

Adjudication of Cancer Outcomes in UK Biobank

R Brannan^[1], M Chapman^[1], N Cooper^[1], J Thomas^[1], H Møller^[2]

1. National Cancer Intelligence Network 2. King's College London

UK Biobank is a long-term study of the impact of lifestyle, environment and genetics in adults aged 40-69. Achieving the full value of this resource will require UK Biobank to accurately classify clinical outcomes among their 500,000 participants over their lifetimes.

- UK Biobank aim to identify disease events by cross-referencing reported outcomes across multiple sources of information, starting with health outcomes ascertained through lower-cost electronic sources and proceeding to more resource-intensive methods for validation and classification.
- In the UK, the validation of cancer outcomes benefits from well developed conventions for defining occurrence and classification compared with many other non-communicable diseases.
- The NCIN and UK Biobank pilot project on the Adjudication of Neoplastic Outcomes has set out to establish mechanisms for identifying and phenotyping cancer outcomes amongst UK Biobank participants.

PILOT OBJECTIVES

NCIN and UK Biobank set out to:

1. Demonstrate linkage of a subset of the UK Biobank cohort to the National Cancer Data Repository for England and assess the quality of this link.
2. Identify and phenotype prevalent and incident cancer cases (using date of entry into the cohort) and to classify these cases by the certainty of diagnosis.
3. To work with the cancer research community to define an initial dataset for the sub-classification of cancer cases in UK Biobank
4. Identify the hospital trusts treating cancer patients in the selected cohort and work with the UK Biobank co-coordinating centre to help assess the feasibility of obtaining additional clinical data, diagnostic samples and imaging information.



RESULTS

LINKAGE: The NCIN have successfully linked the NCDR (1990-2010) to UK Biobank participants' identifiers. 99.9% of the cohort had a valid NHS number.

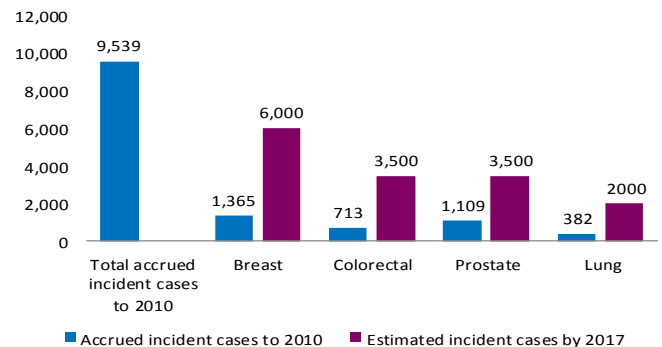
CANCER CASES: Within the NCDR, 46,370 tumours were identified in 42,830 unique patient records.

INCIDENT CASES: 9,539 incident cases have been accrued since entry into the cohort, this includes around 1,400 breast, 700 colorectal, 400 lung and 1100 prostate incident cases.

ACCESS TO DIAGNOSTIC TISSUE: 83% of incident cases have a linked HES record. Analysis using 'treating hospital' as a proxy to identify where diagnostic tissue may be held on UK Biobank participants has identified pathology labs to approach for access to cancer tissue.

MINIMUM CANCER DATASET: Minimum cancer dataset has been identified to enable UK Biobank users to phenotype cancer cases. UK Biobank users will be able to access this data by the end of 2013.

Accrued incident cases in UK Biobank since (Date of entry into cohort to Dec 2010)



MINIMUM CANCER DATASET

- Primary site
- Morphology
- Date of Diagnosis
- Basis of Diagnosis
- Grade
- Laterality
- TNM stage
- Treatment
- Duke stage of (C18-20)
- FIGO Stage of (C51-57)
- Clark level of (C43, C44)
- NPI score Of C50
- Breslow of (C43, C44)
- Gleason grade of C61

METHOD

All England resident participants in UK Biobank (n=441,850) were matched against the NCDR (1990-2010) using NHS number, date of birth, postcode and sex.

All Incident and prevalent cases in the UK Biobank-NCDR linked dataset were identified using date of entry into the cohort (date of assessment centre visit) as Day 0 for the analysis. These cases have been further broken down by site and year. Through the Cancer Outcomes Working Group and a workshop held in July 2012, a minimum cancer dataset for UK Biobank linkage has been defined.

CONCLUSIONS AND NEXT STEPS

Linkage to the NCDR demonstrates that retrospective and prospective cancer case ascertainment is readily achieved for UK Biobank. It is intended that the UK Biobank will be enhanced using the proposed minimum cancer dataset once the new iteration of the NCDR is available (1990-2011). Work surrounding access to diagnostic archives held by NHS pathology departments is on-going and will be reported in 2014.

REFERENCES: Allen N, Sudlow C, Downey P, Peakman TC, Danesh J, Elliott P, Gallacher J, Green J, Matthews P, Pell J. (2012) UK Biobank: Current status and what it means for epidemiology, Health Policy and Technology,1(3):123-126

ACKNOWLEDGEMENTS: The NCIN would like to thank Dr Cathie Sudlow, Dr Tim Peakman, Dr Naomi Allen and colleagues at UK Biobank.