MEETING SUMMARY

COVID-19 DATA PLATFORM
DATA ACCESS COMMITTEE INAUGURAL MEETINGS

Videoconference – 21 September 2020, 08:00-10:00 BST

Videoconference – 23 September 2020, 14:00-16:00 BST

MEETING ATTENDEES

DAC Members	Secretariat
Prof Yaseen Arabi [23SEP]	Laura Merson [21SEP/23SEP]
Dr Vivian Avelino-Silva [23SEP]	Matthew Brack [21SEP/23SEP]
Dr Abi Beane [23SEP]	Dr Ruth Bird [23SEP]
Prof Oumar Gaye [23SEP]	
Prof Oommen John [21SEP]	
Dr Francis Makiya [21SEP]	
Dr Eva Marwali [21SEP]	
Dr Kathryn Mngadi [23SEP]	
Dr Kalaiarasu Peariasamy [23SEP]	
Prof Luis Felipe Reyes [23SEP]	
Dr Sally Shrapnel [21SEP]	
Dr Nicola Stingelin-Giles [21SEP/23SEP]	
Prof Annelies Verbon [21SEP]	
Robert Terry (Chair) [21SEP/23SEP]	

Apologies: Prof Alistair Nichol

OVERALL SUMMARY

IDDO are grateful for the time taken by DAC members at these meetings to provide questions, comments and suggestions to support the development of DAC governance and processes for the COVID-19 Data Platform.

Specifically, outcomes of these meetings were as follows:

1. The role of the DAC was outlined by the Chair, emphasising the DAC role of facilitating access to data.
2. The background and aims of the platform were introduced, noting the wish from contributors that data be shared more widely to advance science – the purpose of the COVID-19 Data Platform.
3. The platform's governance framework was described, highlighting key elements of protections for the data contributor and patients, oversight, security and intellectual property.
4. DAC members reviewed the *Terms of Reference* and *Data Access Guidelines* and posed questions, with no reservations recorded for any members present. Further review will continue until 28 September.
5. The DAC and Secretariat agreed next steps for launching the DAC and the COVID-19 Data Platform, as outlined in the Summary of Action Items on page 2 below.

SUMMARY OF ACTION ITEMS

Issue	Tasks to be completed	Person(s) Responsible	Target Date
Any follow-up questions or queries	If DAC members have any additional questions, queries or wish to follow-up with IDDO on any of the issues discussed please do not hesitate to contact us	All	As required
Meeting Minutes (Part 1)	Slides and summary of the meeting, next steps and action items.	Secretariat	COMPLETE
Meeting Minutes (Part 2)	Full summary of both meetings shared with DAC members.	Secretariat	COMPLETE
DAC comments on governance documents	DAC members are requested to return all comments on the *DAC Terms of Reference* and *Data Access Guidelines* by the end of this week.	DAC	28 September 2020
Finalise governance documents	The Secretariat will incorporate DAC comments and share the *DAC Terms of Reference* and *Data Access Guidelines* with the DAC for final approval.	Secretariat / DAC	9 October 2020
Share *Data Transfer Agreement*	The Secretariat will alert the DAC once the *Data Transfer Agreement* is complete and share for reference.	Secretariat	9 October 2020
Access to online DAC review space	DAC members will be given access to the online review space in Microsoft Teams and provided with further guidance on the review process.	Secretariat	9 October 2020
Share data inventory for launch	The Secretariat will keep the DAC informed of the date for launch of the platform and provide access to the data inventory once it is ready.	Secretariat	16 October

MEETING NOTES

PLEASE REFER TO MEETING SLIDES FOR FURTHER DETAILS

A. WELCOME AND INTRODUCTIONS

1. Introductory comments

a. DAC members were welcomed by the Secretariat (Laura Merson and Matthew Brack) and Chair, Robert Terry, and thanked for agreeing to be a part of the Committee.

b. The Chair noted that IDDO has some of the best experience out there in developing and running data sharing platforms, and with this latest iteration we are very close to something that we know will work, accepting that in the current environment everything is subject to change.

c. It was emphasised that this is a Data Access Committee, whose job is, wherever possible, to facilitate access and use, and not the corollary of preventing access and thinking of reasons not to share the data.

