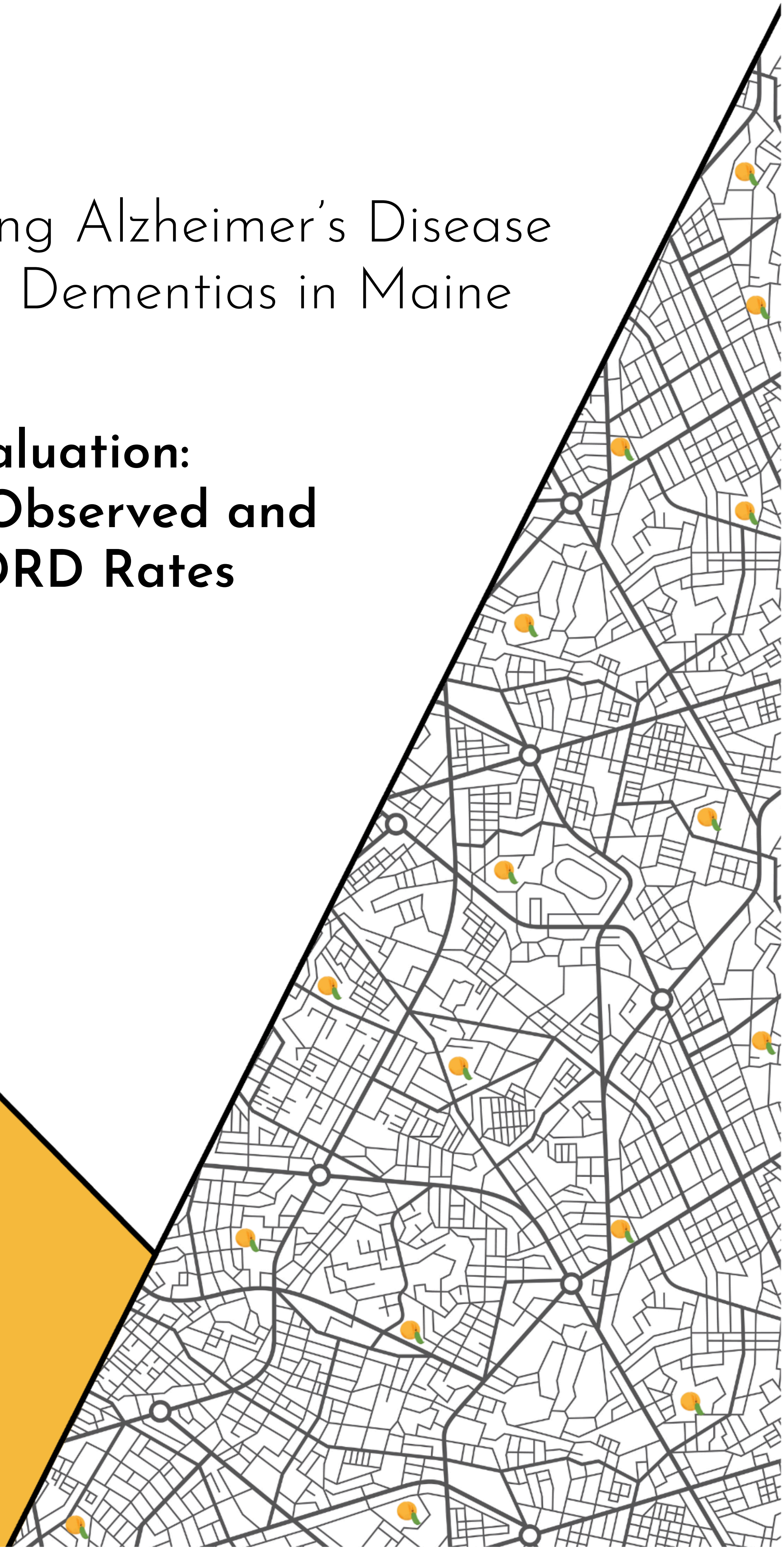
Apricot

Characterizing Alzheimer's Disease and Related Dementias in Maine

Estimate Evaluation: Comparing Observed and Expected ADRD Rates

July 2024



Introduction

This project investigated the potential utility of transformation and geographic downscaling of existing approaches to ADRD estimation in Maine. The working thesis of the exploration was that the general unavailability of ADRD estimates at the sub-county level may mask key insights and limit the potential impacts of public health ADRD surveillance, detection, and intervention efforts.

