Ohio Opportunity Index Update
June 2024
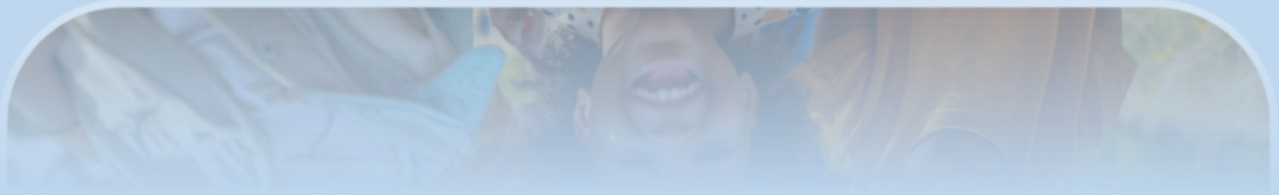
Ohio Opportunity Index: Version 3

Introduction

The Ohio Opportunity Index (OOI) is a composite, area-based measure of neighborhood socioeconomic, structural, and environmental conditions representing the social determinants of health. Its purpose could range from supporting population health focused agencies, organizations, and individuals with activities that include targeting assessments, interventions, and evaluations, facilitating research, and informing policy decisions by offering a comprehensive geographic view of the social determinants of health at a census tract level.

This brief report documents an update of the OOI to the third version (Version 3 or V3, developed in 2024) as part of an ongoing update cycle in which the OOI is updated every other year. While this report does indicate numerous basic components or features of the OOI, it largely focuses on high-level details of the methods and what has changed from V2 to V3 of the OOI. The sections below include the procedural steps of constructing the index, some data output that results from the procedures for the purpose of verifying and validating the product, and a brief closing discussion.

Procedures Constituting the Construction of the Ohio Opportunity Index

Aside from the domain and measure differences between the OOI and the Children's Opportunity Index (COI), the following set of procedural steps are identical to those used to construct the COI Version 2, which included several methodological changes that were described in the report packaged with the COI V2 data files. Those changes are represented in the documentation of procedures below but are not called out as new.

Step 1: Select Domains of Opportunity

The following domains were carried over from the OOI Version 2 to compose the multi-dimensional construct reflecting socioeconomic, structural, and environmental factors in Ohio: 1) Transportation, 2) Education, 3) Employment, 4) Housing, 5) Health, 6) Environment, and 7) Crime. The research team believes these domains still comprehensively cover area-level social determinants of health.

Step 2: Select or Refine Indicators to Measure Each Domain

The indicator assignments in each domain are described below. We highlight indicators that were carried over from OOI Version 2 as well as new or removed indicators for OOI Version 3. Those that were carried over were updated with the most current data available. In most cases the most current data is from either 2022 or 2023. Table 1

summarizes the commonalities and differences between versions 2 and 3 of the OOI in terms of measure inclusion and domain assignment.

Transportation: The OOI Version 2 transportation domain contained four variables in OOI Version 2. The research team carried forward all four variable selections: access to public transit, average commute to work time, households without vehicle access, and traffic proximity. A change to the 2024 public transit access indicator was the inclusion of taxi use.

Education: The education domain contained five variables. We carried forward four variables: educational attainment, average school performance, average free and reduced lunch participation rate, and high school graduation rate. A change to the variable 'residential internet connectivity availability' was that we replaced it with a new dataset representing internet infrastructure (as opposed to purchased internet subscription in V2) with a theoretical connection speed threshold of at least 300 Mbps. The threshold was chosen due to its ability to distinguish different parts of the state in visual geographic analysis.

Employment: The employment domain contained four variables. We carried forward four variables: low-wage job access (total entry-level jobs divided by total people with high-school or less education), access to workforce or job training sites, unemployment, and poverty. Aside from updating to the most current data, no changes were made to the indicators included in Version 3.

