2025 EECMS Summer Internship Application Form

Main Supervisor	Tudor Groza, Rare Care Centre / Duke-NUS PRISM
Other supervisors (if applicable)	A/Prof Saumya Shekhar Jamuar, Duke-NUS PRISM, SingHealth Prof Gareth Baynam, Rare Care Centre
Project Title	Automating rare disease genetic variant curation and management for international data sharing
Student location(s) for the project	KK Hospital for Women and Children, Singapore
Duration of project (ideally six weeks)	6 weeks
Project Description	Rare diseases, though individually uncommon, collectively impact millions of individuals worldwide. The genetic underpinnings of these diseases frequently involve rare variants that require extensive curation. The existing database on the Undiagnosed Diseases Program (UDP), Singapore, while being comprehensive, currently lacks automation capabilities, which leads to extensive manual labour. The aim of this project is to increase the efficiency of curating and managing the UDP patient data by developing a user-friendly, semi-automated platform. The project will focus on the collection and standardization of both genomic and phenotypic data. Genomic data - and clinical test results - will be sourced from the EHR and the local testing laboratory. Standardization of this data is essential for ensuring interoperability and reliable interpretation across different platforms. Similarly, phenotypic data will be collected from the clinical team. Detailed phenotypic descriptions are vital as they provide context to the genomic data and support an accurate diagnosis. The team uses the Human Phenotype Ontology to ensure that phenotypic information is consistently categorized and described. To enhance the efficiency and accuracy of the curation process, the project will integrate external tools, currently employed in a manual manner by the team. These include multiple variant annotation endpoints and standardised curation guidelines, which aggregated, enable the team to flag significant genetic variants, suggesting pathogenicity (disease causation), and linking relevant phenotypic data. The development of a UDP variant curation and management database that relies on automatic data integration and standardization will be an invaluable resource for the team, and in general for the rare disease community, by representing a step towards international data sharing, which in turn will advance the understanding of rare diseases, facilitate more accurate diagnoses.

Please email completed form to Tele Tan at t.tan@curtin.edu.au