This report covers three key aspects of the work completed during this project. First, we summarize a demographic data analysis of Maine's 65+ population. Second, we outline the processes used to produce two expected (model-based) rates, one proxy measure, and one observed rate of ADRD in Maine. Third, we summarize findings from a comparative analysis of observed and expected ADRD rates at the state and county levels. Finally, we offer recommendations for next steps to refine and improve these ADRD estimates.

Key Findings

1. The proportion of Maine's population that is aged 65+ is a demographic outlier compared to the rest of the United States. Thus, understanding the demographic characterization of Maine's 65+ population can contribute useful context for Maine-specific ADRD estimation. Bringing publicly-available demographic data into a future model would enable more robust, Maine-specific estimates.

2. The improved performance of the Dhana (2023) model following transformation of the education variable highlights the potential value of model fine-tuning. This finding suggests that more substantial tinkering with model inputs – and specifically, aligning them with known demographic details about Maine's 65+ population – will likely yield stronger performance in a Maine-specific context.

3. The comparative analysis component of this work highlights the utility of the "observed/expected" framework commonly used in public health surveillance to identify geographic areas of interest – that is, areas that display unusually high or low observed rates when compared to estimates of expected rates.

4. Our work found significant heterogeneity in ADRD estimates at the sub-county level. The observed variation provides evidence in support of this investigation's working thesis—that county-level ADRD estimates may mask important geographic differences—and illustrates the potential utility of downscaling estimates to increasingly granular geographies.

1. Overarching Context: Characterizing Maine's 65+ Population

Before beginning modeling and estimation, we conducted a demographic analysis of Maine's 65+ population, as this is the primary age group impacted by ADRD. We contextualized this population by *zooming out* to compare Maine to New England the rest of the United States and by *zooming in* to explore variation within Maine. Using publicly-available ACS (2018-2022) and Census (2020) data, we analyzed sex, race, living arrangements, home ownership status, group quarters populations, and some two-way interactions of these variables (e.g. living alone by sex) for Maine's 65+ population.

Zooming Out:

Maine is often described as the "oldest state in the nation." Using ACS and Census data, we found that Maine (in 2020) had 297,029 residents aged 65+, which comprised 21.8% of Maine's total population. Only Puerto Rico's 65+ population (22.3%) makes up a larger share of its total population. Moreover, the population proportion of older Mainers continues to grow – 2022 postcensal updates estimated the number of 65+ residents rose by an additional 15,195 to 312,224, or 22.5% of the state population. Table 1 summarizes a few additional facets of Maine's 65+ population in a New England and United States context. A particularly noteworthy finding in the context of ADRD outreach efforts is that Maine leads the nation in the proportion of its total population that is 65+ and living alone.

Table 1: Zooming Out – Maine's 65+ Population in a NE, US Context

Contextualizing Maine's 65+ Population

Selected Demographic Characteristics of 65+ Populations in Maine, New England, and the US

	Maine	New England	US
Proportion of Population That Is 65+	21.8%	19.2%	17.4%
Proportion of Population That is 85+	2.4%	2.3%	1.9%
Proportion of 65+ Living in Group Quarters	4.0%	3.8%	3.3%
Proportion of 65+ That Is a Householder Living Alone	27.0%	27.0%	26.1%
Proportion of Total Poplation That Is 65+ and Living Alone	5.9%	5.2%	4.6%

Data Source: 2018-2022 ACS
Produced by Apriqot, Inc.

