## Appendix. CIHI’s Population Grouping Methodology

CIHI’s Population Grouping Methodology 1.1 is the second edition of their Canadian health risk predictive modelling platform. Following efforts in the US (Medicare in the late 1990s through early-2000s)1 and Europe (the Netherlands and Germany in the mid-late 2000s)2 to produce models suitable for risk-adjusting payments based on morbidity, the CIHI model uses patient-level administrative healthcare data to summarize clinical complexity of the population and model costs to the publicly-funded healthcare system. The Canadian model was developed using pooled data from three provinces (Ontario, Alberta and British Columbia) for fiscal years (FY) 2010 and 2011 (concurrent period; April 1, 2010-March 31, 2012) and FY 2012 (prospective period; April 1, 2012-March 31, 2013). Patient clinical profiles, built with diagnosis codes from all available encounter types, were used to explain concurrent period costs and predict costs in the prospective period.

The foundation of the model is the diagnosis grouper, which takes into account approximately 10,000 ICD-9 and 18,000 ICD-10-CA diagnosis codes along with DSM-IV codes and some elements from RAI-MDS 2.0© and RAI-MH©, recorded in every available setting, and groups them into 226 Health Conditions (HCs) using algorithms developed in consultation with physician experts. HCs are groupings of clinically similar diseases that encompass chronic and acute illnesses, disabilities, medical emergencies, signs and symptoms, and other health states, such as pregnancy. HCs are assigned based on diagnosis codes recorded on discharge abstracts for inpatient, outpatient and emergency department hospital visits; assessments from the RAI-MDS 2.0© and the RAI-MH©; and physician encounter data. The methodology include data quality checks for physician encounter data to help ensure health conditions are assigned for ‘diagnosed’ and not merely ‘suspected’ conditions. Grouping rules ensure that each diagnosis code links to a single health condtion. Patients may have multiple HCs attached to them. However, a set of “clinical override rules” is applied by which less serious (or redundant) HCs are recoded as null in the presence of a more serious HC that is considered to be part of the same disease process (e.g., if a patient has HCs for both seizure and epilepsy, only the epilepsy flag will remain after overrides are applied).

For CIHI model development, HCs were assigned using diagnosis data recorded for inpatient stays (including mental health hospitalizations), day surgery/procedures, ED visits, and physician encounters (fee-for-service and shadow billed claims). Costs were estimated by CIHI as the sum of payments recorded for inpatient stays (excluding mental health hospitalizations), day surgery, ED visits and physician care. In the case of shadow billed physician encounters, costs were imputed based on payments for the same fee codes billed under fee-for-service.

For health system users with at least one health condition, costs were modelled as a function of 226 HCs and 460 HC interaction terms using ordinary least squares regression. For health system users with no conditions and health system non-users, costs were modelled separately based on age and gender only. The coefficients from the final additive models comprise the model weights that may be applied to new datasets in order to create risk scores, which are predictions of patient cost relative to others in the study sample or population.

CIHI has reported that their population grouper model explained nearly half of the variance in the dependent variable using age, gender, health conditions and condition interactions in the concurrent period (R2=0.475) and nearly one-tenth in the prospective period (R2=0.094) for their three-province model validation sample (CIHI 2017). An independent study using more complete healthcare costs data for Ontario than was available to CIHI for model development found that the model performed somewhat better than the Johns Hopkins ACG® model that has been used in Ontario for many years.3

1. Pope GC, Ellis RP, Ash AS, et al. Diagnostic Cost Group Hierarchical Condition Category Models for Medicare Risk Adjustment: Final Report. Prepared for Health Care Financing Administration, December 21, 2000.
2. Buchner F, Goepffarth D, Wasem J. The new risk adjustment formula in Germany: Implementation and ﬁrst experiences. Health Policy 2013;109:253-62.
3. Cheng S, Austin P, Wodchis W, et al. Evaluation of Population Groupers. ICES Report, September 2016.