B. COVID-19 DATA PLATFORM: BACKGROUND, OBJECTIVES AND TIMELINES

1. Background and objectives

a. This project is a collaboration between many groups, but in particular the International Severe Acute Respiratory and emerging Infections Consortium (ISARIC) and the Infectious Diseases Data Observatory (IDDO), both research groups at the University of Oxford.

b. ISARIC is a network of research networks with a focus on respiratory epidemics and pandemics, born of the 2009 H1N1 pandemic when it was realised that the global research community is not prepared to execute research in these types of outbreaks.

c. In January 2020, ISARIC had the network, partners and trust in place to spawn a global collaboration around COVID-19 data collection efforts, helping to put together standardised case report forms for emerging infections in collaboration with the WHO and many other groups.

d. ISARIC patient data contributions now amount to over 113,000 individuals diagnosed with COVID-19 from 612 sites in 46 countries, with a growing focus on low- and middle-income countries (LMICs). ISARIC have been producing regular public reports and academic publications on these data internally, but looking for more external use of data to maximise science.

e. To do this ISARIC teamed up with IDDO to develop a data platform to enable broader use of these data. IDDO has a number of data sharing platforms, which began with the Worldwide Antimalarial Resistance Network (WWARN) platform 12 years ago, which now holds the majority of global malaria ACT clinical trial patient data.

f. IDDO has expanded to other data platforms including Ebola and visceral leishmaniasis with a focus on poverty-related diseases. There is now a large infrastructure and a huge amount of experience behind the ethics, governance, technical and processes that are needed for very robust data sharing frameworks within IDDO.

g. IDDO focuses on individual patient level data (IPD) – the specific variables for each patient submitted, not aggregate-level data. IDDO cleans, standardises and maps the data to an international standard called CDISC (Clinical Data Interchange Standards Consortium). This is important to increase the efficiency of re-use due to the heterogeneity of global data.

h. Data are released through independent DAC approval and within a data governance structure in order to enable researchers and policymakers to understand the disease better and support the pandemic response.

i. *How will the data be accessed? Where will the analysis of data be done?* A metadata inventory will be publicly available at launch. This is the first place that people will go to see what types and volumes of data are available. From this they will select the data variables and populations they require access to in order to answer their specific question, and those data will be listed on an application for this committee to review, along with an analysis plan of how they will use those data for their question. If the DAC approves the application, the approved data that they require will be transferred to them under a Data Transfer Agreement and they will use those data to execute that analysis to produce results which will be made publicly available. They will then destroy the dataset.

j. *Will the DAC be able to encourage collaboration between two groups addressing the same research question?* This is for the DAC to determine. The platform aims to minimise competition between researchers and ensure that researchers are answering as many questions as possible about this pandemic. If overlapping or duplicate questions come in then collaboration could be encouraged should the DAC wish to. Protocols will be published as they are approved, and so prospective applicants will be able to see the analysis that has already been approved.

2. Timelines

a. ISARIC contributors have been invited to join the data sharing platform, with 213 sites signed up thus far. This number is expected to increase. Those who have agreed and signed off on the paperwork for

the platform, the metadata about those data will be listed on an inventory, showing volumes and types of data available.

b. The DAC is constituted and we are aiming to sign off on the governance framework this week. The platform is in the final stages of technical development and it is planned to launch access to data on the platform next month in October. Applications for access to data would be expected for DAC review after that time.

C. COVID-19 DATA PLATFORM GOVERNANCE

1. Platform governance

a. Within the context of a data sharing platform, two types of governance are referred to at IDDO:
   i. *Platform governance:* This covers the purpose and direction of the platform, its aims and objectives, and how the committees function.
   ii. *Data governance:* This describes the mechanisms used by the platform to govern security and privacy, and access to data.

b. Key points concerning platform governance:
   i. This is the ISARIC COVID-19 Data Platform, and ISARIC make decisions on the platform's direction. The platform is hosted and managed by IDDO.
   ii. The DAC is independent, with decisions by consensus or vote where needed. The Chair has the casting vote. Details are found within the *DAC Terms of Reference*.
   iii. External parties, such as the ISARIC Senior Management Team and IDDO Board may be consulted if the DAC agree this is needed.
   iv. Platform activities, including those of the DAC, are supported by the Secretariat based at IDDO.
   v. The governance framework is based on IDDO experience developed from the Ebola Data Platform and other IDDO platforms over the last 12 years.