Zooming In:

When examining Maine's 65+ population at the county and sub-county levels, some demographic characteristics vary widely, but others show high degrees of similarity. For example, while residents aged 65+ make up 21.8% of the total state population, the range at the county-level is quite wide – from 18.6% (Androscoggin) to 29.5% (Lincoln). To highlight an example of cross-county similarity, the vast majority of the 65+ population is white across the state, ranging from 95.1% (Washington) to 97.2% (Aroostook). Table 2 summarizes county-level variation for a few selected demographic indicators.

Table 2: Zooming In – Demographics of Maine's 65+ Population, by County

County-Level Variation in Maine's 65+ Population

Selected Demographic Characteristics of 65+ Population in Maine, by County

County	Proportion of Population That Is 65+	Proportion of Population That Is 85+	Proportion of 65+ Population That Is White	Proportion of Households With at Least 1 Person 65+
Androscoggin	**18.6%**	2.5%	95.8%	**31.5%**
Aroostook	25.3%	**3.3%**	**97.2%**	40.0%
Cumberland	19.7%	2.4%	95.2%	32.4%
Franklin	23.6%	2.5%	96.4%	37.4%
Hancock	26.4%	2.7%	96.9%	40.6%
Kennebec	21.2%	2.4%	96.2%	33.0%
Knox	26.9%	3.0%	96.3%	42.0%
Lincoln	**29.5%**	3.0%	96.9%	**45.1%**
Oxford	22.8%	2.4%	96.4%	38.8%
Penobscot	19.7%	2.2%	96.1%	33.0%
Piscataquis	27.5%	2.7%	96.3%	43.1%
Sagadahoc	23.7%	2.6%	96.0%	36.7%
Somerset	22.1%	2.1%	96.3%	36.7%
Waldo	24.2%	**2.0%**	96.5%	40.1%
Washington	26.1%	2.7%	**95.1%**	41.3%
York	21.4%	2.2%	96.3%	36.1%

Data Source: 2018-2022 ACS
Produced by Apriqot, Inc.

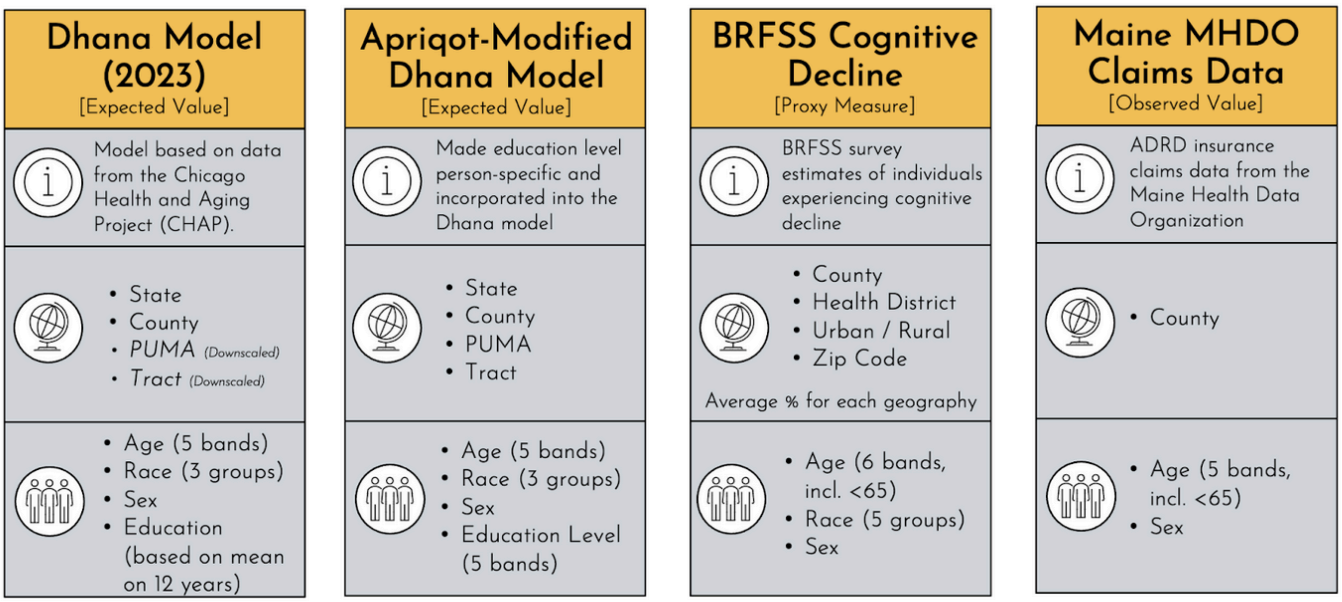
This observed pattern of homogeneity in some demographic characteristics and broad heterogeneity in others extends to the sub-county level. Identifying the aspects of the 65+ population that are highly variable at different geographic levels could inform a future Maine-specific ADRD model by illuminating which characteristics may warrant downscaling and which may provide only limited insight at a finer geographic scale.

