Approach to Model COVID-19 Severity Risk and Prevalence of Underlying Health Conditions

September 2020

Tim Dall, MS
Executive Director

Ritashree Chakrabarti, PhD
Senior Consultant

Frank Chen, PhD
Principal Consultant

Ryan Reynolds, MA
Senior Consultant

Kari Jones, PhD
Associate Director

IHS Markit
1300 Connecticut Ave., NW
Suite 800
Washington, DC 20036 USA

Study Objective

PFCD engaged IHS Markit Ltd. to provide national, state and county-level estimates of how many adults have chronic conditions or other characteristics that are recognized risk factors for COVID-19 severity. Furthermore, this study estimated the number of adults with these chronic conditions whose conditions are poorly managed and thus could put them at greater risk for COVID-19 severity. This document provides an overview of the approach, data, and assumptions used to generate the estimates.

Overview of Methodology

The approach to generate estimates on chronic conditions, disease management and COVID-19 risk severity consists of three major elements.

- A constructed population database contains demographic, socioeconomic, and health risk factor information for a representative sample of the population in a geographic area (county, state, national) in 2020.
- For health risk metrics—such as poor control status of underlying medical conditions, and select diseases—where geographic data is unavailable, prediction equations were developed that relate individuals' health status to their know and observed characteristics.
- Population categorization by COVID-19 risk severity.

A. Construction of the Population Database

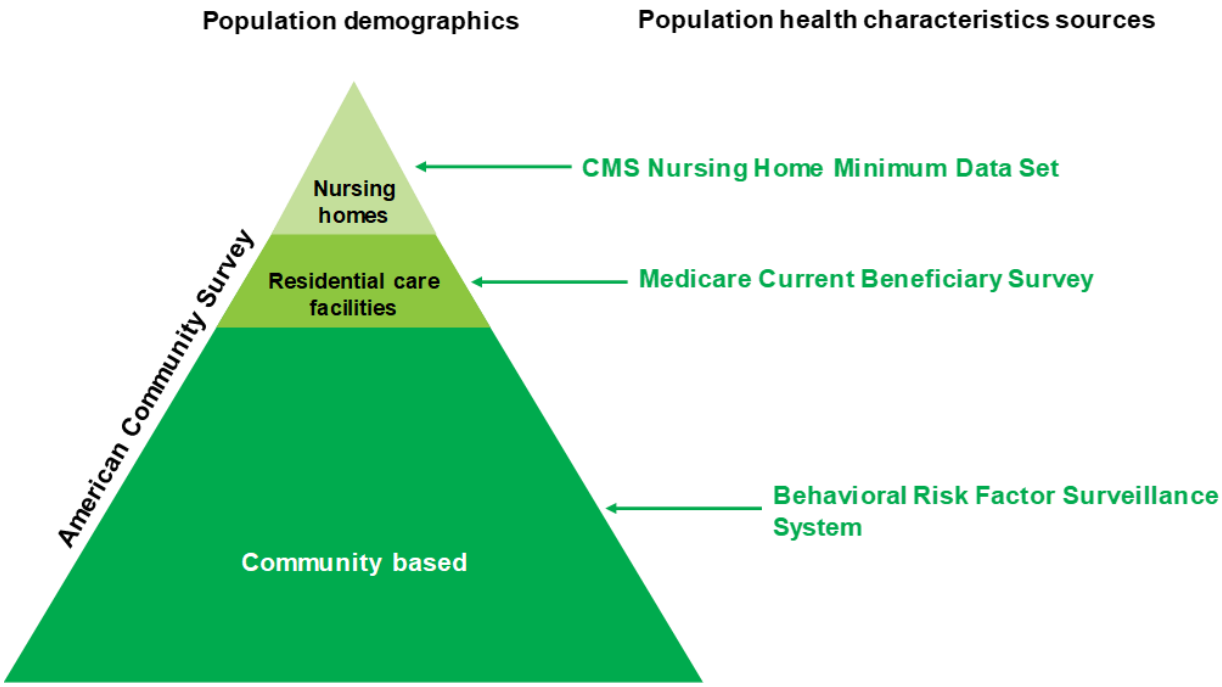
The county level estimates are based on a representative population database constructed for each of the 3,142 counties in the U.S. Construction of the county level population files starts with combining data from multiple sources, to create preliminary state population files that contain a representative sample of the population in each state by demographic, household income, medical insurance, and residency institution status (resides in community, in residential care facility, in nursing home). Then, the population data are re-calibrated to produce a representative sample of the population for each county in a state with the demographics and prevalence of health conditions (disease, lifestyle choices) benchmarked to external sources. More detail about the constructed population file is available elsewhere.[1,2]

