



**Definitions of Chronic Obstructive Pulmonary Disease for
UK Biobank Phase 1 Outcomes Adjudication**

**Date: March 2018
Version: 1.0**

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

Definitions of Chronic Obstructive Pulmonary Disease, UK Biobank Phase 1 Outcomes Adjudication

Algorithms allow a participant to have more than one pathological type.

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

Definitions & Abbreviations:

COPD	Chronic Obstructive Pulmonary 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 (SMR 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:

Chronic Obstructive Pulmonary Disease (COPD) is described by the World Health Organization (WHO) as:

'...a lung disease characterized by chronic obstruction of lung airflow that interferes with normal breathing and is not fully reversible.'

Previous medical terms used to describe this condition include “emphysema” and “chronic bronchitis”.

COPD is predominantly caused by smoking; however, other factors, including occupational exposures and genetic factors, may also contribute to the development of COPD².

Around 1.2 million people in the UK have a diagnosis of COPD³ and every year there are 140,000-hospital admission with the condition, making it the second largest causes of emergency admissions to hospital³. 30,000 people die from COPD in the UK every year³.

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 use of self-report code dates is discussed in Appendix 1.

A. CHRONIC OBSTRUCTIVE PULMONARY DISEASE

(1) COPD prior to baseline assessment ('prevalent COPD')

(a) COPD 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 after the date of baseline assessment.

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

Setting the date of prevalent COPD 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) COPD following baseline assessment ('incident COPD')

Excluding those with COPD detected prior to baseline assessment:

(a) COPD 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) COPD 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 COPD 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.

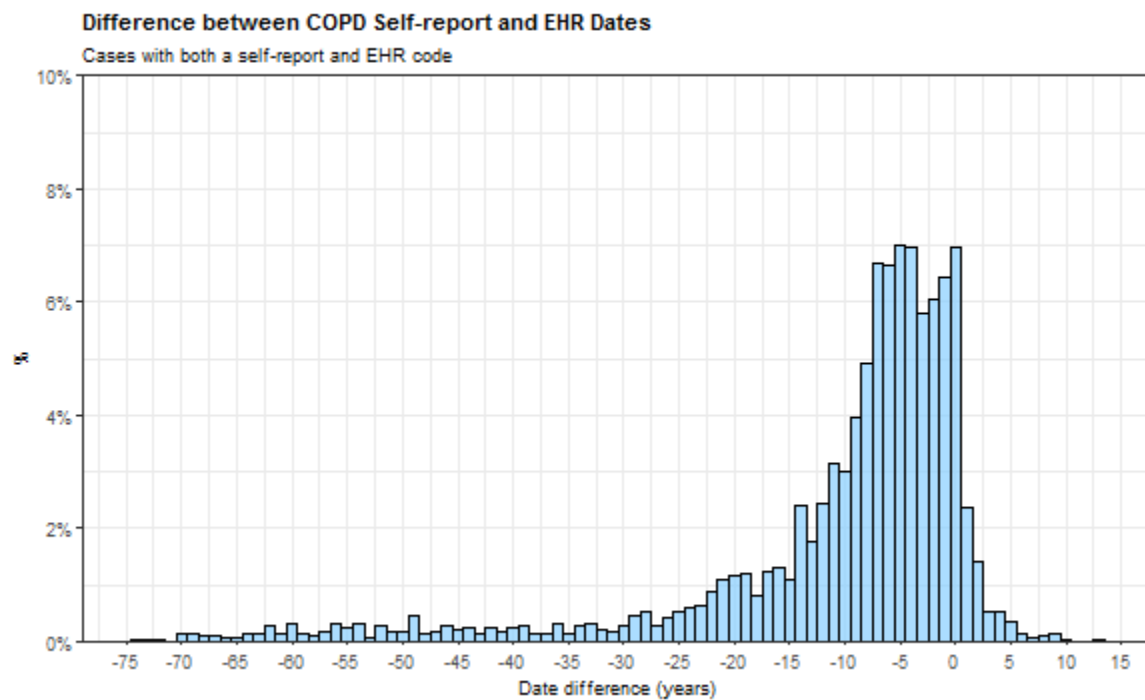
Table 1. Code Lists for COPD

UK Biobank Self Report Codes			
Code Type	Code	Biobank Code Text	COPD
UK Biobank Self Report	Field 20002 Code 1112	Chronic obstructive airways disease/COPD	ü
UK Biobank Self Report	Field 20002 Code 1113	Emphysema/chronic bronchitis	ü
UK Biobank Self Report	Field 20002 Code 1472	Emphysema	ü
ICD 9 Codes			
Code Type	ICD 9 Code	ICD 9 Text	COPD
ICD 9 Code	492	Emphysema	ü
ICD 9 Code	492.0	Emphysematous bleb	ü
ICD 9 Code	492.8	Other emphysema	ü
ICD 9 Code	496.X	Chronic airway obstruction, not elsewhere classified	ü
ICD 10 Codes			
Code Type	ICD 10 Code	ICD 10 Text	COPD
ICD 10 Code	J43	Emphysema	ü
ICD 10 Code	J43.0	MacLeod syndrome	ü
ICD 10 Code	J43.1	Panlobular emphysema	ü
ICD 10 Code	J43.2	Centrilobular emphysema	ü
ICD 10 Code	J43.8	Other emphysema	ü
ICD 10 Code	J43.9	Emphysema, unspecified	ü
ICD 10 Code	J44	Other chronic obstructive pulmonary disease	ü
ICD 10 Code	J44.0	Chronic obstructive pulmonary disease with acute lower respiratory infection	ü
ICD 10 Code	J44.1	Chronic obstructive pulmonary disease with acute exacerbation, unspecified	ü
ICD 10 Code	J44.8	Other specified chronic obstructive pulmonary disease	ü
ICD 10 Code	J44.9	Chronic obstructive pulmonary disease, unspecified	ü

Appendix 1

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 COPD cases that have both. Negative values indicate that the self-report date is earlier than the EHR. In the vast majority of cases (91%), 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. World Health Organization. *COPD Definition*. [Online] <http://www.who.int/respiratory/copd/definition/en/>. [Accessed 16 July 2018].
2. National Institute for Health and Care Excellence. *Clinical Knowledge Summaries: Chronic Obstructive Pulmonary Disease*. [Online]. <https://cks.nice.org.uk/chronic-obstructive-pulmonary-disease#!backgroundsub:1>. [Accessed 16 July 2018].
3. Snell N, Strachan D, Hubbard R, et al. *S32 Epidemiology of chronic obstructive pulmonary disease (COPD) in the UK: findings from the British lung foundation's 'respiratory health of the nation' project*. *Thorax* 2016;71:A20.