with less than 30 values will be excluded and filtered out when doing the comparisons. Missing values will be removed, replaced or calculated for, depending on the variables under which they occur.

The interpretation of measurement accuracy will be focused on acceptable limits of percentage difference of ±3% in controlled settings and percentage difference of ±10% in free-living settings, as outlined in previous work (Feehan et al, 2018).