As we discuss in the next section, a major shortcoming of existing ADRD models is that they are based on populations that do not represent the Maine population to which they are applied to generate estimates. Our demographic analysis indicates that a key step for producing highly-relevant and actionable ADRD estimates is incorporating as much detail as possible about the unique features of Maine's 65+ population. While layering this demographic information onto the models is outside the scope of the current project, bringing Maine demographics into an ADRD prediction model is a potential next step.

2. Approach: Developing Expected and Observed Estimates

The next phase of the project focused on developing four ADRD-related estimates: two sets of model-based estimates, one proxy measure, and one observed ADRD rate. Figure 1 below summarizes the sources for these four sets of estimates.

Figure 1: Summary of Four ADRD Models and Estimates



Model-Based Expected Value #1: Dhana Model (2023)

Using data from the Chicago Health and Aging Project (CHAP), Dhana and colleagues (2023) produced nationwide ADRD estimates at the state and county levels.[1] Their model uses age, race, sex, and education as input variables (see Table 3).

Table 3: Input Variables for Dhana Model

Age	Race	Sex	Education
• 65-69 • 70-74 • 75-79 • 80-84 • 85+	• White • Black or African American • Hispanic	• Male • Female	• 1 standard deviation (SD) difference from average years of education (12.3) of Census respondents

[1] Dhana K, Beck T, Desai P, Wilson RS, Evans DA, Rajan KB. Prevalence of Alzheimer's disease dementia in the 50 US states and 3142 counties: A population estimate using the 2020 bridged-race postcensal from the National Center for Health Statistics. Alzheimer's & Dementia. 2023 Oct;19(10):4388-95.

Apriqot replicated the methods used by Dhana and colleagues and matched the model coefficients. After reproducing the model, Apriqot downscaled it to produce estimates at the PUMA and census tract levels by applying the model to ACS 2018-2022 microdata.

Model-Based Expected Value #2: Apriqot-Modified Dhana Model
One clear and easily addressable shortcoming of the Dhana model is the treatment of education as an input variable. In the original Dhana model, the education coefficient, measured in standard deviation (SD) differences from the mean years of education, was highly significant. Despite this, they produced state- and county-level estimates using the mean years of education reported in the 1990 Census—12.3 years. As education is an important predictor of ADRD, we wanted to correctly incorporate it in the model by applying more accurate estimates of education levels at the individual level.

To that end, we modified the education variable to include person-specific education levels – <9 years, 9-12 years, 12-14 years, 14-16 years, and 16-19+ years. We then noted the resulting education coefficient and used it to replace the original operationalization of education in the existing Dhana model. We used the Apriqot-Modified Dhana Model to produce ADRD estimates at the state, county, PUMA, and tract levels.

