Outbreak data analysis platform

Data held

ISARIC4C has developed a clinical and research data integration platform to facilitate integrative analyses of multi-omic, serial disease profiling, stratified by viral variant, clinical phenotype and outcome. This is hosted on nationally-leading, exabyte-scale computational infrastructure including state-of-the-art security systems for protection of identifiable data, and high-performance CPU/GPU computing (the Edinburgh Parallel Compute Centre, EPCC, and ARCHER/ARCHER2).

The ISARIC4C analysis platform exists to *encourgage and facilitate research* by collating, linking and presenting data together with high-performance computational capacity. This platform now serves as a hub for a coordinated UK national research response to COVID-19. Data are included from:

- ISARIC4C tier 0: (unconsented) prospective clinical data from 204333 cases
- ISARIC4C tiers 1 and 2: serial multiomic assays from research samples of blood, respiratory secretions, urine, and stool from 2306 cases
- COG-UK: (unconsented) summary variant data from COG-UK viral sequencing study is already included for matched patients
- GenOMICC study complete data: microarray and whole genome sequence data from 12454 cases
- PHOSP complete data: follow-up clinical and biological data generated by the Post-Hospitalisation for COVID-19 follow-up study (1075 cases)
- UK-CIC: deep immunological phenotyping data from across the UK Coronavirus Immunology Consortium, using ISARIC4C samples and local collections.

Linkage to clinical data

Research data within the analysis platform is already linked to: - NHS Scotland primary, secondary care and death records - NHS Digital health records data

In future, plans are in place to transfer data to link with: - ICNARC and SICSAG critical care audit databases - NIMS National Immunisation Dataset - Pillar 1 testing - Pillar 2 testing - ONS

Research outputs

The ISARIC Coronavirus Clinical Characterisation Consortium (4C) is the largest observational study of hospitalised patients with COVID-19 anyhwere in the world. Through acquisition, integration and analysis of clinical, biological, genetic and virological data on patients with Covid-19 in UK hospitals, ISARIC4C has provided essential weekly updates to SAGE that guide the public health response, and enabled understanding of the clinical features,[1] prognostication,[2] disease biology[3] and host genetics.[4]

Analysis platform structure

There are two routes of access to the analysis platform: 1. NHS Trusted Research Environment (Safe Haven) for access to personal clinical data and data collected without explicit consent. 2. Rapid-access flexible compute for access to non-disclosive research data collected with explicit consent.

Within both of these environments there is an additional division in the data: 1. Publishable “open access” data which any user can use and report as they wish, according to data protection and privacy rules; 2.

