Long COVID Computational Challenge (L3C) Frequently Asked Questions

Challenge Overview

1) **How are COVID-19 and PASC/Long COVID defined in this Challenge?**
   Answer: This Challenge uses the following definitions:
   - **COVID-19 Positive:** Patients who have had a positive COVID-19 PCR test (not just a diagnosis)
   - **Post-acute sequelae of COVID-19 (PASC)/Long COVID:** Patients who have an ICD code U09.9 recorded in the dataset are labeled as true positive patients with PASC/Long COVID

2) **What does the model need to produce?**
   Answer: Solutions should provide a percentage likelihood of developing PASC/Long COVID among patients who have tested positive for SARS-CoV-2.

Challenge Eligibility

3) **How do I determine if I register as an individual, team, or entity?**
   Answer:
   - **An Individual is a solo participant.** To be eligible to receive a cash prize, the Individual must be a citizen or permanent resident of the United States. Note: Even if you are a single participant, you will need to be on a registered Team to participate.
   - **A Team is defined as a group of one or more individual(s) who are competing in the challenge together.** If the Team is made up of members from different institutions, each institution must have their own signed data use agreement (DUA) with National COVID Cohort Collaborative (N3C). By default, the participant who creates a team is the "Team Captain" who will register and submit on behalf of the Team members. The Team Captain is responsible for all communications with the Challenge sponsors and, in the event of winning a cash prize, will be paid the prize in full. To be eligible to receive a cash prize, the Team Captain must be a citizen or permanent resident of the United States.
   - **An Entity is defined as a group of individuals competing on behalf of a legally established organization, institution, or corporation.** Each participating Entity is required to identify a Point of Contact who will register and submit on behalf of the Entity. The Point of Contact is responsible for all communications with the Challenge sponsors. In the event of winning a cash prize, the prize will be paid directly to the Entity, not to the Point of Contact. The Entity must have a signed DUA with N3C.
   - Regardless of registration, participants are required to acknowledge whether federal funding will be used in the development of the Challenge submission.
   - All Participants must certify they have read and understand and agree to abide by the official eligibility criteria, participation rules, and requirements for the Challenge as stated in this announcement. See the L3C Challenge for a complete list of eligibility rules.

4) **If competing as a Team, how should I structure my team?**
   Answer: It will be up to each team entering to determine the composition of their respective teams and ensure the team meets the eligibility criteria. There is no limit on the number of members who can participate on a team. Registrants should consult with their institution to

determine any requirements they may have regarding participation. See the L3C Challenge for a complete list of rules.

5) **Can participants outside of the United States participate?**
Answer: In the case of Individuals and Teams competing in this challenge, non-U.S. citizens and non-permanent residents are not eligible to win a monetary prize (in whole or in part). However, non-U.S. citizens and non-permanent residents can participate as a member of a team that otherwise satisfies the eligibility criteria. Their participation as part of a winning team, if applicable, may be recognized when the results are announced. In the case of Entities competing in this Challenge, those that are not incorporated in and do not maintain a primary place of business in the United States are not eligible to win a monetary prize (in whole or in part). See the L3C Challenge for a complete list of rules.

6) **What are the criteria for an independent citizen scientist to participate in the Challenge?**
Answer: If an individual is not affiliated with an institution, they are termed citizen scientist. If a citizen scientist wishes to participate in the challenge, it is recommended they partner with an institution that can sign a Data Use Agreement (DUA) with N3C. Citizen scientist can review the guidance on L3C Challenge and the Synapse site for more information.

7) **I've submitted the Data Use Request, but my request is still "Pending Approval", what should I do?**
Answer: Go to the How to Participate page on the Synapse site and make sure you've followed the first three Synapse registration steps. In particular, make sure you have followed the instructions carefully for joining the *existing* Data Use Request under Step 5 of the above linked How to Participate page.

8) **My Data Use Request has been accepted, but I'm not able to access the project yet, what should I do?**
Answer: The N3C Data Enclave team needs a day or two to provision your workspace. You should receive an email a day or two after you've been approved with details about how to access the project.

National Covid Cohort Collaborative (N3C)

9) **Who can access N3C?**
Answer: The N3C Data Enclave can be accessed by anyone who has completed the onboarding process and has submitted a DUR that is approved by the Data Access Committee. More information can be found at the N3C Website.

