



**Definitions of Motor Neurone Disease and the Major Diagnostic Subtypes,
UK Biobank Phase 1 Outcomes Adjudication**

Date: March 2018

Version: 1.0

**Documentation prepared by:
Kathryn Bush, John Nolan and Cathie Sudlow
On behalf of UK Biobank Outcome Adjudication Group**

Definitions of Motor Neurone Disease; UK Biobank Phase 1 Outcomes Adjudication

Data sources on which the algorithm relies are UKB baseline assessment data (verbal interview); linked hospital admissions data (HES APC, SMR01, PEDW); death register data.

Definitions & Abbreviations:

MND	Motor Neurone Disease
HES APC	Hospital Episode Statistics - Admitted Patient Care (England)
SMR01	Scottish Morbidity Records – General / Acute Inpatient and Day Case Admissions (Scotland)
PEDW	Patient Episode Database for Wales
EHR	Electronic Health Records
Finished Consultant Episode	The basic counting unit for statistics of admitted care Hospital EHR data (= a row of data in the data extracts provided) is a finished consultant episode (FCE).
Code date	The start date of the FCE is taken as the code date.
ICD 9	International Classification of Diseases, Version 9 (SMR01 only)
ICD 10	International Classification of Diseases, Version 10
Prevalent Case	First known hospitalisation with a relevant diagnostic code prior to recruitment, or self-reported event at recruitment.
Incident Case	First known hospitalisation with a relevant diagnostic code post recruitment, or cause-specific death, in those without indication of prevalent event as defined above.

Background:

The term 'Motor Neurone Disease' (MND) describes a group of conditions, which affect the motor nerves and central nervous system. They are progressive, degenerative, neurological conditions that lead to muscle wasting, weakness and changes in muscle tone. Amyotrophic lateral sclerosis (ALS) is the most common form of Motor Neurone Disease and in the UK; the terms are often used synonymously.

Although ICD9 and ICD 10 codes exist for the subtypes of MND, in the UK only the parent codes for MND are used in routine clinical practice. This means that the subtypes of MND cannot be identified from the ICD codes. Internationally a wider range of ICD codes are used and UK data from general practice (Read Code data) allows for sub-typing.

Below is a description of the algorithm to identify participants with MND codes in the UK Biobank population.

A full list of the ICD and Biobank self-report codes used can be found in Table 1 at the end of this document.

The estimated accuracy of the algorithm is included in Appendix 1.

The use of self-report codes is discussed in Appendix 2.

A. MOTOR NEURONE DISEASE

(1) Motor Neurone Disease prior to baseline assessment ('prevalent Motor Neurone Disease')

(a) Motor Neurone Disease detected by hospital admission EHR (with or without self-report) : One (or more) of the ICD (9 or 10) codes listed in Table 1, in HES APC, SMR01 or PEDW linked records in the primary or any secondary position where either

- The first ICD code date is prior to the date of baseline assessment.

OR

- The participant has self-reported the condition at the baseline assessment, but the first ICD code date is post the date of baseline assessment.

(b) Motor Neurone Disease by self-report only: The participant has self-reported Motor Neurone Disease at baseline assessment, but without evidence of Motor Neurone Disease from linked HES APC, SMR01 or PEDW data (as defined above).

Setting the date of prevalent Motor Neurone Disease Diagnosis:

- If a participant has both an ICD code and a self-report code, the earliest recorded date regardless of source is used.
- If a participant has both an ICD code and a self-report code, but the self-reported date is missing, the ICD code date is used unless it is post the date of baseline assessment, in which case the default missing date is used.
- If the participant has ICD code(s) only, the earliest ICD code date is used.
- If the participant has self-report code(s) only, the earliest self-reported date is used.
- Missing dates are set to 1/1/1900.

(2) Motor Neurone Disease following baseline assessment ('incident Motor Neurone Disease')

Excluding those with Motor Neurone Disease detected prior to baseline assessment:

(a) Motor Neurone Disease detected by hospital admission EHR: One (or more) of the ICD (9 or 10) codes in HES APC, SMR01 or PEDW linked records, in the primary or any secondary position, with code date post the date of baseline assessment.

(b) Motor Neurone Disease detected by death register only: No ICD codes in HES APC, SMR01 or PEDW linked records, but one (or more) ICD codes in death register records, in the underlying cause or any other position.

Setting the date of incident Motor Neurone Disease Diagnosis:

- If a participant has ICD codes in both hospital admission and death register records, the earliest recorded code date regardless of source is used.
- If ICD code(s) recorded in hospital admission only, the earliest ICD code date is used.
- If ICD code(s) recorded in death register only, the date of death is used.