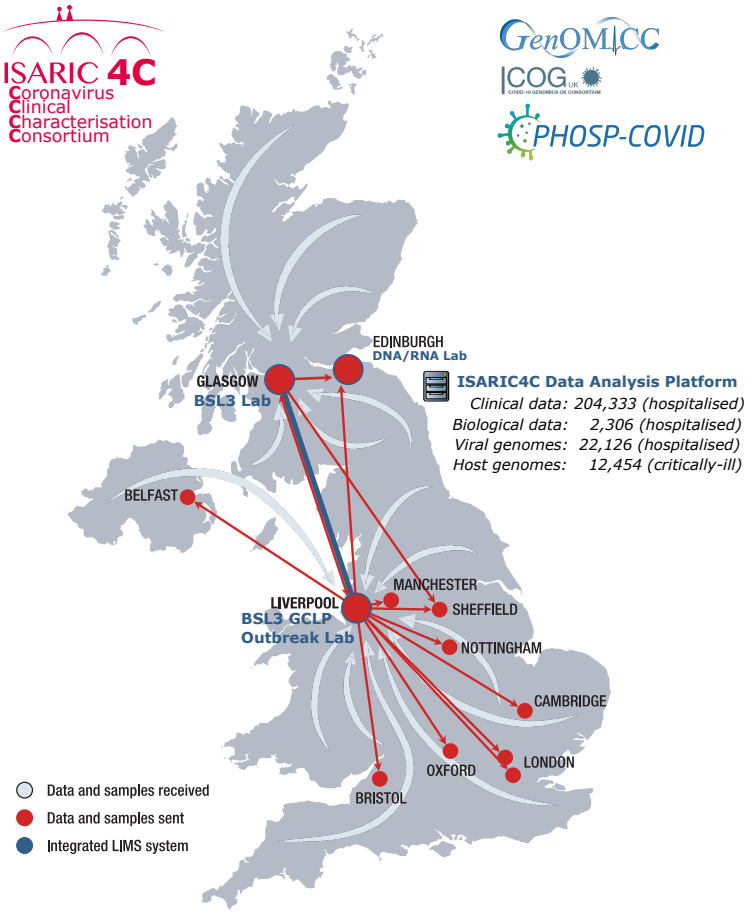

Figure 1: ISARIC4C study and data analysis platform

Embargoed active research data, shared by academic investigators and available for linked analysis but not for publication without agreement from all contributors.

This design is intended to build trust in order to encourage immediate contributions of research data from academic collaborators.

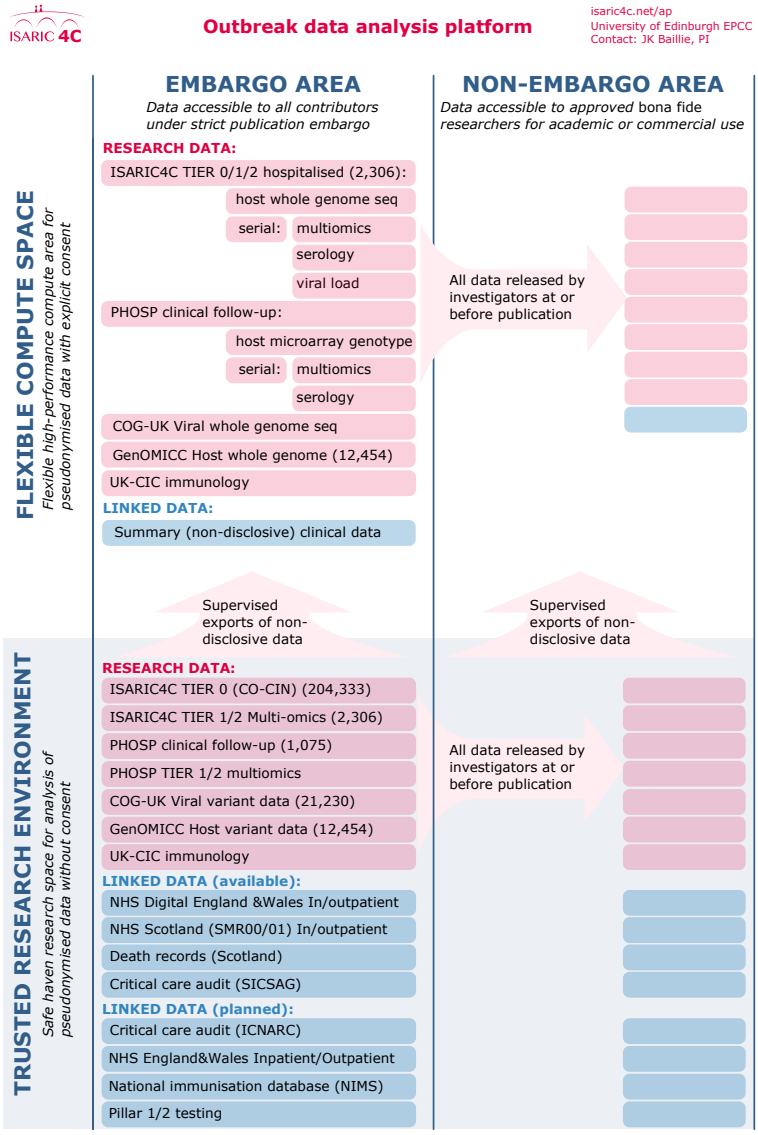

Figure 2: Structure of the ISARIC4C Analysis Platform

Future plans

Rapid addition of viral sequence data from the COG-UK platform will enable real-time detection of the clinical impact of new viral strains, in-depth biological study of reinfection, and host:pathogen interactions at a genetic and mechanistic level.

