Published Date: 18-Jul-2025

NOTE: These guiding principles will be updated periodically as needed. Please check the website regularly for updates or contact the Secretariat at <u>research@precise.cris.sg</u>

Our Commitment to Responsible Data Sharing

The PRECISE-SG100K resource is a unique multi-ancestry Asian population cohort dataset of ~100,000 Singaporean residents, jointly created by PRECISE, SG100K Institutions and GIS.

As joint stewards of this national resource, PRECISE and the SG100K PIs strive to maximise public good from the PRECISE-SG100K dataset while respecting participant consent and privacy, upholding scientific rigour, aligning with national priorities, and fulfilling commitments to funders and existing collaborators.

About the PRECISE-SG100K Resource

SG100K reached the milestone of recruiting 100,000 participants in Apr 2025 and whole genome sequencing of these participants was completed in Jun 2025. All data components are currently undergoing extensive quality checks and curation, both prior to *and* following linkage to electronic health records on the TRUST platform.

The first release of the integrated dataset for 100,000 participants is projected to be released in the first half of 2026.

The PRECISE-SG100K integrated dataset includes:

- **Research phenotype data** from four longitudinal population health cohorts ("SG100K Institutions") supported by the SG100K OF-LCG award:
  - <u>Health for Life in Singapore (HELIOS), NTU</u>
  - <u>Singapore Population Health Studies (SPHS), NUS</u>
  - <u>Singapore Epidemiology of Eye Diseases (SEED), SERI</u>
  - <u>SingHeart, NHCS</u>

- **Whole genome sequencing** (short read 30X coverage sequencing) generated in partnership with the Genome Institute of Singapore (GIS, A*STAR) under Phase II of Singapore's National Precision Medicine Programme.

Note: SG100K Institutions, GIS and PRECISE are collectively referred to as "PRECISE-SG100K".

The PRECISE-SG100K dataset is also linked to **electronic health records (EHRs)** that includes routinely collected clinical data and other health-related information. This integrated research dataset is accessible only through the TRUST platform with additional approval of the TRUST Data Access Committee.

Principles Governing Access to PRECISE-SG100K Resource

The guiding principles for accessing the PRECISE-SG100K resource is anchored on Singapore's Research, Innovation, Enterprise master plan. Our overarching goal is to advance precision medicine and related health research and innovation to improve health outcomes, in Singapore, Asia and the world. We are committed to the responsible and ethical sharing of data, grounded in principles of public benefit, participant trust, and scientific transparency.

To achieve these, we will avail de-identified data and research results through controlled data releases to qualified institutional researchers and appropriate research projects. This will be done in a manner that safeguards participant privacy, benefits Singapore, while upholding our ethical obligations, requirements of data guardians, and our commitments to funders and existing academic and industry collaborators.

Near-term

As the PRECISE-SG100K resource is being generated and curated, we are facilitating early data access for:

- **Investigator-Initiated Studies**. Funded by an NMRC OF-LCG grant that brought together four longitudinal cohorts in Singapore, the SG100K cohort is an Investigator-Initiated Study focused on determinants and prediction of cardiovascular disease and diabetes in Asian populations. To fulfil the grant award obligations and deliverables, the researchers from SG100K Institutions who have created the SG100K cohort have an initial period of priority to pursue research on these chronic conditions.

- **Call for Proposals.** To enable research across a wide range of use cases beyond cardiovascular disease and diabetes, early access to the intermediate data release of 50,000 participants by Singapore public academic and clinical researchers was enabled through an open Call for Proposals (application closed in January 2024). The PRECISE-SG100K Scientific Committee selected 36 projects, comprising 8 Flagship Projects and 28 Driver Projects.

- **Strategic Industry Partnerships**. Coordinated by PRECISE Business Development Team, strategic industry collaborations are being established to strengthen precision medicine research capacity and capability in Singapore.

Mid-term

We anticipate that the PRECISE-SG100K dataset will become broadly open to Singapore public academic and clinical researchers directly through the TRUST platform from 2027.

Access by overseas researchers or industry partners will remain contingent on the application being led by a Singapore research team.

All applications that request use of linked electronic health records will be subject to additional review by TRUST DAC.

Early Access Mechanisms (2024 – 2026)

- **Access by public researchers in Singapore**: PRECISE-SG100K will launch a 2nd Call for Proposals in first half of 2026 to provide early access to the PRECISE-SG100K resource of 100,000 participants dataset. This 2nd Call for Proposals is open only to Singapore public academic and clinical researchers.

  Please check the PRECISE website and newsletters regularly for updates on the 2nd Call for Proposals. Until then, PRECISE-SG100K is not actively accepting new research proposals.