2. Data governance

a. Data are governed by a framework of guidance, legally-binding documents and processes:
   i. *Privacy Impact Assessment:* To be able to handle this type of individual patient data, IDDO has had to work with the University of Oxford to identify risk factors, described in a *Privacy Impact Assessment*.
   ii. *Data Security Model:* Risk factors identified within the *Privacy Impact Assessment* were then addressed in a *Data Security Model* describing actions to ensure data protection and security provisions.
   iii. *Terms of Submission:* Once a secure environment for data curation has been established, the platform can receive data from a Data Contributor with their signed *Terms of Submission* (see more detail below).
   iv. *Data Access Application Form:* Once data has been received to the platform, a potential Data Recipient can submit an application form for review by the DAC.
   v. *Data Access Guidelines:* These guidelines provide clear criteria against which the application is assessed by the DAC.
   vi. *Data Transfer Agreement:* If approved by the DAC in accordance with the *Data Access Guidelines*, data are transferred under a legally-binding *Data Transfer Agreement* between the Data Recipient institution and the University of Oxford.

3. Terms of Submission and Data Transfer Agreement

a. The *Terms of Submission* and *Data Transfer Agreement* are key documents in governing the relationship between the Data Contributor and the Data Recipient. These are both legally-binding agreements that require authorised signatories from an institution that is a legal entity.

b. The relationship between Data Contributor and Data Recipient is defined by and overseen through these data agreements with IDDO and the University of Oxford, where the University is the legal entity.

c. The two agreements mirror each other in many aspects in order to match expectations and requirements between Data Contributor and Data Recipient. In particular, each of them address:
   i. Limitations on data access and use, protecting contributors and patients.
   ii. Oversight of compliance with terms of agreement.
   iii. Security of data quality and privacy.

d. Examples of limitations on data access and use:
   i. *Terms of Submission:*
      1. Submitted under ethics and regulatory approvals.
      2. IDDO can only use data for curation and managed access for research.
      3. No limitations on Data Contributor use of data.
   ii. *Data Access Agreement:*
      1. Recipient may only use Dataset for research conducted within designated Research Team.
      2. No re-identification, linking or sharing the dataset with anyone not covered by the *Data Transfer Agreement*.
      3. Ensure ethics and regulatory approvals are obtained before the dataset is used.
      4. Maintain compliance with applicable data protection regulations while observing ethical standards.

e. Examples of oversight of compliance:
   i. *Terms of Submission:*
      1. IDDO provides reasonable assistance to Data Contributor in enforcing terms of the *Data Transfer Agreement*.
      2. Ensures data use in compliance with ethical and regulatory approvals, and governance framework.
      3. Agreement can be terminated by the Data Contributor or IDDO at any time.
   ii. *Data Transfer Agreement:*
      1. Data Recipient must comply with instructions on use of data conveyed by Data Contributor via Secretariat.
      2. Notify IDDO of changes to research and obtain written approval from DAC to pursue changes.
      3. Describe scientific and societal benefits generated from the Dataset at end of duration of *Data Transfer Agreement*.
      4. Disclose to IDDO all research results by end of duration of *Data Transfer Agreement*.

f. Examples of security of data quality and privacy:
   i. *Terms of Submission:*
      1. Personal Data only processed for purposes of curation and de-identification.
      2. Controlled access to personal data on secure server – never shared with third party and destroyed after curation and de-identification.
      3. Compliance with data protection legislation.
   ii. *Data Transfer Agreement:*

1. The Data Recipient takes appropriate security measures to protect the dataset from unauthorised access.
2. Data are to be stored on encrypted, access-limited, password-protected servers.
3. The Data Recipient shall not disclose the dataset and its information beyond the Research Team without DAC written consent.

