

UK Biobank

Updated derivation of MET scores

Version 1.0

<http://www.ukbiobank.ac.uk/>

June 2023



1. Overview

The MET (Metabolic Equivalent Task) score data in [Category 54](#) was derived according to International Physical Activity Questionnaire (IPAQ) [guidelines](#). In early 2023 it was noticed that the published returned dataset differed somewhat from the methodology published by IPAQ. The published data in Category 54 was re-derived accordingly in June 2023 from the relevant data in [Category 100054](#) to better align with IPAQ guidelines. This document describes the changes made in the derivation of MET score data from the original.

In addition to updating the existing values, data was added from post-baseline assessments to the point of publication.

2. Differences in derivation

2.1. Treatment of outliers

The original derivation recoded all time values for walking, moderate exercise and vigorous exercise exceeding 180 minutes to 180 minutes. This is outlined in section 7.4 of the IPAQ guidelines, to permit a maximum of 21 hours each week (3 hours * 7 days) to be reported for each category. However, section 7.2 previously states that extreme outliers where the sum of walking, moderate and vigorous time variables exceeds 960 minutes (16 hours) should be excluded from analysis. The original derivation recoded these values to 180 minutes rather than removing them. The update excludes these cases from the data.

2.2. Treatment of missing data

Section 7.1 of the IPAQ guidelines states that if data is missing from a case for any time or day variables, then that case should be removed from analysis. However, the original published return recoded the missing values to 0 if the participant had provided complete time and day answers for other activities. For example, if a participant missing time per day spent walking had answered all questions for moderate and vigorous exercise, the days and time per day spent walking were recoded to 0. The 2023 updates removes cases where any response was missing (after data cleaning to remove outliers, etc.) to align with the guidelines.

2.3. Responses of “Unable to walk”

[Data-Field 864](#) (“Number of days/week walked 10+ minutes”) allows for a response of “Unable to walk” in the corresponding [Data-Coding 100307](#), which was not explicitly provided for in IPAQ guidelines. The original derivation removes these values from the dataset, along with “Do not know” and “Prefer not to answer.” In the updated derivation, this has been interpreted as 0 days per week spent walking 10+ minutes.

3. Results of the revised derivation

The above changes result in approximately 17,000 more missing baseline values in this dataset than the previously published version, plus a smaller number of (< 500) additional values. Researchers may therefore wish to refresh their data if using extractions released before July 2023.

The touchscreen questionnaire responses from which these variables are derived remain available in [Category 100054](#), so researchers are able to recreate the original derivation if they wish.