

## UK Biobank & Access to Primary Care Data: March 2020

### Background

UK Biobank is a major health resource with the aim of improving the prevention, diagnosis and treatment of a wide range of serious and life-threatening illnesses. As a prospective cohort study, UK Biobank provides a powerful resource to help scientists discover why some people develop particular diseases and others do not. An important aspect of this relates to being able to follow the health of the UK Biobank participants long-term.

Between 2006 and 2010, 500,000 people in the UK who were aged between 40-69 years agreed to join UK Biobank. Each of them underwent physical measures, gave blood, urine and saliva samples, and provided detailed information about themselves. All 500,000 UK Biobank participants explicitly consented to have their health followed through linkage to their medical and other health records (including primary and secondary care) for any health-related research that is in the public interest.

UK Biobank is an open access resource for use by bona fide researchers. UK Biobank data is only available to approved researchers for approved research and it is then provided in a de-identified form for that research under a Material Transfer Agreement with UK Biobank.

### Initial Approach to Primary Care Data Linkage (based on “opt-out”)

Of the 500,000 participants, 446,000 were recruited in England, 36,000 in Scotland and 21,000 in Wales. Since 2012, UK Biobank has routinely linked to death, cancer registry and secondary care data across England, Wales and Scotland. It has also been able to get access to primary care records in Wales (through the SAIL system) and in Scotland (through Albasoft) but not to date in England.

As primary care data are not held centrally in England (where the large majority [89%] of participants live), UK Biobank sought these data directly from the main GP system suppliers who hold them. The only data that were being sought were solely coded data about diagnoses, prescriptions and investigations (not any letters, notes or other written material).

In November 2015, UK Biobank entered into an agreement with TPP and after a long period of consultation adopted the following approach:

- a letter was sent from UK Biobank and the TPP Chief Medical Officer to each GP practice, endorsed by the Royal College of GPs, setting out the details of the prospective extraction;
- the letter set out the basis for and the extent of the prospective extraction and provided a guarantee that UK Biobank would only extract data on a GP practice patient who was also a participant;
- this letter provided the GP practice with the opportunity to ask any questions or seek further information (which questions were raised by a relatively small number of GP practices, but were satisfied by the answers provided); and
- the extraction then went ahead a number of weeks later. In the event that a practice objected – and only one practice chose to do so – data were not extracted from the practice.

Prior to this extraction proceeding, UK Biobank consulted extensively with the RCGP as it was considered (correctly) to be pivotal to obtain the RCGP's buy-in and endorsement. This resulted in the RCGP being a co-signatory of the letter that was dispatched to all the TPP GP practices.

As a result of the TPP extraction – which allowed UK Biobank to receive extracted data on ~170,000 consented UK Biobank participants – UK Biobank engaged with the other main system suppliers in England (INPS & EMIS) to reach agreement based on a similar approach to obtain data from their GP practices.

In November 2016, UK Biobank engaged with INPS (via Apollo as their data processor) to secure a similar agreement that led to a successful extraction of primary care data for ~30,000 UK Biobank participants whose records were being managed by INPS GP practices. There were no objections raised by any of the ~600 INPS GP practices contacted in relation to the prospective extraction.

Then in July 2017, UK Biobank sought to finalise a similar agreement with EMIS however there was a discussion with EMIS's National User Group (NUG) who raised two issues:

1. the EMIS NUG raised certain questions about the UK Biobank consent and these were addressed to the satisfaction of the NUG; and
2. the EMIS NUG suggested that (on data protection grounds) it would be desirable to get the individual GP practices to opt-in to the proposed extraction and that otherwise it would be difficult for them to recommend the de facto opt-out approach that UK Biobank had hitherto used with TPP and INPS.

In response to these concerns, UK Biobank confirmed that it would review the situation with the UK Information Commissioner's Office (ICO).

To this end, UK Biobank engaged with both the Commissioner herself and the ICO's legal Counsel. The upshot of these discussions was that the ICO issued a robust endorsement of UK Biobank's methodology and approach used with TPP and INPS. The ICO confirmed that using the language of opt-in and opt-out was not constructive (as this was a right of the data subject not the data controller) and stated specifically that the UK Biobank consent is compliant with the (then) existing law and the GDPR. The ICO did mention that, as the GP practices were data controllers, it would be necessary for there to be a clear instruction from them to the systems suppliers (the data processors) to release the data to UK Biobank. This instruction could be specific or general in nature (which follows the language of the GDPR).

The Joint GP IT Committee (of both the BMA and the RCGP) unilaterally intervened in late 2017 and expressed certain concerns about the process. The Committee insisted that the only way, in their opinion, that the data could be accessed would be with the proactive assent (a de facto opt-in) of all the relevant individual GP practices (some 6,000 or so) in England.

This effectively suspended all primary care data extraction from the system suppliers whilst an alternative approach could be worked out.

### **Revised Approach to Primary Care Data Linkage (based on "opt-in")**

Following further discussions with the RCGP/BMA Joint GP IT Committee and with NHS England and NHS Digital (who adopt a de facto opt-in approach for data requests from GP practices) UK Biobank considered that it should evaluate the viability of an opt-in process.