The core micro data file is the 2018 American Community Survey (ACS). ACS provides demographic and socioeconomic characteristics of a representative sample of the population in each state, as well as information on medical insurance type, household income, and whether the person lives in a community or institutional setting. To add health risk factors and information on disease presence that is unavailable in ACS, we match each person in ACS with a similar person in the Behavioral Risk Factor Surveillance System (BRFSS), Medicare Beneficiary Survey (MCBS), or Centers for Medicare and Medicaid Services Long-Term Care Minimum Data Set (NHMDS).

The matching process uses an iterative statistical sampling algorithm. For the non-institutionalized population, each individual in the ACS file is matched with someone in the 2017 or 2018 BRFSS from the same sex, age group (15 age groups), race, ethnicity,

insurance type, household income level (8 income categories), and state of residence. Individuals categorized as residing in a residential care facility or nursing home are randomly matched to a person in the MCBS or NHMDS, respectively, in the same state, age group, sex, and race and ethnicity strata. The match probability for BRFSS and MCBS reflects the surveys' sample weights, with survey participants having higher sample weight more likely to be sampled. Exhibit 1: Population Database Mapping Algorithm show the various data sources and iterations used for creating the state representative population database.

Exhibit 1: Population Database Mapping Algorithm



The next step after creating the state population file is construction and calibration of the county level population files. The U.S. Census Bureau produces annual data on the total population in each county by five-year age bands, sex and race/ethnicity. We re-weight the sample weights for each metropolitan and non-metropolitan individual in a state's population file to match the demographics of the population characteristics in each metropolitan and non-metropolitan county, respectively, using the published Census Bureau population data. This produces a weighted sample that is representative of the demographics in each county. Further, county-level estimates of disease prevalence are calibrated at the individual level to match with external published information for each county. BRFSS data from state BRFSS surveys are the primary source for external county-level statistics used for calibrating prevalence of diseases and risk factors in the population files. Using states' projections of county-level population change, the population files are re-weighted to the 2020 population.

The resulting constructed population file contains a representative sample of the year 2020 population in each county by demographics, insurance type, prevalence of disease and health risk factors, with household income and residence type (community, residential care, or

nursing home) reflective of the counties' demographics. The population file includes data on health risk factors/conditions used this analysis: diabetes, hypertension, coronary heart disease, history of heart attack, history of stroke, asthma, arthritis, history of cancer, hypercholesterolemia, current smoker and obesity.

Although people with chronic obstructive pulmonary disease (COPD) and chronic kidney disease (CKD) are considered to be at an increased risk of severe illness from COVID-19, these conditions are not captured in the constructed population files.[3] We therefore used predictive equations (Exhibit 2) to impute the probability of having those conditions based on logistic regression analysis of the 2017 and 2018 BRFSS files.

Exhibit 2: Odds Ratios of Equations to Impute COPD and CKD Probability

	Explanatory Variables	COPD	CKD
Race-Ethnicity	Hispanic	0.72**	1.19**
	Non-Hispanic Black	0.90**	1.19**
	Non-Hispanic Other race	1.00	1.22**
	Non-Hispanic White	1.00	1.00
	Male	0.94**	0.95**
Age	18-34 years	1.00	1.00
	35-44 years	1.73**	1.66**
	45-64 years	5.19**	3.09**
	65-74 years	9.38**	4.75**
	75+ years	12.93**	7.14**
Diagnosed With	Smoker	4.75**	1.14**
	Hypertension	NA	1.05**
	Diabetes	NA	3.30**
	Asthma	6.13**	NA
	Insured	0.97	1.10*
	Medicaid	1.91**	1.23**
Body Weight	Normal	1.00	1.00
	Overweight	0.91**	1.03
	Obese	1.38**	1.32**
	Metropolitan	0.92**	1.03

Notes: Logistic regression analysis using the 2017-2018 BRFSS; Statistically different from 1.00 at the 0.05 (*) or 0.01 (**) level. NA=not applicable. The comparison group is non-Hispanic White, female, age 18-34, insured, without the chronic conditions or risk factors indicated, normal body weight, residing in a nonmetropolitan area.

