Title: Applying a disease-specific annotation protocol for *VHL* gene curation using Hypothes.is

Authors/Affiliations:
Sarah Ridd,[1] Kirsten M. Farncombe,[2] Veronica Andric,[1] Safa Ansar,[1] Sean Delong,[1] Eric Li,[1] Samantha Macpherson,[1] Dena Salehipour,[1] Deborah I. Ritter,[3] Courtney Thaxton,[4] Raymond H. Kim[5]

[1]Department of Medicine, Division of Medical Oncology, University Health Network, Toronto, Ontario, Canada
[2]Toronto General Hospital Research Institute, University Health Network, Toronto, Ontario, Canada
[3]Department of Pediatrics, Baylor College of Medicine and Texas Children's Hospital, Houston, Texas, United States
[4]Department of Genetics, University of North Carolina, Chapel Hill, North Carolina, United States
[5]Division of Medical Oncology and Hematology, Princess Margaret Cancer Centre, University Health Network, Sinai Health System, Division of Clinical and Metabolic Genetics, The Hospital for Sick Children, Ontario Institute for Cancer Research, Department of Medicine, University of Toronto, Toronto, Ontario, Canada

Introduction: Von Hippel-Lindau (VHL) disease is a rare, inherited disorder that predisposes individuals to develop tumors in various organs throughout the body. Variants of uncertain significance in the *VHL* gene continue to be encountered despite the Clinical Genome Resource VHL Variant Curation Expert Panel (VCEP)'s best efforts to efficiently classify them for clinical decision making. To address the challenges associated with the lack of variant and clinical phenotype data-sharing among clinicians and scientists, the VHL VCEP is developing gene-specific conditions for *VHL* variants with the aim of improving patient care. We have adapted ClinGen's current standard operating procedure (SOP) for variant curation to account for specific characteristics and disharmony observed within VHL.
Methods: Hypothes.is is an online platform which presents a centralized hub to house aggregate information on *VHL* variants, systematically highlighting the variants found in the clinical literature and emphasizing American College of Medical Genetics (ACMG) standards. Our protocol standardizes legacy *VHL* variants, and highlights patient-specific information including age, family information and inheritance, disease assertion and phenotypic associations, and population database frequencies. We have screened for 435 unique patient-specific *VHL* genotype-phenotype publications dated up to September 2019.
Results: Our team has currently used the Hypothes.is platform to annotate 422 of the 435 papers. We have identified 584 unique variants in *VHL* with accompanying genotype-phenotype information. Of these, only 45% (263) were associated with ClinVarIDs, while 46% (269) had at most a CAID, and 9% (52) were not found on either of these notable databases, yet still determined to be relevant by our team.
Discussion: The considerable variability in *VHL* variant access across platforms poses a challenge to researchers and clinicians. We successfully established an accessible hub of genotype-phenotype information for VHL disease at a patient-specific level through our adaptation of a VHL-specific Hypothes.is annotation protocol. This standardized approach to annotating *VHL* variants has been a successful tool integrated into the VHL VCEP and their efforts towards curation and pathogenicity classification. Using Hypothes.is, we strive to mitigate the challenges imposed on VHL treatment by bringing together the strew of information on potentially causative *VHL* variants found in patients.