Housing: The housing domain contained six variables. We carried forward six variables: median rent, median home value, concentration of existing Low-Income Housing Tax Credit (LIHTC) units, housing stock built pre-1980s, residential overcrowding, and residential mobility. Aside from updating to the most current data, no changes were made to the indicators included in Version 3.

Health: The health domain contained five variables. The OOI V3 carried forward five variables: age-adjusted mortality rate among Medicaid members, preventable ED visits, diabetes hospital admissions, geographic access to healthy foods, and access to medical providers. The OOI V3 included one change. The diabetes hospitals admissions variable was modified so the denominator now represents Medicaid members with a diabetes diagnosis any time in the period covered by the data. It thus represents diabetes-related hospital admissions among those with diabetes, which could be summarized as uncontrolled diabetes.

Environment: The environment domain contained four variables. Three variables were carried forward from the OOI Version 2: access to green space, PM2.5 (particulate matter of 2.5 micrometers or less) levels in air, and walkability. The OOI Version 3 included three changes. First, it no longer includes percent high density urban landcover due to ambiguity in its likely effect on opportunity for healthy living. Because other measures in this domain capture much of what urban land cover does, but in more specific ways, the research team opted to remove urban land cover from the domain

(and the index altogether). Second, the research team added AirToxScreen data (https://www.epa.gov/AirToxScreen) to enrich the air quality component of the domain. From these data, we constructed two measures representing (1) Cancer risk and (2) non-cancer health risk (inclusive of respiratory, neurological, liver, developmental, reproductive, kidney, ocular, endocrine, hematological, immunological, skeletal, spleen, thyroid, and whole body hazards) due to air quality. The research term chose to add these measures instead of replacing PM2.5 with them because correlation analysis suggested that the PM2.5 data were measuring something distinct. Finally, the access to green space measure was enriched by the inclusion of a second data set—USA Parks from ESRI—that contained substantially more park locations than we had previously included in the measure. The previous parks shape file was augmented with this new data because, while they do overlap somewhat, they each contain distinct park areas.

Crime: The crime domain contained five variables. Four variables were carried over from the OOI Version 2: robbery; burglary, larceny-theft and motor vehicle theft; public drunkenness and DUI; and drug involved crimes. Changes to the OOI Version 3 included splitting the variable 'homicide, aggravated assault and sexual assault' into two distinct variables: 'homicide, aggravated assault' and 'sexual assault.'

Step 3: Imputation for Individual Measures

Missing values in each measure within each domain were imputed by the following procedure. A missing value for a tract was replaced with the mean value of the tract's adjacent neighbor tracts. If the adjacent neighbors also had no available data (rare), then those neighbors' adjacent neighbors were used to calculate a mean for the imputed value.

Step 4: Standardization of Variables and Combined Domains

As in OOI Version 2, all variables were obtained at the Census tract level. Because the measure values for Census tracts with zero population (e.g., airports) tend to be outliers, and because we largely are not concerned with opportunity in such tracts, the measure values of these tracts are set to be missing (if not missing already) and are not included in calculations for variable standardization to avoid skewing the standardization procedure (see step 5 for how zero-population tracts are treated). To combine all variables within a domain into a single measure, all variables were standardized by converting them into z-scores. Importantly, some standardized variables were reversed by multiplying the values by negative one to ensure large and small values had consistent meanings with respect to opportunity across indicators. Next, the standardized variables were combined using an unweighted mean within each domain for each census tract. The domain scores were then re-standardized to have a mean of zero and a standard deviation of one. These standardized domain scores are the ones that are made available in the published OOI data along with the overall OOI score.