B. Estimating Poorly Controlled Status of Chronic Conditions

The constructed population database lacks information on county-level measures of undertreatment or poorly managed chronic conditions. Such information can help assess COVID-19 severity risk. Recently published research suggests that poor glycemic control is independently associated with greater risk for adverse outcomes among people infected with COVID-19.[4–7] Although studies show that hypertension is associated with greater risk for

COVID-19 related outcomes, the research to date does not indicate whether poor control of one's blood pressure independently affects COVID-19 severity risk.[8] Some research suggests that medications used to help address hypertension could contribute to improvement of clinical outcomes of COVID-19 patients with hypertension.[9,10] Likewise, while there is limited information on whether hypercholesterolemia increases COVID-19 severity risk[11], managing cholesterol is important because it is related to high CVD risk which is associated with poor outcomes from coronavirus infections.[12,13]

We developed prediction equations of poor control for the chronic conditions listed in Exhibit 3. The estimates of poor control of chronic conditions are not included as risk factors for categorizing the population by COVID-19 risk severity, given the early stages of research looking at the evidence of potential connections between poor disease control and risk of severe illness from COVID-19. However, the estimates help underscore concerns that populations with underlying health conditions often have high rates of poor control for these chronic conditions which could potentially exacerbate severity of complications.

The prediction equations, summarized in Exhibit 4, are based on logistic regression analysis of pooled (2011-2018) National Health and Nutrition Examination Survey (NHANES) data. NHANES is a comprehensive national database which collects detailed biometric information for each individual through survey and diagnostic tests and exams—including checking blood pressure, cholesterol levels, and blood glucose levels. Five regressions were estimated, one each for the population with diabetes, hypertension, hypercholesterolemia, cardiovascular disease, and asthma. Adults with multiple conditions were included in each regression that was applicable to their conditions. The dependent variable for each regression is whether the person's condition is poorly controlled (1=poor control, 0=otherwise) using the criteria in Exhibit 3.

Exhibit 3: NHANES Criteria for Identifying Poorly Managed Disease Status

Population Diagnosed With	Criteria for Poor Control	Proportion of Diagnosed Adults
Diabetes	Hemoglobin A1c ≥ 7[14]	54%
Hypertension	Systolic blood pressure $\geq$130, or Diastolic blood pressure ≥ 80[15]	32%
Hypercholesterolemia [a]	Total cholesterol ≥ 240 or LDL ≥ 160[16]	17%
Cardiovascular Disease	Above recommended blood pressure levels or cholesterol levels	26%
Asthma	Had an emergency care visit for asthma during the over past year	19%

Source: 2011-2018 NHANES. Note: [a] High cholesterol is not considered an independent risk factor for COVID-19 severity based on current CDC guidelines. However, managing cholesterol level is recommended for people with or at high risk for COVID-19 because it is related to high CVD risk which is associated with poor outcomes from coronavirus infections.

Odds ratios in Exhibit 4 that are greater than 1.0 indicate patient characteristics associated with higher odds of poor control relative to the reference population. For example,

controlling for other patient characteristics, minority populations have higher odds of uncontrolled hypertension, CVD, and asthma relative to the non-Hispanic White population.

Exhibit 4: Odds Ratios of Equations to Impute Probability Condition is Poorly Controlled