- **Requests from strategic national programmes**: PRECISE-SG100K recognises that research proposals and data access requests from strategic national programmes will arise periodically. We welcome these requests for review.

  Please direct enquiries to research@precise.cris.sg.

- **Requests involving overseas collaborators or industry partners**: During the early access period (2024 – 2026), research proposals and data access requests involving overseas collaborators or industry partners are reviewed by PRECISE-SG100K on a case-by-case basis, with the expectation that the overseas collaborator or industry partner must demonstrate return of tangible benefits to Singapore.

  Please direct enquiries to research@precise.cris.sg.

Data Access Governance

- The PRECISE Management Team, which includes both the Executive Director of PRECISE and the Lead PI of the SG100K study, performs an initial evaluation of research proposals and data access requests from: i) strategic national programmes and ii) overseas academic collaborations. Where necessary, the Management Team consults relevant ecosystem stakeholders before recommending proposals for review by the PRECISE-SG100K Scientific Committee and PRECISE-SG100K Data Access Committee.

- The PRECISE-SG100K Scientific Committee reviews and endorses research proposals based on strategic and/or scientific merit. The Scientific Committee

may inform applicants of any competing or overlapping research interest with existing approved projects; these may include overlaps with SG100K cohort's research on cardiovascular disease and diabetes.

- The <u>PRECISE-SG100K Data Access Committee (DAC)</u> reviews and approves data access requests to the PRECISE-SG100K resource that <u>do not</u> require linkage to electronic health records (EHR)* to ensure they are in line with the permitted use of the dataset and are intended for public good.

Access to linked EHR data on TRUST will require additional approval by the <u>TRUST DAC</u> which assess the projects based on public interest and benefit to Singapore.

- The <u>PRECISE Business Development Team</u> reviews research proposals and data access requests with industry involvement or collaboration, in consultation with relevant ecosystem stakeholders if needed, before they are submitted for review by the PRECISE-SG100K Scientific Committee and PRECISE-SG100K Data Access Committee.

- The <u>PRECISE Chief Scientific Officer's Office</u> (CSOO) serves as Secretariat for data access, and is the initial primary point of contact for data availability and collaboration enquiries. Please contact <u>research@precise.cris.sg</u>.

Please see illustration for PRECISE-SG100K Data Access Governance process below.

Matters Pertaining to Publication Policy and Intellectual Property

All publications arising from the current early access route must be reviewed by the PRECISE-SG100K Data Access Committee and PRECISE-SG100K Scientific Committee prior to submission to confirm that the data use and publication objectives are within the scope of the approved data request. Projects involving use of TRUST Data will also require pre-publication review by TRUST.

Researchers shall follow the authorship and acknowledgement instructions listed in the Data Access Agreement.

PRECISE-SG100K will not claim rights to inventions and associated Intellectual Property Rights (IPRs) developed by researchers using the PRECISE-SG100K resource, unless the team and/or its members had made inventive contribution to the development of the IPs. Any resulting IPRs must not unreasonably restrict access to the dataset for academic research and healthcare purposes.

Exclusions

Biological samples collected from SG100K are under the custodianship of the respective population cohorts. Access to biological samples require approval from the individual population cohorts. Given that samples are a finite resource, SG100K prioritises:

- Assays conducted on all participants or a large subset of randomly selected participants
- Validated assay methodologies with minimal depletion and maximal output.

Newly generated data must be returned in full to the SG100K cohort(s) contributing the samples, along with a sublicensable, non-exclusive license for the SG100K cohort(s) to use the data in future research. Any additional requirements around sample access must be negotiated and agreed with the SG100K cohort(s) prior to sample release. PRECISE is pleased to facilitate introductions and discussions for access to biological samples from more than one cohort. For access to biological samples from a single cohort, applicants should approach the individual cohorts directly.

Similarly, access to other SG100K cohort resources, including various additional datasets, generation of new phenotypic information and participant recontact, is under the purview of the individual SG100K cohorts.

Please contact research@precise.cris.sg for more information.

While PRECISE-SG100K does not currently support linkages to other cohorts or datasets, we welcome suggestions for future data linkage opportunities.

Additional research phenotype or molecular data will be generated with various (new) funding sources. These data will be made accessible to the broader research community after fulfilling exclusivity (priority) periods or access requirements jointly established by the funders and PRECISE-SG100K.

Disclaimer

PRECISE-SG100K recognises that it is impossible for these guiding principles to cover all scenarios and reserves the right to update them from time-to-time.

These guiding principles are not intended to be cover all details related to data access. For questions or more information, please email the Secretariat for data access policy and operations at research@precise.cris.sg.

PRECISE-SG100K Data Access Governance process