4. Intellectual Property (IP)

a. The approach to IP, common across all IDDO platforms, is that the platform is designed with the assumption that the majority of data uses will be for academic research that will be openly available for public consumption and advancing science.

b. The *Data Transfer Agreement* says that use of the data will not be for any purpose that leads to any form of IP licensing. That is not to say that use would be prohibited for companies that wish to contribute to ending the pandemic.

c. If a company or group wishes to do this, they must return to the University of Oxford to negotiate appropriate terms around that. Meanwhile, it is very difficult to define one rule that will apply to all cases. Each case will be discussed first with the DAC based on precedent from real examples.

d. The DAC agreed that this was a sound position to take on IP, acknowledging that as the project develops we may need to examine rules and regulations further. With the pressure on academia and emerging commercial collaborations it is important to keep IP under review in this context.

e. The Chair noted a key consideration that there is not any application of IP that may restrict the use by others of the same data. If someone can take these data and do something with them through IP that promotes the pandemic response, that should be encouraged, provided the IP does not prevent further re-use of data.

f. The DAC asked about the availability of data and preference for those Data Contributors who made the effort to prepare the data. In the case of this platform, there exists within ISARIC a collaboration where all those contributing may use the entirety of the ISARIC data to produce science. Separately, those within that collaboration who wish their data to be more openly available to third parties have added their data to this platform hosted by IDDO to increase utility of the data further.

g. The DAC recommended that the sharing of analysis code should be mandatory. While it is difficult to police this, the platform asks that any research results be published within an open-access publication. A request to additionally make analysis code openly available will be added.

5. Additional questions

a. *If a commercial company requests data, can they request data as a company or do they need to be affiliated with an academic researcher?* A request from a commercial entity will be assessed applying the same guidelines as any other applicant. The fact they are commercial is not a basis for refusal. The important elements will be that they have a quality protocol and the capacity to undertake the research and comply with the requirement that the outputs are made open access. As long as the purpose is health and research it would not be a barrier to access, but it will be for the DAC to decide.

b. *In relation to the Data Transfer Agreement and alignment with University of Oxford data policies influenced by EU GDPR, how will the DAC balance stringent Oxford guidance recognising that these standards may not be practical for some institutions to implement?* The platform supports LMICs and other countries outside the EU by managing the GDPR requirements within IDDO. This is achieved by removing all personal identifiers before sharing data. Some variables that may increase the risk of

identification (e.g. pregnancy status, date of infection, age category) will be verified by IDDO through a statistical disclosure analysis on the dataset before it is released to ensure that the data are below acceptable risk as defined by the European Medicines Agency (0.09% chance of any person being identified). The key responsibility of the Data Recipient is data security, the requirements of which are defined in the Data Transfer Agreement.

c. *Does the applicant need to provide evidence that institutional approval is provided before the application or after the approval?* Currently the application form asks whether the applicant's institution has agreed to execute the Data Transfer Agreement and that is taken on good faith – this approach has worked across IDDO platforms to date. This can be reviewed if it becomes a problem. A pre-signed DTA or a letter of support from the institution, for example, could be provided with the application if the DAC felt this was necessary.

D. ROLE OF THE DATA ACCESS COMMITTEE

1. Role of the Data Access Committee

a. The Chair presented the role of the DAC, emphasising that it is the role of the DAC to facilitate access to data wherever possible. This is a sub-set of the larger ISARIC dataset that is intended to be public and widely used by anyone who has a legitimate reason to do that. The default is that we want to share.

b. The DAC's job is to think from a starting point of: is there a legitimate reason that may cause us some concern? That would be anything that might compromise the continuing operation of the platform, for example: research that was in some way unethical; any kind of breach of the privacy and security standards of the platform; or any research falling outside expected norms and standards.

c. It is also the role of the DAC to make decisions in line with the expectations of what the Data Contributors considered these data could be used for, and what research has been prioritised by the published research agendas. It is the DAC's role to ensure that re-use lies within those principles.

d. One of the issues will be an assessment of the science itself. The DAC need to hold carefully their assessment of the science. Other DACs have provided feedback to help an application improve on the science, but that wouldn't necessarily be a starting point to block access.

e. The platform may need to be flexible on how experience and data analysis can be reflected on the application. Ministries of health have requested data through analysts who have not had a PhD. The statistical analysis plan sent to the DAC should help to judge competence.

f. The DAC supported the testing of a secure analytic environment for the data as a means of increasing equity across data users through provision of software through the environment, with fewer data security concerns and more streamlined access, also holding potential as a training opportunity for LMICs.