	Explanatory Variables	Diabetes	Hypertension	Cardiovascular disease	Asthma
Race-Ethnicity	Hispanic	1.07	1.37**	1.80**	1.95*
	Non-Hispanic Black	0.85	1.95**	1.90**	2.83**
	Non-Hispanic Other race	0.87	1.23	1.15	2.33*
	Non-Hispanic White	1.00	1.00	1.00	1.00
	Male	1.06	1.10	0.72*	1.00
Age	18-34 years	1.00	1.00	1.00	1.00
	35-44 years	2.70*	2.35**	4.53*	1.14
	45-64 years	2.09	3.21**	6.33*	0.82
	65-74 years	1.59	4.25**	5.64	0.74
	75+ years	1.39	6.83**	9.60**	0.75
	Smoker	0.96	1.01	1.13	1.43
Diagnosed with	Hypertension	0.85	NA	NA	1.00
	High cholesterol	0.75*	0.89	NA	0.74
	Coronary heart disease	1.06	1.06	NA	1.05
	History of heart attack	0.96	0.78	NA	1.48
	History of stroke	0.79	0.90	NA	1.07
	Diabetes	NA	0.93	0.69*	1.10
	Arthritis	0.69**	0.85	0.99	1.12
	Asthma	0.85	1.08	0.71	NA
	History of cancer	0.71	0.97	1.10	1.04
	Insured	0.77	0.57**	0.78	1.60
	Medicaid	1.02	0.99	0.78	0.91
Household Income	<$10,000	0.94	1.34	2.88**	2.52*
	$10,000 to <$15,000	1.44	1.27	1.45	1.03
	$15,000 to < $20,000	1.19	1.22	1.39	1.60
	$20,000 to < $25,000	1.41	0.93	1.30	1.12
	$25,000 to < $35,000	1.78*	1.14	1.40	1.48
	$35,000 to < $50,000	1.63*	1.12	0.84	1.74
	$50,000 to < $75,000	1.58	1.21	1.27	0.24*
	$75,000 or higher	1.00	1.00	1.00	1.00
Body Weight	Normal	1.00	1.00	1.00	1.00
	Overweight	1.12	0.74**	0.68	1.64
	Obese	1.10	0.73**	1.35	1.69

Notes: Logistic regression analysis using the 2011-2018 NHANES; Statistically different from 1.00 at the 0.05 (*) or 0.01 (**) level. NA=not applicable. The comparison group is non-Hispanic White, female, age 18-34,

insured, without the chronic conditions or risk factors indicated, normal body weight, with household income of $75,000 or higher.

The estimated regression coefficients are applied to the population files to generate estimates of uncontrolled conditions at the county level controlling for disparities in demographics, presence of relevant risk factors and geographic variation.

C. Categorizing Population by Risk

Although everyone is at risk for COVID-19, there are certain individuals who are at an elevated risk to develop complications from COVID-19 that could result in hospitalization or mortality. The Centers for Disease Control and Prevention (CDC) published guidelines to identify individuals who are at an increased risk for COVID-19 severity.[3] Generally, older adults and people with underlying medical conditions are the most vulnerable populations. A growing body of research has identified specific conditions (e.g., chronic kidney disease, COPD, obesity, serious heart conditions, diabetes) as increasing a person's risk of severe illness, regardless of age.

Based on the CDC recommendations[3] and the City Block COVID-19 risk rules-based model criteria[17] we used a combination of age and presence of certain chronic conditions to stratify the adult (age 18 or older) population into three risk categories:

- High risk: having two or more of the following risk factors including *age>=65, diabetes, hypertension, cardiovascular disease, asthma, obesity, smoker, COPD and chronic kidney disease.*
- Moderate risk: having 1 or more of the following risk factors including *age>=65, diabetes, hypertension, cardiovascular disease, asthma, obesity, smoker, COPD and chronic kidney disease.*
- Low risk- individuals not *meeting the criteria for high or moderate risk.*

The three categories are used to produce summary reports and maps at the national, state, and county levels to look at the characteristics and distribution of the population by COVID-19 risk. In particular, we generated national and state maps of the proportion of the adult population in each county who is at high risk for severe COVID-19 complications. The maps were generated with four categories for counties ranging from 10% to 70% in 15% increments. These ranges were developed for the maps in order to show the geographic variation at the national level, and to illustrate the variation across counties in states.

References

1. Dall TM, Reynolds RL, Chakrabarti R, Iacobucci W, Jones K. Health Workforce Model Documentation [online]. IHS Markit; 2020. Accessed at: https://cdn.ihs.com/www/pdf/1118/Health-Workforce-Microsimulation-Model.pdf.

2. Health Resources and Services Administration. Technical Documentation for Health Resources Services Administration's Health Workforce Simulation Model [online]. Rockville, Maryland: U.S. Department of Health and Human Services; 2020. Accessed at: https://bhw.hrsa.gov/sites/default/files/bhw/health-workforce-analysis/research/projections/technical-documentation-health-workforce-simulation-model.pdf. Accessed May 15, 2020.