Proxy Measure Expected Value: BRFSS Cognitive Decline
In addition to the two Dhana models, we also constructed an ADRD proxy measure using the Behavioral Risk Factor Surveillance System (BRFSS).[2] We developed the measure using responses to the cognitive decline optional module currently administered every two years in Maine.[3] Specifically, we used a question in the 2020 and 2022 cognitive decline modules: "During the past twelve months, have you experienced confusion or memory loss that is happening more often or is getting worse?" We used this question as a proxy measure estimate for ADRD due to its subjective nature and lack of clear linearity between a positive response and a diagnosis of ADRD. However, one advantage of the BRFSS data is its geographic disaggregation at the county, Maine health district, urban/rural, and zip code levels. Accordingly, for each geographic area, we computed direct estimates of the percentage of people experiencing cognitive decline by dividing the weighted sum of "yes" responses by the weighted sum of those answering the question, with the survey weight BRFSS population control totals taken from Intercensal population estimates and the American Community Survey.

---

[2] Olivari BS, Baumgart M, Taylor CA, McGuire LC. Population measures of subjective cognitive decline: A means of advancing public health policy to address cognitive health. Alzheimer's & Dementia: Translational Research & Clinical Interventions. 2021;7(1):e12142.

[3] Lee SH, Moore L. BRFSS statistical brief: Cognitive decline optional module.

Observed Value: MHDO Claims Data

We used Maine Health Data Organization (MHDO) claims data as "observed" ADRD rates at the county level, as this was the best dataset option available to us during this investigation. MHDO constructs this dataset using insurance claims from the All-Payer Claims Data medical claims database. Variables included in the county-level dataset include age, gender, insurance type, disability status, and year.

To prepare the observed ADRD rates we computed the number of individuals with claims by dividing the number of individuals with claims made in 2020 by the total county population over 65 years. Figures for total county population came from the 2020 National Center for Health Statistics (NCHS) bridged race file, which is the same file used for race estimation in the Dhana model. This produced individual ADRD claims rates for each Maine county as well as a statewide figure, providing the observed measures needed to conduct a comparative analysis of observed, proxy, and expected ADRD rates at both the state and county levels.

One caveat is that claims-based rates may not precisely estimate underlying population prevalence of chronic conditions. This imprecision can result from underdiagnosis of the condition, differential latency in the timing of the diagnosis, and data quality issues for demographic and geographic information in administrative data. For ADRD specifically, underdiagnosis is a known issue that likely impacts the quality of the MHDO claims data.[4]

[4] Alzheimer's Association. Alzheimer's Disease Facts and Figures. 2024. Available from: https://www.alz.org/media/documents/alzheimers-facts-and-figures.pdf

3. Comparative Analysis of ADRD Estimates

We conducted comparative analysis to evaluate the relative performance of the two Dhana-based models and the BRFSS proxy measure against the observed ADRD rate captured in the MHDO claims data.

Table 4: Statewide ADRD Estimates

Comparing Statewide Estimates of ADRD

Estimates of ADRD Rates, By Sex and Age Group

Age Span	Estimates			
	Apriqot-Modified Dhana Model	Original Dhana Model	BRFSS Cognitive Decline	MHDO Claims Data
Male				
45-59	NA[1]	NA[1]	10.2	0.2
60-64	NA[1]	NA[1]	11.2	0.6
65-74	3.5	4.1	8.8	1.6
75-79	7.7	8.9	10.8	5.9
80+	21.2	23.5	16.4	13.9
Female				
45-59	NA[1]	NA[1]	9.7	0.2
60-64	NA[1]	NA[1]	6.9	0.5
65-74	3.9	4.6	9	1.9
75-79	9	10	6.7	6.6
80+	25.4	26.9	12.9	19.6

[1] Age range not available in input data.