References

1. Docherty, A.B., Harrison, E.M., Green, C.A., Hardwick, H.E., Pius, R., Norman, L., Holden, K.A., Read, J.M., Dondelinger, F., Carson, G., Merson, L., Lee, J., Plotkin, D., Sigfrid, L., Halpin, S., Jackson, C., Gamble,

C., Horby, P.W., Nguyen-Van-Tam, J.S., Ho, A., Russell, C.D., Dunning, J., Openshaw, P.J., Baillie, J.K. & Semple, M.G. Features of 20 133 UK patients in hospital with covid-19 using the ISARIC WHO Clinical Characterisation Protocol: Prospective observational cohort study. *BMJ* **369**, (2020).

2. Knight, S.R., Ho, A., Pius, R., Buchan, I., Carson, G., Drake, T.M., Dunning, J., Fairfield, C.J., Gamble, C., Green, C.A., Gupta, R., Halpin, S., Hardwick, H.E., Holden, K.A., Horby, P.W., Jackson, C., Mclean, K.A., Merson, L., Nguyen-Van-Tam, J.S., Norman, L., Noursadeghi, M., Olliaro, P.L., Pritchard, M.G., Russell, C.D., Shaw, C.A., Sheikh, A., Solomon, T., Sudlow, C., Swann, O.V., Turtle, L.C., Openshaw, P.J., Baillie, J.K., Semple, M.G., Docherty, A.B. & Harrison, E.M. Risk stratification of patients admitted to hospital with covid-19 using the ISARIC WHO clinical characterisation protocol: Development and validation of the 4C mortality score. *BMJ (Clinical research ed.)* **370**, m3339(2020).

3. Thwaites, R., Uruchurtu, A.S.S., Siggins, M., Liew, F., Russell, C.D., Moore, S., Carter, E., Abrams, S., Short, C.-E., Thaventhiran, T., Bergstrom, E., Gardener, Z., Ascough, S., Chiu, C., Docherty, A.B., Hunt, D., Crow, Y., Solomon, T., Taylor, G., Turtle, L., Harrison, E.M., Semple, M.G., Baillie, J.K. & Openshaw, P.J. Elevated antiviral, myeloid and endothelial inflammatory markers in severe COVID-19. *medRxiv* 2020.10.08.20209411(2020).doi:10.1101/2020.10.08.20209411

4. Pairo-Castineira, E., Clohisey, S., Klaric, L., Bretherick, A.D., Rawlik, K., Pasko, D., Walker, S., Parkinson, N., Fourman, M.H., Russell, C.D., Furniss, J., Richmond, A., Gountouna, E., Wrobel, N., Harrison, D., Wang, B., Wu, Y., Meynert, A., Griffiths, F., Oosthuyzen, W., Kousathanas, A., Moutsianas, L., Yang, Z., Zhai, R., Zheng, C., Grimes, G., Beale, R., Millar, J., Shih, B., Keating, S., Zechner, M., Haley, C., Porteous, D.J., Hayward, C., Yang, J., Knight, J., Summers, C., Shankar-Hari, M., Klenerman, P., Turtle, L., Ho, A., Moore, S.C., Hinds, C., Horby, P., Nichol, A., Maslove, D., Ling, L., McAuley, D., Montgomery, H., Walsh, T., Pereira, A., Renieri, A., Shen, X., Ponting, C.P., Fawkes, A., Tenesa, A., Caulfield, M., Scott, R., Rowan, K., Murphy, L., Openshaw, P.J.M., Semple, M.G., Law, A., Vitart, V., Wilson, J.F. & Baillie, J.K. Genetic mechanisms of critical illness in covid-19. *Nature* (2020).doi:10.1038/s41586-020-03065-y