**Data Transfer Agreement: THIRD PARTY DATA PROVISION TO COG-UK DATABASE**

|  |  |  |
| --- | --- | --- |
|  | **Recipient** | **Covid-19 Genomics UK consortium** (“**COG-UK**”)  acting through The Chancellor, Masters and Scholars of the University of Cambridge of The Old Schools, Trinity Lane, Cambridge CB2 1TN, UK (“**Cambridge**”) |
|  | **Provider** | [INSERT PROVIDER LEGAL ENTITY] of [INSERT ADDRESS] |
|  | **Data** | means the data items described in Table M1(Redacted) set out in Annex A as updated and revised from time to time, that are provided by the Provider to the Recipient. |
|  | **A** | COG-UK is a Covid-19 Genomics UK consortium funded by the MRC, NIHR and Wellcome Sanger Institute. As named recipient of the funding, Cambridge is the lead party of the consortium which includes the Wellcome Sanger Institute, the public health agencies for England, Northern Ireland, Scotland and Wales and numerous universities. |
|  | **B** | COG-UK was established to create a Severe Acute Respiratory Syndrome Virus (“**SARS-CoV-2**”) sequencing network across the UK and sequence circa 200,000 samples. It works in conjunction with and to support the public health agencies in public health monitoring and reporting. COG-UK conducts analytical research to address public health questions related to SARS-CoV-2 through developing a better understanding of the genetic makeup of SARS-CoV-2 using the sequence related data from its network together with epidemiological and clinical information. |
|  | **C** | Data including sequence related data and epidemiological and clinical information is provided to COG-UK by consortium members and third parties (the “**COG-UK Data**”). It is compiled and arranged in a database (the “**COG-UK Database**”) which is currently hosted by part of the Cloud Infrastructure for Microbial Bioinformatics (“**CLIMB-COVID**”) funded by the Medical Research Council and managed by the University of Birmingham and Cardiff University (the “**CLIMB Team**”). |
|  | **D** | The Provider is willing to provide, and COG-UK is willing to receive, the Data for incorporation into the COG-UK Database and use for the above described purposes from the Effective Date. |

In consideration of the payment by the Recipient to the Provider of the sum of one pound (£1), the receipt and sufficiency of which is acknowledged by the Provider, the Provider agrees to provide and the Recipient agrees to receive and use the Data in accordance with the terms and conditions set out in this Agreement.

AGREED by the Parties through their authorised signatories:-

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| For and on behalf of  the **COG-UK consortium acting through**  **The Chancellors, Masters and Scholars of the University of Cambridge** | |  | | For and on behalf of [INSERT PROVIDER] | |  |
| Signed | |  | | Signed |  | |
|  | |  | |  |  | |
| Print name | |  | | Print name |  | |
|  | |  | |  |  | |
| Title | |  | | Title |  | |
|  | |  | |  |  | |
| Date | |  | | Date |  | |

# Terms and Conditions

1. **Definitions**
   1. **Data Protection Legislation:** means the Data Protection Act 2018 as amended from time to time and any successor, subsidiary or accompanying legislation in the UK and (for so long as and to the extent that the law of the European Union has legal effect in the UK) the General Data Protection Regulation (EU) 2016/679 and any other directly applicable European Union regulation relating to data protection and privacy.
   2. **Effective Date**: means [1 April 2020].
   3. **Permitted Purposes**: means research into Covid-19 (including SARS-CoV-2 sequencing and analytical research aligned with the Research Themes) and use for public health monitoring purposes.
   4. **Personal Data:** will have the meaning given in the Data Protection Legislation.
   5. **PHA** or **PHAs:** means any or all of (i) Public Health England, (ii) the Regional Agency for Public Health and Social Well-being (i.e. Public Health Northern Ireland), (iii) Public Health Scotland and (iv) Public Health Wales.
   6. **Process:** will have the meaning given in the Data Protection Legislation.
   7. **Research Themes:** means the public health questions set out in Annex B.
   8. **Results:** means all information, techniques, data, results, design, technology, materials, inventions, algorithms and software identified or first reduced to practice or writing in the course of undertaking the Permitted Purposes.
   9. **Steering Group**: means the COG-UK steering group; which can be contacted by email at [eg618@medschl.cam.ac.uk](mailto:eg618@medschl.cam.ac.uk).

In this Agreement any words following the term “including” or any similar expression will be construed as illustrative and will not limit the sense of the words, description, definition, phrase or term preceding that term.