Table 1. Opportunity Index Domains, Variables, and Data Sources from 2023 and 2024

Domains	2023 Variables	2024 Variables	2024 Changes
Transportation	Public Transit Access	**Public Transit Access**	We included taxi use (unclear from previous documentation if this was already included)
	Average Commute Time	Average Commute Time	
	Households without vehicle access	Households without vehicle access	
	Traffic Proximity	Traffic Proximity	
Education	Educational attainment	Educational attainment	
	School performance	School performance	
	Students on free and reduced lunch	Students on free and reduced lunch	
	High School graduation rate	High School graduation rate	
	Residential internet connection availability	**Residential internet connection performance**	Using a new dataset, not limited to residential internet access, using 300 Mbps download speed as the cutoff
Employment	Low-wage job access	Low-wage job access	
	Access to workforce or job training sites	Access to workforce or job training sites	
	Unemployment	Unemployment	
	Poverty	Poverty	
Housing	Median Rent	Median Rent	
	Median Home Value	Median Home Value	
	Concentration of existing LIHTC Units	Concentration of existing LIHTC Units	
	Housing stock built pre-1980s	Housing stock built pre-1980s	
	Residential overcrowding	Residential overcrowding	
	Residential mobility	Residential mobility	
Health	Geographic access to Medical Providers	Geographic access to Medical Providers	
	Geographic access to healthy food options	Geographic access to healthy food options	
	Age-adjusted mortality	Age-adjusted mortality	
	Preventable ED admits/visits	**Preventable ED admits/visits**	Algorithm updated with 2020 specification sheet.
	Diabetes admits/diagnoses	**Diabetes admits/diagnoses**	Previously used a denominator counting the full tract population. The denominator now counts people with any diabetes diagnoses.
Environment	Access to green space	**Access to green space**	In addition to the parks dataset we were using before, we are now also using the USA Parks dataset from ArcGIS. The two were combined to better identify parks.
	PM2.5 levels	PM2.5 levels	
	Walkability	Walkability	
	Percent high density urban landcover	---	Removed urban landcover because it's theoretical contribution to opportunity is mixed and unclear. Other measures in this domain are more specific about health opportunity.
		AirToxScreen: Cancer risk	Added to better represent air conditions
	---	**AirToxScreen: Non-cancer health risk**	Added to better represent air conditions
Crime	Homicide, aggravated assault, & sexual assault	**Homicide, aggravated assault**	Sexual assault was separated to its own category
	Robbery	Robbery	
	Public drunkenness and DUI	Public drunkenness and DUI	
	Drug involved crimes	Drug involved crimes	
	Burglary, larceny-theft & motor vehicle theft	Burglary, larceny-theft & motor vehicle theft	
	---	**Sexual assault**	Separated sexual assault from homicide and aggravated assault

Step 5: Domain-Level Imputation for Zero-Population Tracts

A final imputation step addresses zero-population tracts. Many of the measures in the index are not reliable or meaningful in tracts with a zero population. Moreover, because opportunity is intended to apply to people, it may not be necessary to have a reasonable domain or opportunity score in tracts where people do not live. However, with consideration of the possibility of errors in geocoding of analytic units (e.g., Medicaid members), it could prove valuable to have reasonable opportunity values representing tracts even if no people truly reside in them. Therefore, for all zero-population tracts, and for all domain scores, the research team imputed the mean of the domain in tracts neighboring the zero-population tract.

Step 6: Ranking and Transformation of Domain Scores

When combining domain scores to form an overall index, the research team followed previous iterations of the OOI by transforming domain scores prior to combination. The purpose of the transformation was to impart a desirable cancellation property to the index. The transformation first ranked the values in each domain and divided the ranks by the number of tracts to get scores in the range 0 to 1 (*adjusted rank score*) with 1 being the most deprived score. This resulted in a uniform distribution that was then transformed into an exponential distribution for the purpose of combining into an overall OOI using the following formula:

$$domain\ score\ =\ -23 * \log\left(1 - adjusted\ rank\ score * (1 - e^{-\frac{100}{23}})\right)$$

The exponential form of the domain scores is positively skewed in such a way that when multiple transformed domain scores are averaged together (as we did when constructing the OOI), a low level of opportunity with respect to one domain cannot be completely cancelled out by a correspondingly high level of opportunity on just one other domain. It will require more than one high-opportunity domain to counterbalance the depriving effect of a single low-opportunity domain. The exponential transformed domain scores are not provided in the published data file because their use is not recommend. But they can be reproduced from the data provided. Instead, we recommend using the standardized, but otherwise untransformed domain scores provided in the published data file.