Table 1. Code Lists for Motor Neurone Disease

UK Biobank Self Report Codes			
Code Type	Code	Biobank Code Text	Motor Neurone Disease
UK Biobank Self Report	Field 20002 Code 1259	Motor Neurone Disease	ü
ICD 9 Codes			
Code Type	ICD 9 Code	ICD 9 Text	Motor Neurone Disease
ICD 9 Code	335.2	Motor Neurone Disease	ü
ICD 10 Codes			
Code Type	ICD 10 Code	ICD 10 Text	Motor Neurone Disease
ICD 10 Code	G12.2	Motor Neurone Disease	ü

Appendix 1

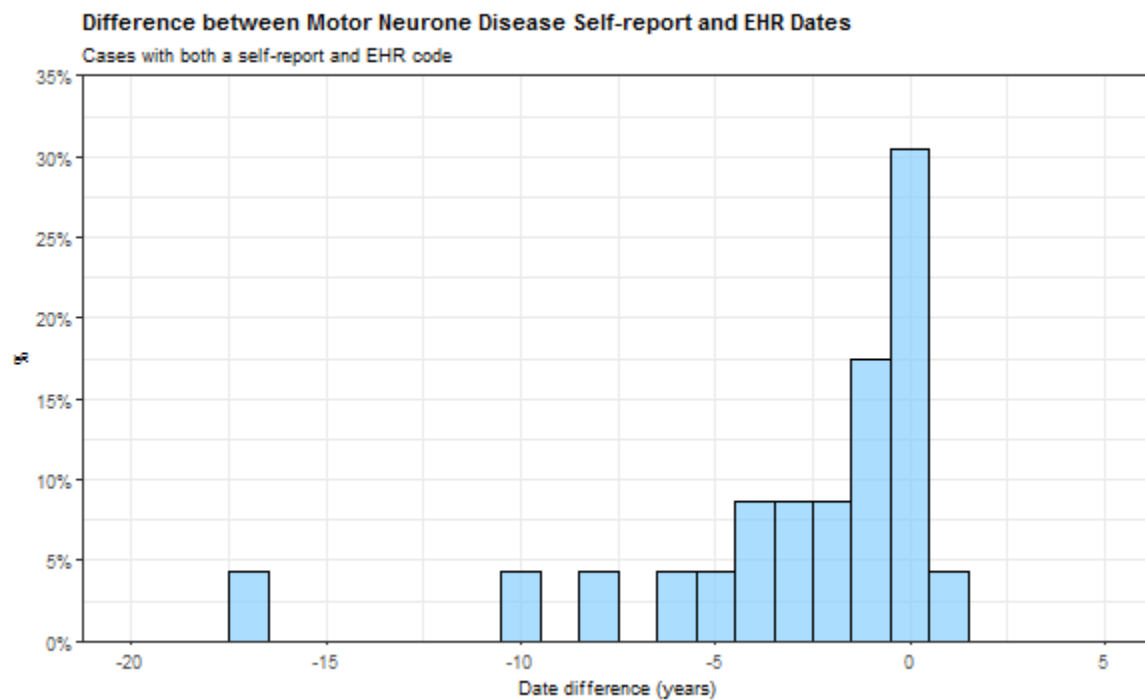
A 2017 systematic review¹ looked at the positive predictive values (PPV) and sensitivity of using routinely collected healthcare data to identify Motor Neurone Disease. This review included 13 studies from around the world. Estimates of PPV ranged from 55-92%. The only UK based study of primary care data reported a PPV of 85%.

A validation study of the accuracy of MND codes in UK Biobank participants is planned and should provide further information with regards to the estimated accuracy of MND codes in UK Biobank.

Appendix 2

The self-report date is taken from the UK Biobank field [20008](#) ("Interpolated Year when non-cancer illness first diagnosed"). At the nurse led interviews, nurses were instructed to record either a year or an age at which the diagnosis occurred. Where an age was provided, a best-fit fractional year was then calculated.

For cases that have both a self-report and EHR code, this algorithm assigns the earliest of the two code dates as the event date for the case. The histogram below shows the difference (in years) between self-report and EHR dates for the subset of Motor Neurone Disease cases that have both. Negative values indicate that the self-report date is earlier than the EHR. In the majority of cases (78%), the earliest date is the self-reported date.



Acknowledgements:

We would like to acknowledge the contributions of the 'UK Biobank Follow-up and Outcomes Working Group', whose work provided the foundations of this document:

Chair: John Danesh, Cambridge University,
Naomi Allen, UK Biobank, Oxford University,
Mark Atkinson, Swansea University,
Ekaterini Blaveri, Cancer Research UK,
Rachael Brannan, National Cancer Intelligence Network,
Carol Brayne, Cambridge University,
Sinead Brophy, Swansea University,
Nish Chaturvedi, University College London,
Rory Collins, UK Biobank, Oxford University,
Simon deLusignan, Surrey University,
Spiros Denaxas, University College London,
Parul Desai, Moorfields Eye Hospital,
Sophie Eastwood, University College London,
John Gallacher, Cardiff University,
Harry Hemingway, University College London,
Matthew Hotopf, Kings College London,
Martin Landray, Oxford University,
Ronan Lyons, Swansea University,
Mark McGilchrist, Dundee University,
Henrik Moller, Kings College London,
Terence O'Neil, Manchester University,
Mike Pringle, Nottingham University,
Tim Sprosen, Oxford University,
David Strachan, St George's University, London,
Cathie Sudlow, UK Biobank, Edinburgh University,
Frank Sullivan, Dundee University,
Rebecca Woodfield, Edinburgh University,
Qiuli Zhang, UK Biobank, Edinburgh University,
Secretariat: Robin Flaig, UK Biobank Edinburgh University.

References:

1. Horrocks S, Wilkinson T, Schnier C, et al. Accuracy of routinely-collected healthcare data for identifying motor neurone disease cases: A systematic review. Le W, ed. *PLoS ONE*. 2017;12(2):e0172639. doi:10.1371/journal.pone.0172639.