2. Data Access Committee review process

a. DAC review process is planned as follows, supported by the Secretariat:
   i. *Data Access Application:* This is received from a prospective Data Recipient and is reviewed first by the Secretariat to confirm that all elements of the application are complete. Any incomplete applications are returned to the applicant for re-drafting. Only complete applications are shared with the DAC for review.
   ii. *DAC review:* An email from the Secretariat will alert members to applications ready for review, along with a link to an online Data Access Request Evaluation (DARE) form. As per the *DAC*

Terms of Reference we are asking DAC members to review applications within three working days. Members are asked to provide feedback on the reasoning for any decision other than 'approval'.

iii. *External support:* The ISARIC Senior Management Team and IDDO Board are available for consultation if the DAC agree that this is needed.

b. An online private shared space is being prepared for DAC members that will include all applications for review, governance documentation for reference and guidance for the review process. This is a new working environment for this DAC in order to facilitate and more rapid and straightforward review process.

c. In particular, an online review form is being prepared that will correspond exactly with the DARE form shared in the meeting pre-reads. This will be shared via a link and data from completed forms submitted by the DAC will populate a spread sheet that can be reviewed by the Chair.

d. The Chair emphasised that the system was designed in order to be as efficient as possible, allowing for 'asynchronous' reviewing (so all members can do their review at different times), to let the DAC complete the six reviews required to approve an application in the three working-day turnaround time.

e. *Is it permitted for a DAC member to be a co-author on research resulting from an application to access data on the platform?* If that were a possibility, before access was granted the platform would want to ensure that the six approvals on that application were provided by DAC members who would not be involved in any of the outputs. Involvement in research using the platform data would require the DAC member to abstain from the review process due to conflict of interest. If the conflict happened in arrears, this could be perceived as a conflict externally and so it would need to be ensured that the application had received a sufficient number of approvals from the other DAC members, who would need to be aware of the conflict of interest.

f. *If there is competition for a research question, is there a time limit on research conducted with platform data to ensure that the research is completed?* Within the Data Access Application Form there is a space for a publication plan and timeline. The DAC can then decide whether that plan and timing is reasonable. The terms of the Data Transfer Agreement will expire at the end of the period indicated in the publication plan and agreed by the DAC. Data is not consumable, so there is also nothing preventing release of a similar dataset to another group at the same time.

g. *Will the platform be able to police use of data, particularly concerning use of data for research not included in the application form?* The platform is not intending to police what researchers do, but everyone is involved in the COVID-19 community and sees the research outputs – if the larger ISARIC community saw that research has been produced that was against community regulations then action will be pursued. The DAC will not review research prior to publication due to the potential workload and the risk of creating a double jeopardy if there was a conflict with the journal decision. But this can be reviewed.

E. DATA ACCESS COMMITTEE TERMS OF REFERENCE

a. The Chair asked the Committee for feedback on the *Data Access Committee Terms of Reference*, noting that these governance documents are to be openly available to the public and subject to change.

b. *Has the generation of synthetic data been accounted for?* If someone were going to use the data for that purpose, they would need to specify that in their application. That question would come to the DAC and the outputs of that research would need to be considered. The Secretariat could support

that discussion within the DAC. All research outputs must be shared back to the platform in accordance with the *Data Transfer Agreement*, including derivative data.

c. *How is the scientific merit weighed in DAC decisions?* The DAC is not a scientific review committee, though that needs to be considered in line with broader principles of the research agenda. If the question is far outside of the scope of the data, we would raise those concerns. Where the protocol may not be perceived as robust by certain members, feedback may be provided, but this is not necessarily a starting point for refusal. Unless they will not be able to answer the question submitted in the proposal, we do not see a reason to refuse access. It is approached as an opportunity to support good science, rather than stop other methods from being tested and applied.

d. *What if not all centres have explicitly agreed to share data on this platform? What about attribution?* The DAC will only be providing access to data from those Data Contributors who have signed a Terms of Submission agreed to data access via DAC review. For attribution, it is required within the *Data Transfer Agreement* that the institution who has contributed the data be acknowledged, as a minimum. It is also a requirement that anyone who uses the data contact the lead coordinator at each hospital to invite them to participate in that analysis. We are not prescriptive on how that plays out, it is between the Data Recipient and the Data Contributor.

