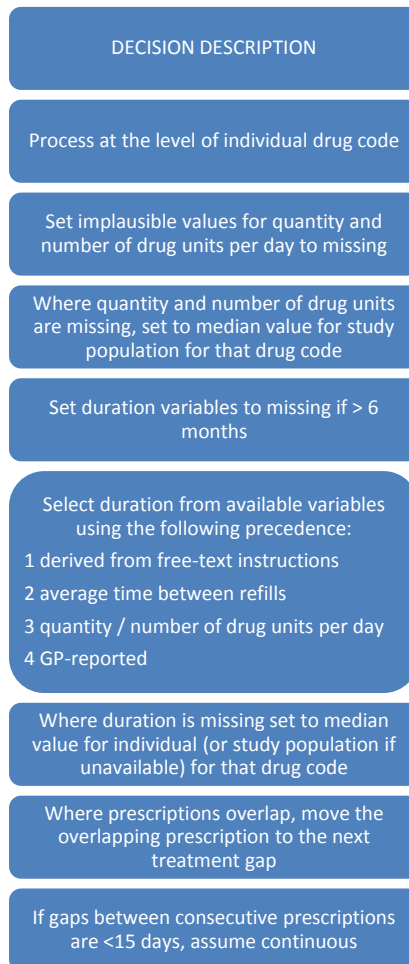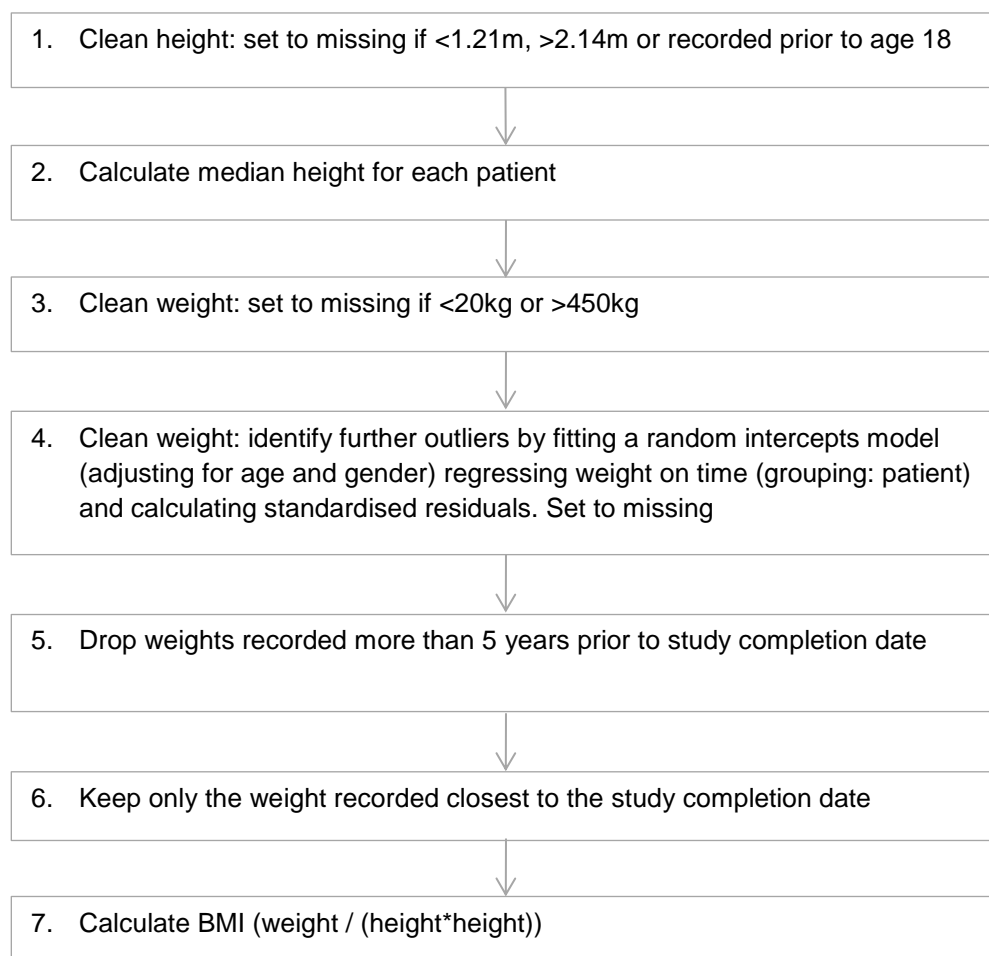# Supplementary File 3 Definition of variables using CPRD datasets

3.1 Exposure to oral glucocorticoids / disease-modifying anti-rheumatic drugs

An algorithm, coded using Stata, was used to define current use of these medications from the CPRD prescription data. The decisions taken are summarised in the flow diagram below. The algorithm can be shared by contacting the authors.

DECISION DESCRIPTION

Process at the level of individual drug code

Set implausible values for quantity and number of drug units per day to missing

Where quantity and number of drug units are missing, set to median value for study population for that drug code

Set duration variables to missing if > 6 months

Select duration from available variables using the following precedence:

1 derived from free-text instructions

2 average time between refills

3 quantity / number of drug units per day

4 GP-reported

Where duration is missing set to median value for individual (or study population if unavailable) for that drug code

Where prescriptions overlap, move the overlapping prescription to the next treatment gap

If gaps between consecutive prescriptions are <15 days, assume continuous

3.2 Defining body mass index

1. Clean height: set to missing if <1.21m, >2.14m or recorded prior to age 18

↓

2. Calculate median height for each patient

↓

3. Clean weight: set to missing if <20kg or >450kg

↓

4. Clean weight: identify further outliers by fitting a random intercepts model (adjusting for age and gender) regressing weight on time (grouping: patient) and calculating standardised residuals. Set to missing

↓

5. Drop weights recorded more than 5 years prior to study completion date

↓

6. Keep only the weight recorded closest to the study completion date

↓

7. Calculate BMI (weight / (height*height))

Defining BMI within the Clinical Practice Research Datalink. RA rheumatoid arthritis, BMI body mass index. Method developed by Ruth Costello, ARUK Centre for Epidemiology, The University of Manchester.

Figure reproduced from:

Joseph, R. M., Movahedi, M., Dixon, W. G. and Symmons, D. P. M. (2016), Smoking-Related Mortality in Patients With Early Rheumatoid Arthritis: A Retrospective Cohort Study Using the Clinical Practice Research Datalink. Arthritis Care & Research, 68: 1598–1606. doi:10.1002/acr.22882