1. **Provision & Use of Data**
2. The Provider will provide Data, duly anonymised, to the Recipient from time to time for incorporation into the COG-UK Database and use for the Permitted Purposes. The Provider agrees and acknowledges that when Data is incorporated into the COG-UK Database (currently hosted on CLIMB-COVID) it cannot subsequently be removed or deleted.
3. Nothing in this Agreement will affect ownership of the Data. However, the COG-UK Database (i.e. the compilation and arrangement of COG-UK Data but not its individual parts) is jointly owned by Cambridge, the PHAs and the Wellcome Sanger Institute.
4. The Provider confirms that it has complied and will at all times comply with all legal obligations relating to its provision of Data to the Recipient for use for the Permitted Purposes as described in this Agreement.
5. Without prejudice to clause 2.3, the Provider agrees and acknowledges that, where it collects or arranges for the collection of the samples or data in order to generate the Data provided to the Recipient, it will ensure that the collection is authorised under applicable laws on public health monitoring (or under other lawful basis) and under any necessary ethics approval.
6. The Provider grants the Recipient a perpetual, irrevocable, non-exclusive, royalty free licence to incorporate the Data into the COG-UK Database and use the Data (and grant sub-licences to use the Data) for the Permitted Purposes.
7. The Provider agrees and acknowledges that pursuant to the licence in clause 2.5 the Recipient may without limitation grant sub-licences to:
   * 1. COG-UK consortium members and third parties to access and use the Data on CLIMB-COVID for research projects aligned with the Research Themes, subject to approval of the Steering Group and relevant PHAs; and
     2. PHAs to download and use the Data for their own public health monitoring and reporting purposes.
8. Before providing any Data to the Recipient (including by uploading Data to CLIMB-COVID itself) the Provider will:
   * 1. check the Data for quality control and quality assurance purposes; and
     2. ensure that no single item of Data contains any direct identifiers (i.e. any value which directly identifies an individual).

2.8 The Recipient undertakes:

2.8.1 to use and permit the use of the Data solely for the Permitted Purposes in accordance with this Agreement and all applicable laws;

2.8.2 subject to clause 3, to keep the Data confidential, except for data items described in Annex A with a privacy level ‘Public’;

2.8.3 keep the Data secure by implementing organisational and technological measures appropriate to the nature and sensitivity of the Data to protect against the unauthorised or accidental access, use or disclosure of the Data;

2.8.4 subject to clause 2.9, not to attempt to identify any individual from the Data or to communicate with any individual identified from the Data, or to link or attempt to link the Data to other data or information if doing so might create Personal Data.

2.9 The Provider agrees and acknowledges that clause 2.8.4 shall not apply in respect of any download of the Data carried out by any PHA pursuant to clauses 2.5 and 2.6.2.

2.10 The Provider acknowledges that activity on CLIMB-COVID is monitored and information about the source of data input into CLIMB-COVID may be shared with members of the COG-UK consortium and their nominated representatives.

1. **Confidentiality & Publication**
   1. The Recipient will not be in breach of any obligation to keep the Data confidential if and to the extent that it:
2. is used or disclosed as expressly permitted by this Agreement;
3. is or becomes publicly known without any breach of this Agreement or other undertaking to keep it confidential; or
4. is required to be disclosed by law or order of court or competent authority, provided in all cases the Provider is notified as early as possible prior to disclosure.
   1. The Provider will not have or assert any right or interest in the Results. The Recipient is free to publish and permit the publication of Results arising from use of the Data provided that the Data is not disclosed in any such publication.
5. **Termination & Survival**
   1. This Agreement takes effect on the Effective Date and will terminate on expiry of twenty-eight (28) days written notice from one Party to the other.
   2. Rights granted to the Recipient under this Agreement will continue indefinitely and will be extended to any new member of the COG-UK consortium or third party as described in this Agreement.
   3. Any provision of this Agreement that expressly or by implication is intended to survive termination of this Agreement will remain in full force and effect, including the licence granted under clause 2.2.
6. **Liability**
   1. Except as provided in clauses 2.3 and 2.4, the Provider:
7. provides the Data “as is” and makes no representation and gives no warranty of any kind either express or implied in relation to the Data, including warranties of accuracy or fitness for a particular purpose, or that the Data will not infringe any patent, copyright, trademark or other proprietary rights, accordingly the Provider will not be liable for any loss arising from any reliance placed on the Data by the Recipient;
8. will not be liable to the Recipient for any use made of the Data by the Recipient, including any analysis, interpretations, conclusions or Results and any reports or publications of the Results.
   1. Nothing in this Agreement limits or excludes either Party’s liability for (a) death or personal injury resulting directly from negligence, (b) fraud or fraudulent misrepresentation, or (c) for any other liability which by law cannot be limited or excluded.
   2. The liability of either Party for any breach of this Agreement will not extend to loss of business or profit or to any indirect or consequential loss or damage.