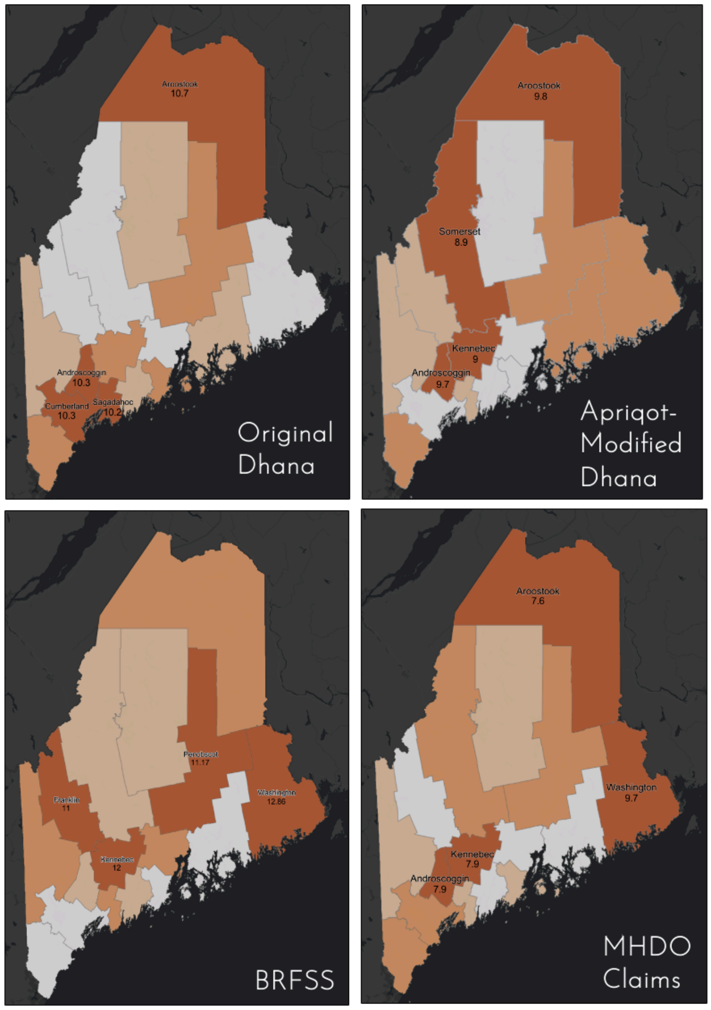
Produced by Apriqot, Inc.

Table 4 summarizes statewide "expected" (Apriqot-Modified Dhana, Dhana, BRFSS Cognitive Decline) and "observed" (MHDO Claims Data) ADRD estimates by sex and age span. For both sexes, the Dhana models correspond with the pattern of increase seen in the MHDO Claims data as age increases, with the overall-lower estimates produced by the Apriqot-modified Dhana model being slightly more aligned than the original Dhana model. Unlike the Dhana models, the BRFSS proxy measure provides estimates for younger individuals – but these predicted rates are significantly higher than those observed in the MHDO claims data and it seems unlikely that ADRD would be more prevalent amongst adults 45-64 than adults 65-74. Another oddity in the BRFSS estimates is that they predict lower incidences of ADRD for females aged 75 and over, while both Dhana models and the claims data predict that older women would have higher ADRD rates than men.

In summary, this phase of the comparative analysis indicates that the two Dhana models outperform the BRFSS proxy measure when predicting statewide ADRD rates by sex and these five age categories.

Turning to the county level, Figure 2 presents quartile maps of the ADRD estimates.

Figure 2: Maps of County-Level ADRD Estimates



The maps in Figure 2 display predicted ADRD rates by quartiles. That is, the four counties with the lowest predicted rates of ADRD are shaded in the white, the next four in the light peach shade, and so on until the four counties with the highest predicted rates, which are highlighted in dark orange. Looking at the maps in Figure 2, there are areas of convergence between many of the estimates in Androscoggin, Aroostook, and Kennebec counties – areas where ADRD rates are predicted to be high – and in Piscatiquis and Waldo counties, where ADRD rates are predicted to be lower. There are also areas of significant divergence in terms of quartile prediction, particularly Somerset, Washington, and Cumberland counties.