e. *How does avoidance of duplication of research work in practice? Is there some place where all the analysis will be made public?* All approved applications will be made publicly available and prospective applicants will be directed to that resource, which will also include all ISARIC analysis plans. Enabling duplication of science could be a risk factor affecting the ability of the platform to operate, and with so many questions to answer in relation to COVID-19 it is for this Committee to decide how best to spread the scientific expertise across as many questions as possible.

f. *What about demonstrating reproducibility in science for validation, which requires duplication?* It would need to be determined if the reproducibility was for methodological reasons or whether it appears to be a duplication of the same protocol. The question is open for all to answer by the DAC, and this is why the Committee has been formed from such a representative group of members, in order to be able to make such informed decisions. The aim of the platform is not to support unnecessary competition but better science.

g. *If the DAC are challenged when making a decision, can they seek advice from external parties?* If the DAC are presented with a problem that can't be solved within the Committee, the *Terms of Reference* give the DAC freedom to consult outside of the Committee if the DAC agrees it is needed.

h. No DAC members reported any reservations regarding the *Terms of Reference*.

i. *Would more than six DAC members be required for quorum during the first review of applications to promote learning and cohesion?* Once the DAC reaches six approvals with no objections, that should be sufficient. If there are many objections, that may need a discussion.

F. DATA ACCESS GUIDELINES

a. The Chair asked the Committee for feedback on the *Data Access Guidelines*, reminding the Committee that they have the power to review and update these guidelines.

b. *Does the requirement for 'an academic record consistent with the proposed analysis' exclude early career researchers? What track record are you looking for?* The research and supervisory track record requested is to help ensure that the research can be undertaken. As these platforms go forward, they could be excellent resources for PhD or even MSc students. By requesting an appropriate supervisor with suitable academic standing, this is also another reassurance at this early stage that the data will

be used appropriately by the right people. Data Contributors have been concerned that their data might be misused for poor science. It is not to prohibit use by junior resources.

c. No DAC members reported any reservations regarding the *Data Access Guidelines*.

d. The Chair reminded the Committee that the *DAC Terms of Reference* and *Data Access Guidelines* will be made publicly available to demonstrate transparency and efficiency, understanding that there is still resistance to sharing data in the research community.

e. *What is the anticipated volume of data access requests?* This is not known, but a high volume are anticipated. This is why a fairly flexible approval system has been designed. There are 16 DAC members and we require six members to approve without anyone objecting, so not everyone has to review every application in order to distribute the burden of review. We hope that all members can review at least half of all applications. The whole process will be kept under review between DAC members and the Secretariat.

f. *Will the data be anonymised?* The platform receives data that are called 'pseudonymised' under GDPR. These data go through an anonymization verification to find any direct identifier that might have been erroneously included. After curation and before data are released, they go through a statistical disclosure process to ensure that patients cannot be reasonably re-identified.

G. NEXT STEPS

1. *DAC comments on governance documents:* DAC members are requested to return all comments on the DAC Terms of Reference and Data Access Guidelines by Monday 28 September. The Secretariat will incorporate DAC comments and share the *DAC Terms of Reference* and *Data Access Guidelines* with the DAC for final approval the following week. These will remain iterative documents and may still be amended at a later date.

2. *Data Transfer Agreement:* The Secretariat will alert the DAC once the *Data Transfer Agreement* is complete and share for reference. Meanwhile DAC members can consult the existing agreements made publicly available by IDDO, on which the COVID-19 agreement is based. It was noted that the Ebola Data Platform agreement places strong emphasis on benefits for LMICs. COVID-19 is a different disease and population, so will have a different approach. If data from LMICs are requested, then we ask for some words on capacity strengthening and how that is considered in the use of these data and dissemination of outputs.

3. *Access to online DAC review space:* DAC members will be given access to the online review space in Microsoft Teams and provided with further guidance on the review process during October.

4. *Share data inventory for launch:* The Secretariat will keep the DAC informed of the date for launch of the platform and provide access to the data inventory once it is ready.

5. *Sharing experience with the wider community:* The DAC outputs belong to the Committee and we hope that the DAC will be interested to share their experience in operationalising data sharing with the wider community. Anyone wishing to write up these outputs and take a role in leading the academic thinking around how initiatives like this should move forward would be warmly encouraged to do so.