10) **What data elements are included in the N3C Data Enclave?**
Answer: The N3C enclave includes clinical observations, lab results, medication records, procedures, and visits. Please visit the following websites for a full list of elements
- N3C Data Dictionary

To learn more about N3C data, please visit the following resources:
- N3C Website
- N3C External Data
- N3C Cohort Information

11) **What guidance can you provide on working with the N3C Data Enclave tools and data?**
Answer: The L3C Challenge will host office hours on Thursdays at 10am PST, starting on October 13th and ending on December 15th. Visit the Synapse site to join or sign up for the Zoom

meeting. Additionally, there will be three tutorials to provide an overview on the challenge and using the Enclave:

September 29, 2022
Topic: N3C Orientation and Challenge Instructions
Recording Link

October 3, 2022
Topic: Code Workbooks and PySpark
Recording Link

October 10, 2022
Time: 10:00am PST
Topic: Code Performance Optimization
Registration Link

These will be recorded and made available here.

Additionally, you can view the Resource Page on the Synapse site and consult N3C training resources.

12) What tools and computing resources are available in the N3C Data Enclave?
Answer: The N3C Data Enclave includes an expansive suite of tools geared to help you better discover, explore, and analyze N3C clinical data. For more information visit the N3C website. Researchers can utilize familiar software, such as R and Python, to gain insights within the Enclave, and can also take advantage of core Enclave tools built specifically to enable analysis within the environment. The software tools suited for the N3C Data Enclave were selected for their popularity and ease of use. Additional tools are available on the N3C website. You can also attend the L3C Challenge Office Hours. Please visit the Synapse website to register for the Challenge and gain access to and learn how to use all available resources.

13) What is the prediction window for the Challenge?
Answer: The model should make a prediction of a person's risk for developing long COVID in the future following an initial acute COVID diagnosis. That prediction should be based on the person's clinical history prior to the initial acute COVID diagnosis, and/or based on data available from the time of acute COVID diagnosis until 4 weeks following the acute COVID diagnosis. Data beyond 4 weeks following an acute COVID diagnosis should not be used to predict long COVID risk associated with that particular acute COVID episode.

14) How are patients with multiple COVID diagnoses handled?
Answer: Use the date of the first COVID diagnosis for the beginning of the prediction window. The uncensored training data will contain all the patient information, but the censored data will be cut off four weeks after the initial COVID diagnosis.

15) How can I bring outside code into the enclave?
Answer: If you have code or software that you would like to use for the challenge, there are mechanisms that can be used to import your code into the N3C enclave. Please read the guidance on "Importing External Software" on the Synapse site.

Submission

16) Is there a required language for the code?

Answer: There is not a required language, but Challenge participants will have to use the coding tools available in the N3C Data Enclave, which can run R, Python, and SQL.

17) **How will participants submit their code?**
Answer: Participants will submit a Code Workflow in the N3C Data Enclave. You will create a folder named "submission" in your team folder. You are expected to place your final workbook in this folder. If there is no "submission" folder than you will not be considered a submitting team.

Publication

18) **Can we publish our findings?**
Answer: Yes. Challenge participants that intend to publish a Challenge related preprint (to a preprint server) or paper (to a journal or conference) are highly recommended to:
- Acknowledge the National Institutes of Health whenever publicizing the work under this Challenge in any media or publication by including an acknowledgment as follows: "The [analysis/model/algorithm] described in this [publication/report/presentation] was conducted as part of activities associated with the NIH Long COVID Computational Challenge (L3C). The L3C Challenge was supported by the Rapid Acceleration of Diagnostics (RADx®) Initiative and by the efforts of the National COVID Cohort Collaborative (N3C) and NCATS".
- Acknowledge the use of NCATS N3C data and adhere to data user code of conduct, please see N3C data access guidelines and N3C data user code of conduct.
- The Publication Policy can be found at the Synapse site

19) **What will the Challenge Flagship Manuscript include?**
Answer: The Challenge Flagship Manuscript, published after the Challenge has completed, will describe the findings of the Challenge. It will include details of methods and findings of specific teams, and those teams will be listed as co-authors on the paper.

Resources

20) **What communication platform will be used to communicate among collaborators?**
Answer: You can engage collaborators in project specific discussions on the Discussion page on the Synapse site.

21) **Where can I find the recording of the RADx-rad Long COVID Computational Challenge (L3C) webinar?**
Answer: The recording and slides of the webinar can be found at RADx Events.

Technical Issues

22) **I am having technical problems while registering for the Challenge. Who should I contact?**
Answer: Technical issues should be sent to the Challenge Administration mailbox:
RADxLongCOVIDChallengeAdmin@synapse.org.