Table 5 presents the comparative analysis for estimates at the county level. The yellow boxes indicate the county predicted by each model to have the highest rate of ADRD, while the blue boxes highlight the counties estimated to have the lowest ADRD rates. In terms of highest expected rates of ADRD, both Dhana models predict Aroostook County (10.7% and 9.8%), while the BRFSS estimated Washington County (12.9%), which does have the highest individual claims rate in the MHDO data (9.7%). When looking at lowest anticipated ADRD rates, the BRFSS predicts Knox County (8%), while both Dhana models estimate Waldo County (9.1% and 7.5%), which did have the lowest individual claims rate in the MHDO data (4.8%).

Table 5: County-Level ADRD Estimates

Comparing County-Level Estimates of ADRD

Estimates of ADRD Rates

County	Estimates: Original Dhana Model	Estimates: Apriqot-Modified Dhana Model	Estimates: BRFSS Cognitive Decline	Estimates: MHDO Claims Data
Androscoggin	10.1	9.7	9.4	7.9
Aroostook	10.7	9.8	10.2	7.6
Cumberland	10.3	8.1	8.7	7.3
Franklin	9.6	8.3	11	5.7
Hancock	9.8	8.4	9.2	5.1
Kennebec	10.1	9	12	7.9
Knox	10.1	8.1	8	6.2
Lincoln	10	7.8	9.7	5.7
Oxford	9.9	8.2	10.9	6.3
Penobscot	10.1	8.4	11.2	6.9
Piscataquis	9.7	8	9.8	6.3
Sagadahoc	10.2	8.2	10.2	6.4
Somerset	9.5	8.9	9.7	6.9
Waldo	9.1	7.5	10.8	4.8
Washington	9.6	8.5	12.9	9.7
York	10.1	8.4	8.4	6.4

Produced by Apriqot, Inc.

To better understand the alignment between the expected and observed ADRD rates, we computed both Pearson and Spearman's Rank correlation coefficients. In terms of the Pearson correlations (Table 6), the Apriqot-modified Dhana model (0.585) correlates more highly with the MHDO claims data than the BRFSS Cognitive Decline measure (0.430) and the Original Dhana model (0.324). When looking at Spearman's Rank (Table 7), which is computed based on estimated rank order of ADRD rates in the counties, the Apriqot-modified Dhana model (0.713) again outperforms both the Original Dhana model (0.427) and the BRFSS Cognitive Decline proxy measure (0.254).

Table 6: Pearson Correlation Coefficients

County-Level Pearson Correlation Comparison

	Original Dhana	Apriqot-Modified Dhana	BRFSS Cognitive Decline	MHDO Claims Data
Original Dhana	1.000	0.505	−0.295	0.324
Apriqot-Modified Dhana	0.505	1.000	0.076	0.585
BRFSS Cognitive Decline	−0.295	0.076	1.000	0.430
MHDO Claims Data	0.324	0.585	0.430	1.000

Produced by Apriqot, Inc.

Table 7: Spearman's Rank Correlation Coefficients

County-Level Spearman's Rank Correlation Comparison

	Original Dhana	Apriqot-Modified Dhana	BRFSS Cognitive Decline	MHDO Claims Data
Original Dhana	1.000	0.242	−0.291	0.427
Apriqot-Modified Dhana	0.242	1.000	0.201	0.713
BRFSS Cognitive Decline	−0.291	0.201	1.000	0.254
MHDO Claims Data	0.427	0.713	0.254	1.000

Produced by Apriqot, Inc.