3. Centers for Disease Control and Prevention. Coronavirus Disease 2019 (COVID-19): People with Certain Medical Conditions [online]. 2020. Accessed at: https://www.cdc.gov/coronavirus/2019-ncov/need-extra-precautions/people-with-medical-conditions.html?CDC_AA_refVal=https%3A%2F%2Fwww.cdc.gov%2Fcoronavirus%2F2019-ncov%2Fneed-extra-precautions%2Fgroups-at-higher-risk.html. Accessed August 31, 2020.

4. Bode B, Garrett V, Messler J, et al. Glycemic Characteristics and Clinical Outcomes of COVID-19 Patients Hospitalized in the United States. J Diabetes Sci Technol. 2020;14:813–821.

5. Brufsky A. Hyperglycemia, hydroxychloroquine, and the COVID-19 pandemic. J Med Virol. 2020;92:770–775.

6. Zhu L, She Z-G, Cheng X, et al. Association of Blood Glucose Control and Outcomes in Patients with COVID-19 and Pre-existing Type 2 Diabetes. Cell Metab. 2020;31:1068-1077.e3.

7. Wang S, Ma P, Zhang S, et al. Fasting blood glucose at admission is an independent predictor for 28-day mortality in patients with COVID-19 without previous diagnosis of diabetes: a multi-centre retrospective study. Diabetologia [online serial]. Epub 2020 Jul 10. Accessed at: http://link.springer.com/10.1007/s00125-020-05209-1. Accessed August 31, 2020.

8. Schiffrin EL, Flack JM, Ito S, Muntner P, Webb RC. Hypertension and COVID-19. American Journal of Hypertension. 2020;33:373–374.

9. Feigin VL, Abajobir AA, Abate KH, et al. Global, regional, and national burden of neurological disorders during 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015. The Lancet Neurology [online serial]. 2017;16:877–897. Accessed at: http://linkinghub.elsevier.com/retrieve/pii/S1474442217302995. Accessed June 15, 2018.

10. Tadic M, Cuspidi C, Mancia G, Dell'Oro R, Grassi G. COVID-19, hypertension and cardiovascular diseases: Should we change the therapy? Pharmacological Research. 2020;158:104906.

11. Zaki N, Alashwal H, Ibrahim S. Association of hypertension, diabetes, stroke, cancer, kidney disease, and high-cholesterol with COVID-19 disease severity and fatality: a systematic review [online]. Epidemiology; 2020 Jun. Accessed at: http://medrxiv.org/lookup/doi/10.1101/2020.06.16.20132639. Accessed August 31, 2020.

12. Radenkovic D, Chawla S, Pirro M, Sahebkar A, Banach M. Cholesterol in Relation to COVID-19: Should We Care about It? JCM. 2020;9:1909.

13. Banach M, Penson PE, Fras Z, et al. Brief recommendations on the management of adult patients with familial hypercholesterolemia during the COVID-19 pandemic. Pharmacological Research. 2020;158:104891.

14. American Diabetes Association. 6. Glycemic Targets: Standards of Medical Care in Diabetes-2020. Diabetes Care. 2020;43:S66–S76.

15. Whelton Paul K., Carey Robert M., Aronow Wilbert S., et al. 2017 ACC/AHA/AAPA/ABC/ACPM/AGS/APhA/ASH/ASPC/NMA/PCNA Guideline for the Prevention, Detection, Evaluation, and Management of High Blood Pressure in Adults: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. Hypertension. American Heart Association; 2018;71:e13–e115.

16. 2018 AHA/ACC/AACVPR/AAPA/ABC/ACPM/ADA/AGS/APhA/ASPC/NLA/PCNA Guideline on the Management of Blood Cholesterol: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines | Circulation [online]. Accessed at: https://www.ahajournals.org/doi/10.1161/CIR.0000000000000625. Accessed September 16, 2020.

17. Schnake-Mahl AS, Carty MG, Sierra G, Ajayi T. Identifying Patients with Increased Risk of Severe Covid-19 Complications: Building an Actionable Rules-Based Model for Care Teams. NEJM Catalyst Innovations in Care Delivery [online serial]. Massachusetts Medical Society; Epub 2020 May 4. Accessed at: https://catalyst.nejm.org/doi/abs/10.1056/CAT.20.0116. Accessed August 31, 2020.