Step 7: Combining Domains into Measure of Opportunity

Finally, domain scores were added together and divided by the number of domains (seven). The result was reversed—so that higher values represented more opportunity—and then rescaled to range from 0 to 100. The result constitutes the Opportunity Index Score.

Results

Products of the construction procedures are various outputs intended to verify and validate the procedures and the resulting Ohio Opportunity Index. These outputs are generated as HTML documents that are associated with this report as two appendices.

Appendix A provides a thorough description of every individual measure in each of the domains. The data descriptions are organized in tabs by domain. For each tabbed domain, the appendix includes a numeric description of each variable, a visual correlation heatmap describing the bivariate relationships among the variables in the domain, a choropleth map of each measure to visualize the geographic distribution, and a textual description of the measure and its source.

Appendix B provides the code that was used to construct the Ohio Opportunity Index, descriptions or visualizations of intermediate calculations, a visual representation of correlations among the domain scores, choropleth maps of domain scores and the overall index, and evidence of validity of the resulting index.

Specifically, regarding validity of the index, the calculations indicate the OOI V3 has strong positive and statistically significant correlations with the OOI V2 (r = 0.64) and the OCOI V2 (r = 0.66). The OOI V3 also correlates strongly with the Center for Disease Control's Social Vulnerability Index (r = -0.49), but in the negative direction due to the opposing meanings of the two indices. The latter correlation is notably weaker than those with previous versions of the OOI and OCOI, likely because of differing sets of measures, inclusive of Ohio-focused measures derived from Vital Statistics and Medicaid administrative data that are not available nationwide for a national index like the Vulnerability Index. Finally, *Appendix B* also presents a validity test that tests the correlation between OOI V3 and the infant mortality rates among Medicaid members during the years 2022 and 2023. The relationship is negative (as expected) and statistically significant (p < .01) suggesting a clear relationship between the two variables and further validating the index.

Limitations

Limitations of the OOI include the following:

1. It is reasonable to expect that not all seven domains contribute equally to health opportunity. Nevertheless, we combined the domains with equal weight. This choice followed an analysis of quantified expert opinions and a corresponding examination of a version of the OOI V2 weighted according to expert beliefs. Our findings were that the statistical and geographic differences between an unweighted version and an expert-weighted version were minimal, which led us to the simpler option of an unweighted combination scheme. More details about this choice are described in the report included with the data for Version 2 of the OCOI;

2. The measures *preventable emergency department visits*, *diabetes admissions,* and *age adjusted mortality* only contained data from Medicaid service records or Vital Statistics data from Medicaid members as opposed to services and records for all Ohioans. This results somewhat in a Medicaid focus to the index;

3. The index and the domain scores are not comparable across versions of the index. It is therefore not reasonable to compare, for a specific tract or group of tracts, the value from Version 2 of the OOI to the same value in Version 3. This is due to the nature of the index being a *relative* index of opportunity across areas as opposed to an absolute index. The values assigned to tracts do not have an inherent absolute meaning. They only have meaning relative to the values of other tracts in the same version of the measure.

Discussion

This report documents the construction of Version 3 of the Ohio Opportunity Index (OOI V3). The structure of the Index remained largely intact between Version 2 and Version 3. Minor changes were made to some of the measures and domains primarily to better align the measures with the intended meaning of the domain. Data were also updated to the most current available for all measures (2022 or 2023 in most, but not all cases). The results of our validity checks provide evidence that the OOI V3, as intended, measures opportunity for healthy living. We therefore have confidence this tool can be used for its intended purposes. These could include supporting policy makers and researchers who are interested in characterizing small areas of Ohio in terms of their opportunities for healthy living with aims inclusive of targeting policies and interventions, resources, or studies.

Appendices