There are likely both model-based and claims-data based contributions to the error captured in these correlations. While model-based errors are to be expected due the nature of modeling, the claims data underlying the observed rates likely undercounts ADRD cases, as figures are based solely on insurance claims from Maine residents who sought medical care and received a codable ADRD diagnosis. Despite these shortcomings, the stronger performance of the Apriqot-modified Dhana model when compared to the original Dhana model illustrates the beneficial impact of transforming the education input variable and the potential value of model fine-tuning for estimate improvement.

Additionally, the comparative analysis conducted in this project highlights the potential utility of using an “observed/expected” framework, which is commonly used in public health surveillance, to contextualize estimates. When comparing “expected” values to “observed” values, we can identify areas of interest – that is, areas that display unusually high or low observed rates as compared to the expected estimates. Higher than expected values can occur when important risk factors are not included in models, while lower than expected values can identify areas of underdiagnosis. While improvements to both the underlying model and the observed ADRD cases data will further increase the utility of the “observed/expected” framework, the comparative analysis portion of this project highlights that such an approach can help to contextualize published estimates.

Action Steps & Recommendations

Moving forward, there are several steps that could extend this investigation.

1. **Access to Observed Data at the Sub-County Level**
   While the findings of the comparative analysis provide supportive evidence for the value of model fine-tuning to produce improved estimates at the county-level, the absence of sub-county claims data makes it difficult to assess and validate the additional value of downscaling the models to sub-county geographic levels. While the downscaled models suggest significant heterogeneity in ADRD estimates at the census tract level, evaluating the accuracy of the downscaled models using the same comparative analysis procedures conducted at the state and county levels isn’t possible with the current county-level MHDO claims data.

2. **Obtaining Better Age Coverage in MHDO Claims Data**
   The MHDO data summarizes claims in several age bands – 0-44, 45-59, 60-64, 65-74, 75-79, and 80+. Given that ADRD risk highly correlates with aging, it is unfortunate that 65-69 and 70-74 are not parsed as separate age ranges. Given that the models in this investigation predict large increases in ADRD rates between

the 65-74 and 75+ age bands, it would be useful to better understand the nature of the increase taking place within the 65-74 year age group, as these may be particularly critical ages for ADRD interventions.

3. **Producing Downscaled Estimates for Specific Demographic Subpopulations**
   In this investigation, we produced statewide ADRD rates by age and sex. These disaggregated estimates could be downscaled to the county- and sub-county levels. Additionally, race and education could be further incorporated to produce estimates for more specific demographic subpopulations.

4. **Incorporating Maine-Relevant Demographic, SDOH, and Health Risk Factors**
   We believe the strongest potential source of improvement to these initial ADRD models is the incorporation of Maine-specific demographic, health risk factor, and Social Determinants of Health variables. By combining the insights about the Maine population that can be gleaned from publicly-available data with Maine-specific rates of known ADRD risk factors and the ADRD prediction model, we could produce more contextualized, downscaled ADRD estimates.

5. **Developing a Model-Based Estimate for BRFSS Cognitive Decline**
   In the present investigation, the BRFSS cognitive decline estimates are direct estimates. A possible next step to exploring the possible utility of BRFSS as a proxy measure for ADRD is generating a model-based estimate based on responses to the cognitive decline module and comparing the results to the Dhana models and the current observed rates derived from the MHDO claims data.

6. **Identifying Maine Communities with Elevated ADRD Rates or High Potential Rates of Under-Diagnosis**.
   The ultimate goal of downscaling existing ADRD models to finer levels of geography is to provide insight into community-level ADRD rates, which allows for the development of highly-localized surveillance and intervention strategies. Completing any of the action steps outlined above would increase the actionability of the results produced by this approach to ADRD estimation, allowing for more robust identification of communities where there is notable divergence between expected and observed ADRD rates.

In summary, while this project provides initial evidence of the value of demographic downscaling in ADRD estimation at the sub-county level, we believe that there is more work that can be done to further refine these approaches and provide richer, Maine-specific ADRD estimates.