1. **General**
   1. **Costs:** The Data is provided at no cost.
   2. **Entire Agreement and non-variation:** This Agreement constitutes the entire agreement and understanding of the Parties and supersedes all negotiations, understandings or previous agreement between the Parties relating to the subject matter of this Agreement. No addition to or variation, consensual cancellation, novation or assignment of this Agreement and no waiver of any right arising from this Agreement or its breach or termination shall be of any force or effect unless reduced to writing and signed by all the Parties or their duly authorised representatives.
   3. **Third Party Rights:** Except as otherwise expressly stated in this Agreement, no term of this Agreement is enforceable under the Contracts (Rights of Third Parties) Act 1999 by any person who is not a Party.
   4. **Dispute Resolution:** If any dispute arises out of or in connection with this Agreement the Parties will first attempt to resolve the matter informally through designated senior representatives of each Party to the dispute, who are not otherwise involved in this Agreement. A Party may litigate if the dispute has not been resolved informally within a reasonable time not exceeding two (2) months from the date the informal process is requested by notice in writing. Any Party may apply for an injunction, whether or not a dispute has been escalated under this clause.
   5. **Law:** This Agreement will be governed by the laws of England and Wales and the courts of England and Wales will (once the procedures set out in clause 8.4 above have been followed and exhausted) have exclusive jurisdiction to deal with any dispute which may arise out of or in connection with this Agreement.
   6. **Counterparts:** This Agreement may be executed in any number of counterparts each of which will be deemed an original, but all of which together will constitute one and the same instrument. This Agreement may be executed electronically using Docusign and delivered by email.
   7. **No agency:** Nothing in this Agreement shall create, imply or evidence any partnership or joint venture between the Parties or the relationship between them of principal and agent.
   8. **Waiver:** No failure or delay by a Party to exercise any right or remedy provided under this Agreement or by law shall constitute a waiver of that or any other right or remedy, nor shall it preclude or restrict the further exercise of that or any other right or remedy. No single or partial waiver of any right or remedy provided under this Agreement shall preclude or restrict the further exercise of that or any other right or remedy.
   9. **Severability**: If any one or more clauses or sub-clauses of this Agreement would result in this Agreement being prohibited pursuant to any applicable law then it or they shall be deemed to be omitted. The Parties shall uphold the remainder of this Agreement, and shall negotiate an amendment which, as far as legally feasible, maintains the economic balance between the Parties.

**Annex A – The Data**

**TABLE M1 (redacted)[[1]](#footnote-2)**

\* If any of the mandatory fields are missing the relevant sequence data will not be accepted into CLIMB-COVID.

\*\* Consortium Access Data = data items with privacy levels ‘Public’ or ‘Consortium’. COG-UK members can access and use this data for the Research without approval.

Restricted Access Data = data items with privacy level ‘Restricted’. COG-UK members need the approval of the Steering Group and relevant PHAs to access and use this data.

PHAs can access, download and use Consortium and Restricted Access Data relating to their geographical area for their own public health monitoring and reporting purposes without approval.

