Dear colleagues-

Please accept our apologies for the length of this email! There are three data updates we would like to announce regarding the following:

1. Human pre-COVID suspension extended QC
2. PASS1B-06 data freeze
3. PASS1C merged phenotype data

We have also included at (4) a reminder for Consortium members of how to access MoTrPAC GCP Storage.

1) We are pleased to announce that the human pre-COVID suspension extended QC dataset has been prepared and readied for internal consortium use. As outlined in the data sharing plan, and per approvals from the Steering Committee and DSMB, these data should be used to better understand and extend the QC of pre-COVID suspension data generated by the CAS sites and are not to be used for presentation or publication pending review by the Steering Committee. As a reminder, there are three levels of human phenotypic/clinical data:

-batching QC – used for batching at CAS sites

-extended QC (EQC) – used for QC/QA and data exploration across the consortium

-full phenotype data – used for analysis and preparation for presentation/publication subject to approvals of PPD, SC, and DSMB.

At the present time, the ‘full phenotype data’ is being revised by the clinical sites and the DMAQC due to issues found in the DMAQC QC of these data. We anticipate that the majority of pre-COVID suspension sedentary adult ‘full phenotype data’ will be ready for consortium use at the completion of this revision and following preparation of these data for internal sharing; while the ‘full phenotype data’ from PED and HA participants will be held pending additional participant recruitment and data collection in keeping with the recommendations of the Steering Committee and DSMB.

Currently, EQC data are accessible through a MoTrPAC Google Cloud Platform Storage bucket at

|  |
| --- |
| gs://motrpac-pre-covid-human-eqc-data |

and only available to Consortium users who have been granted read-only access to this GCP Storage bucket. Please contact BIC at [motrpac-helpdesk@lists.stanford.edu](mailto:motrpac-helpdesk@lists.stanford.edu) if you need access to the bucket in question.

The readme for these data is available at

gs://motrpac-pre-covid-human-eqc-data/README.md

2) The long-awaited ‘data freeze version’ for PASS1B-06 has been finalized. This data freeze is the result of a truly colossal team effort, and we would like to acknowledge our appreciation of the effort of all who have contributed. These data are expected to form the basis for the PASS1B 6-month landscape manuscript and related companion manuscripts as proposed to and approved by the PPD. As these data have been extensively QC’d, we are hopeful that there will not be a need for an additional versioned dataset. However, if you or your colleagues identify any issues, please document said [issues](https://docs.google.com/document/d/1MnDbmKkvroR6cARF_svUXokP9aJc1Qq2LSPYAvzrck8/edit) (<https://docs.google.com/document/d/1MnDbmKkvroR6cARF_svUXokP9aJc1Qq2LSPYAvzrck8/edit>) and send an email to the BIC at motrpac-helpdesk@lists.stanford.edu.

We have provided read-only access to the data freeze files through the MoTrPAC Google Cloud Platform Storage bucket accessible at

|  |
| --- |
| gs://motrpac-data-freeze-pass |

Results files (e.g., quantitative results files processed by BIC pipelines or provided by CAS-sites) are at

|  |
| --- |
| gs://motrpac-data-freeze-pass/pass1b-06/v1.0/results |

Analysis files (e.g. normalized sample-level data, differential analysis results, intended to be used for landscape manuscript) are at:

|  |
| --- |
| gs://motrpac-data-freeze-pass/pass1b-06/v1.0/analysis |

The data dictionary with the files’ details is available [here](https://docs.google.com/spreadsheets/d/10I4ofH__sMhlONrwdH8ms9S4qkXpQ_wXMjcCa1fBUyY/edit?usp=sharing).

(<https://docs.google.com/spreadsheets/d/10I4ofH__sMhlONrwdH8ms9S4qkXpQ_wXMjcCa1fBUyY/edit?usp=sharing>)

These files are only available to Consortium users who have been granted access to this GCP Storage bucket. Please contact BIC at [motrpac-helpdesk@lists.stanford.edu](mailto:motrpac-helpdesk@lists.stanford.edu) if you need access to the bucket in question.

3) PASS1C merged phenotype data is now available to Consortium users for purposes of extended QC and initial pre-data-freeze analyses and exploration of the integration with PASS1A data. These data are available to registered Consortium users as read-only files through the MAWG bucket at:

|  |
| --- |
| gs://mawg-data/pass1c-06/pheno\_dmaqc/merged2021-09-15 |

As this data is being provided through early data access, if issues are identified please contact the BIC at motrpac-helpdesk@lists.stanford.edu. These data are now ready for Consortium QC.

4) For Consortium members, we provide access to the MoTrPAC GCP Storage through Google Authentication, in which a Google account registered with your institutional email addresses is required (*see instructions below*). We request that large data pulls (>1 Gb) be limited as possible (e.g. once per site) as egress costs can be significant and the Consortium does not have unlimited cloud budget.

### Instructions on registering with Google:

1. Go to https://accounts.google.com

2. Select "Create account" and then the "For myself" option.

3. Select "use my current email address instead" on the following screen and use your institutional email address.

Thanks all for your patience and reading.

Sincerely

-the BIC on behalf of the MoTrPAC Consortium