

Central Linkage Processes: Access to Primary Care Records

Dr Mark Effingham

Chief Operating Officer

UK Biobank

mark.effingham@ukbiobank.ac.uk

Central Linkage Processes: Access to Primary Care Records

This session will provide a brief update on two topics:

1. Completing linkage to primary care records for the cohort
What have we achieved and what is left to complete?
2. Making primary care data available as part of the resource
When will primary care records be available to researchers?

Primary care linkage is critically important

Substantial numbers of many diseases of interest are ascertained only through linkages to primary care data, e.g. dementia, COPD

Condition	Incident cases observed by 2016		Incident cases predicted by 2026	
Dementia	1,800	4,300	18,000	43,400
Stroke	4,600	7,100	18,400	28,500
MI	7,400	8,000	20,500	22,300
COPD	7,600	17,600	23,800	54,700
Parkinson's Disease	1,000	2,000	4,700	9,700

Numbers in BLACK show cases ascertained from linkage to hospital admissions and mortality data

Numbers in RED show cases ascertained from also including linkages to primary care data

Challenges with primary care linkage

Getting access to primary care data has been more difficult...

1. All participants are registered with a GP in the NHS
Each participant has given their consent to linkage
2. But, there is no single source of access
No national authority exists with access to primary care data
3. Our approach has been to work with primary care system providers – *TPP, Vision (via Apollo) and EMIS*

What primary care linkage have we achieved?

Wales (~95%) [*first linked in 2012*]

- Data obtained for 21,000 participants from EMIS and Vision via SAIL

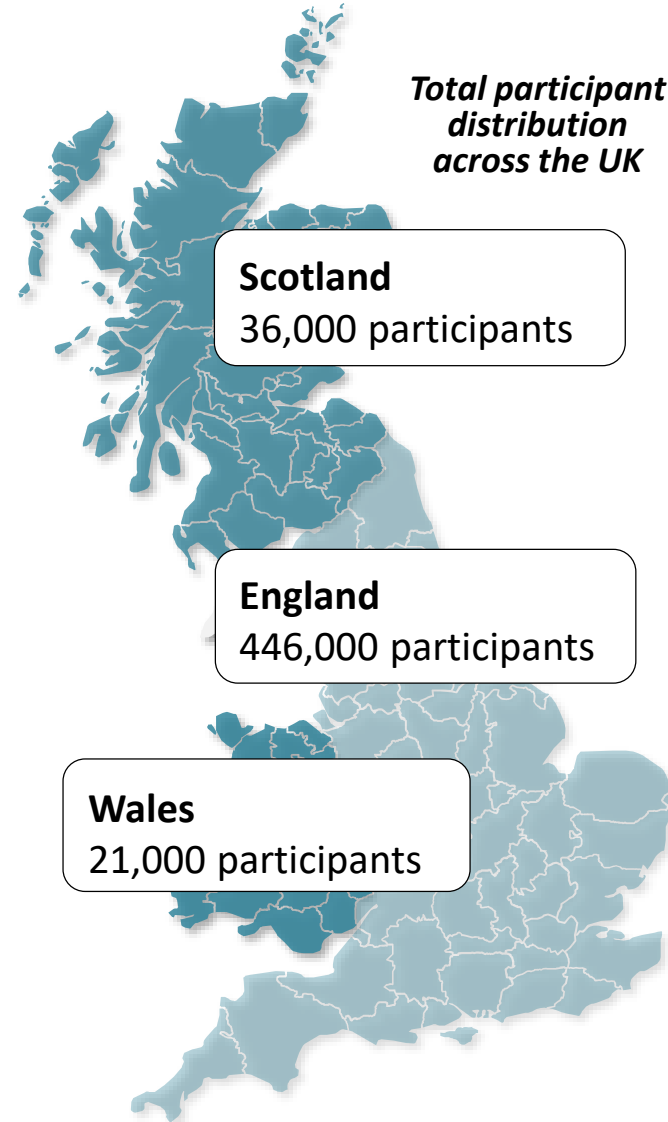
Scotland (~75%) [*first linked in 2013*]

- Data obtained for 27,000 participants from EMIS and Vision via Albasoft

England (~40%)

- Data obtained for 167,000 participants from TPP [*first linked in 2016*]
- Data obtained for 19,000 participants from Vision via Apollo [*first linked in 2017*]

Total participant distribution across the UK



Scotland
36,000 participants

England
446,000 participants

Wales
21,000 participants

Current linkage covers about half of all participants

Why is it taking so long?

- With support of the GP system suppliers, linkage across England worked very well (proceeding on an ‘opt-out’ basis)
 - Only 1 practice (of around 2,600) chose not to participate
- The ‘opt-out’ process was endorsed by the ICO and RCGP, however concerns from the RCGP/BMA Joint IT Committee led to a pause in further extracts (and any repeats)
- UK Biobank has continued to engage extensively with:
 - the ICO, NHS England and NHS Digital; and
 - the RCGP and its RCGP/BMA Joint IT Committee
- A revised ‘opt-in’ approach has now been agreed; success will depend on getting attention from (busy) GPs

Statements of support from (amongst others):

“The College has supported this initiative from the outset. It has the potential to change the way in which we diagnose and treat serious illnesses – and, crucially, prevent them happening in the first place. General practice is under unprecedented pressures and, as a practising GP myself, I fully understand what life on the frontline is really like – but UK Biobank is a trusted and bona fide resource, and I would encourage GP practices to participate wherever possible.”

Professor Helen Stokes-Lampard
Chair of the Royal College of GPs

Completing linkage for the UK Biobank cohort

- The revised approach has considered how best to reduce any administrative burden placed on GPs
- It has explicit support from the RCGP/BMA Joint IT Committee
- We are now writing to each general practice in England to seek their opt-in to provide a data extract for consented participants
- Communication to practices:
 - Pilot mailing to 600 practices by end June 2019
 - Main mailing (to ~6,000 practices) in September 2019
- We hope practices will agree to primary care data being provided to UK Biobank for an extract to occur in Q4 2019 (and then annually thereafter)

When will primary care data be made available?

- An interim release of primary care records will be available in early Q3 (covering ~234,000 participants)
 - > 120 million event records
 - > 50 million prescribed items

Country	Data Provider	Number of participant records	Coding System
Wales	SAIL	21,000	Read v2
Scotland	Albasoft	27,000	Read v2, BNF
England	TPP	167,000	CTV3 (Read v3)
	Vision	19,000	Read v2, DM+D
	EMIS	-	Read v2, DM+D

Total 234,000

- Will contain coded data only e.g.
 - clinical events
 - prescriptions
 - lab test results
 - immunisation records
- Data will be provided in their native coding system
- Only limited data curation applied

How will primary care data be made available?

- This will be an interim release prior to completing linkage across the cohort; subsequent releases are likely during 2020
- Record level data will be available via the online Data Portal (similar to the existing Hospital Episodes Statistics data)
- Our intent is to provide algorithmically derived outcome fields for a large number of diseases (>1,000)
- Timelines:
 - record-level data available early Q3 2019; and
 - derived outcomes either coincident with the interim release, or shortly thereafter

Thank you!

Funding bodies:



Improving the health of future generations