| **Data item** | **MRC-CLIMB field name** | **Description** | **Mandatory for submission to CLIMB-COVID?** \* | **Privacy level**  **(Public, Consortium or Restricted)** \*\* |
| --- | --- | --- | --- | --- |
| Central Sample ID | central\_sample\_id | COG-UK ID provided by Wellcome Sanger Institute | Yes | Public |
| COG-UK Patient ID | biosample\_source\_id | For sequencing labs: COG-UK Identifier for the first sample from this patient / environment  For PHA: unique but anonymous ID for this patient | No | Consortium |
| PHA sample ID | root\_sample\_id | Identifier assigned to this sample by a PHA or submitted to the PHA (if not identifiable) | No | Restricted |
| Local sample ID | sender\_sample\_id | Local ID assigned by submitting organisation (not identifiable) | No | Restricted |
| Date of sample  (collected) | collection\_date | Date sample collected | Either collection\_date or received\_date must be provided | Public |
| Date of sample (received) | received\_date | Date sample received by any lab | Public |
| Sampling strategy | is\_surveillance | Whether sample was chosen according to the surveillance sampling strategy | Yes | Consortium |
| Sample type collected | sample\_type\_collected | Type of sample at collection | No | Consortium |
| Swab site | swab\_site | Site of swab | No | Consortium |
| Sample type received | sample\_type\_received | Type of sample received by sequencing lab | No | Consortium |
| Ct value | diagnostic\_ct\_value / test\_kit / test\_platform / test\_target | Cycle threshold value and detail of measurement from diagnostic RT-PCR test | No | Consortium |
| UK nation | adm1 | UK nation | Yes | Public |
| County | adm2 | Patient’s county of residence | Yes | Consortium |
| Outer postcode | adm2\_private | Outer postcode of patient residence | No | Consortium |
| Household ID | TBC | Anonymous household ID generated by PHA from full address | No | Restricted |
| Age | source\_age | Patient age | No | Consortium |
| Sex | source\_sex | Patient sex | No | Consortium |
| Ethnicity | TBC | Ethnicity | No | Restricted |
| Socio-economic status | TBC | Area level deprivation quintile based on income domain of Indices of Multiple Deprivation | No | Restricted |
| Healthcare worker | is\_hcw | Is the sample from a health care worker? | No | Restricted |
| Employing hospital | employing\_hospital\_name | Name of the hospital if a health care worker who works in a hospital. | No | Restricted |
| Employing trust / board | employing\_hospital\_trust\_or\_board | Name of the trust / health board if a health care worker who works in a hospital. | No | Restricted |
| Care home worker | is\_care\_home\_worker | Is the sample from a care home worker? | No | Restricted |
| Care home resident | is\_care\_home\_resident | Is the sample from a care home resident? | No | Restricted |
| Care home ID | anonymised\_care\_home\_code | Locally assigned anonymous code (up to ten characters) that links samples from the same care home. | No | Restricted |
| Hospitalisation | is\_hospital\_patient | Is the sample from an admitted hospital patient? | No | Restricted |
| Date of admission | admission\_date | Date of admission to hospital | No | Restricted |
| Admitting hospital | admitted\_hospital\_name | Name of the hospital if a hospital patient. | No | Restricted |
| Admitting trust or board | admitted\_hospital\_trust\_or\_board | Name of the trust / health board if a hospital patient. | No | Restricted |
| Admitted with COVID-19 diagnosis | admitted\_with\_covid\_diagnosis | Was the patient admitted with a (suspected) diagnosis of COVID-19? | No | Restricted |
| Critical care admission | icu\_admission | Whether patient has been admitted to ICU | No | Restricted |
| Outcome | TBC | Outcome at 28 days | No | Restricted |
| Investigation name | investigation\_name | An optional field to name a local investigation. | No | Restricted |
| Investigation site | investigation\_site | An optional site identifier where multiple sites appear in the same investigation. | No | Restricted |
| Investigation cluster | investigation\_cluster | An optional cluster identifier for local sites to group samples believed to be in a local cluster. | No | Restricted |

**Annex B – Research Themes**

1. Enhance understanding of epidemiology and transmission
   1. ***Infer the relative contribution to incidence of local transmission versus imported cases.*** In the initial stages of the epidemic, as we currently are, the most immediate high-impact question that can be asked of SARS-CoV-2 whole genome sequencing is the relative contribution of importation and community-transmission to incidence.
   2. ***Define chains of transmission.*** Allowing for the caveats noted above about rates of mutation, viral genome sequences may identify individual chains of transmission at a local level by identifying and distinguishing nucleotide differences that are shared by a transmission chain. This can be used to ascertain if two sets of cases reported close to each other in space and time (e.g. in the same hospital, or the same postcode) represent the same chain of transmission and therefore have a single origin, or if they represent independent transmission chains that happen to have moved into the same location. Outbreaks may also be excluded.
   3. ***Estimate rates of epidemic growth and estimate of unreported cases and rate of sampling.*** Newer phylodynamic methods may be able to estimate the total number of infections in the epidemic and the rate through time at which they are being sampled and reported.
   4. ***Determine within-country spatial movement.*** We predict that widespread, representative, sampling and sequencing of SARS-CoV-2 will allow the reconstruction of general patterns of virus spread among locations and other population sub-groups.
   5. ***Provide insights into mechanisms and drivers of long-distance dispersal.*** These inferences can be made more powerful and precise by augmenting the virus genetic data with detailed information about patterns of national human mobility, obtained from geographic modelling.
2. Enable the monitoring of interventions and treatment
   1. ***Monitor the effect of non-pharmacological inventions on SARS-CoV-2 spread, epidemiology and biology.*** COG-UK will have a number of innovative clinical trials embedded within its capacity.
   2. ***Monitor shifts in antigenicity, and the emergence of drug resistance mutations*** in response to the introduction of new treatments and vaccines.
   3. ***Identify virus genetic markers associated with clinical severity***, noting that hypothesised phenotypes have to be ***interpreted*** with caution pending functional analysis of the role of mutations.
   4. ***Link viral genomics data with human genomics data*** (to be generated in parallel by Genomics England), human respiratory microbiome sequencing (to investigate role of secondary bacterial infections) and electronic health records to estimate multifactorial nature to on disease risk, severity and outcomes.
3. Expand biological understanding and further research
   1. Functional and ***phenotypic*** relevance of observed mutations – although many, perhaps most, hCoV-19 mutations will have no significant effect on virus biology, a subset of changes may alter some function of the virus. Circulating mutations can be identified and triaged using evolutionary and computational structural analysis in order to identify those that might have functional relevance.

1. Please check with Recipient that you have the most recent version of Table M1(Redacted). [↑](#footnote-ref-2)