To this end, in October 2018, UK Biobank initiated a campaign to inform and obtain agreement from practices supported by EMIS and TPP. A letter was dispatched to ~2,600 TPP practices and ~4,000 EMIS practices, with a precursor letter sent to 242 CCGs and 70 Local Medical Committees to advise them in advance of the request.

After several months, only 172 EMIS practices (accounting for around 24,000 out of over 250,000 registered participants) had opted-in and activated their data sharing agreement. This rather low level of response may have been influenced by an internal (and unilateral) communication to all practices in November 2018 from one of the co-chairs of the RCGP/BMA Joint IT Committee. This communication highlighted additional concerns (not previously raised with UK Biobank) that the

proposed process made it difficult for GP practices to meet their data protection obligations and advised them not to agree to the extraction.

UK Biobank then had further consultation with the Joint GP IT Committee and developed a modified process to address these additional concerns, and this was agreed in May 2019. The Committee's agreement was contingent on, inter alia, GP practices being advised to conduct a data privacy impact assessment prior to providing approval.

### **Pilot Mailing to 600 GP Practices using this Enhanced Opt-in Approach**

A pilot to test the modified process commenced on 31 July 2019, with a physical mailing to 600 practices (300 EMIS, 300 TPP) comprising of:

- a letter to the Senior Partner / Practice Manager with information about the process;
- a list of each practice's participants;
- a partially completed data privacy impact assessment (DPIA) for adoption by the practice;
- instructions on how to review participant consent forms online; and
- instructions on how to audit the data extracted.

The letter incorporated quotes of support from the CEO of NHS Digital, the Chief Medical Officer for England, the Chair of the Royal College of GPs, and the Clinical Chair of the NHS Assembly. The pilot was also designed to investigate whether offering a payment to GP practices would increase approval rates, by randomising practices to receive the letter with or without an offer of payment.

Assenting to the data release required the practice to tick an "I agree" box within the data sharing part of their practice management system. The extraction was undertaken directly by the GP system supplier with no further action required by the practice. The response to this campaign remained at the same disappointing level of below 20%. Specifically, as of the end of October 2019 (3 months after the pilot mailing date):

EMIS: 44 GP practices (15%) of practices mailed opted-in, taking the total number of EMIS opt-ins to 233 GP practices – 44 (out of 300) practices targeted within the pilot mailing and 189 arising from the earlier national mailing and EMIS follow-up.

TPP: 47 GP practices (16%) of practices mailed opted-in, taking the total number of TPP opt-ins to 71 GP practices – 47 (out of 300) practices targeted within the pilot mailing and 24 arising from the earlier national mailing and TPP follow-up.

### **Revised Pilot Mailing to a Further 100 GP Practices**

As part of the pilot mailing to 600 practices, several GP practices raised their own concerns with the following consistent themes:

- the need to undertake a DPIA that was time-consuming and appropriate skills are unavailable;
- practices were frustrated about being expected to do this work without reimbursement of their time (though this feedback was also received from practices to whom an offer of reimbursement had been made within the randomised comparison);

- in several cases, practices had defined local data protection policies (as a result of GDPR) that require consent forms to have been signed within the past 6 months (notwithstanding ICO written confirmation that our consent forms remain GDPR-compliant); and
- practices would prefer that a copy of the consent forms be included within the mailing as practices found it complicated to access them through UK Biobank's online system (which is preferable methodology as it avoids sending personal data to the practices).

In January 2020, following the outcome of the pilots and this feedback, UK Biobank developed a revised, streamlined request (in other words as direct and succinct as possible) which was dispatched to a further ~100 GP practices (EMIS – 49; TPP – 48). The revised pack contained:

- a refined letter to the Senior Partner / Practice Manager;
- a more prominent offer to reimburse a practice for their time to review the request;
- further highlighting that the ICO regards our consent forms to be fully GDPR-compliant;
- revised advice about the need for a practice to perform a Data Privacy Impact Assessment;
- a list of each practice's participants; and
- a USB Flash Drive containing a secure, digital copy of consent forms.

The response to this campaign has also remained (similarly) disappointing to-date. As of the beginning of March 2020 (a month after the pilot mailing date):

EMIS: 4 GP (out of 49) practices (8%) of practices mailed have opted-in.

TPP: 3 GP (out of 48) practices (6%) of practices have mailed opted-in.

There have been no requests for reimbursement received. Although the results of the revised pilot are currently preliminary, they so far indicate that seeking the direct engagement of GP practices will not work. The request to GP practices is a request which looks (to them) to be risky and onerous: as such, far easier (and quite understandably) for a busy GP practice to pass.

**CONCLUSION:** Despite explicit consent from all 500,000 UK Biobank participants which the Information Commissioner has explicitly confirmed is valid for access to all of their medical and health-related records, UK Biobank has not been able to obtain access to primary care records for the vast majority of them (i.e. those living in England) using the process proposed by the Joint RCGP/BMA IT Committee.

For those participants for whom we do have primary care data, we know these data help to identify many more cases of all health outcomes (particularly those that are chiefly managed in primary care) and provide much more information about participants (e.g. prescription and diagnosis codes) which are of substantial value to researchers by increasing the sensitivity and specificity of their analyses.