

Primary language spoken by the Medicaid and CHIP population in 2020



CONTENTS

1 – About this brief	2
2 – Population comparison	4
<i>What are the primary languages spoken by Medicaid and CHIP enrollees, and how does this compare to the broader U.S. population?</i>	4
<i>Which states have a higher-than-average proportion of enrollees whose primary language is not English?</i>	5
3 – Spotlight on enrollees with a primary language other than English	6
<i>What is the racial and ethnic composition of enrollees with a primary language other than English compared with enrollees whose primary language is English?</i>	6
<i>What is the age distribution among enrollees whose primary language is not English compared with enrollees whose primary language is English?</i>	7
<i>Are enrollees whose primary language is not English equally likely to be enrolled with the same scope of benefits?</i>	8
<i>Are enrollees whose primary language is not English more or less likely to live in rural areas?</i>	9
<i>What are the common eligibility categories among enrollees whose primary language is not English?</i>	10
4 – Appendix	11

About this brief

WHAT IS THIS BRIEF?

This data brief describes enrollees in Medicaid and the Children’s Health Insurance Program (CHIP) in 2020 whose primary language is not English and compares them to enrollees who speak English as their primary language. CMS is releasing this data brief as part of ongoing agency efforts to measure disparities in access to care and make focused, evidence-based investments to improve health equity for the Medicaid and CHIP population. Over 10 percent of Medicaid and CHIP enrollees have a primary language other than English, which is slightly lower than the U.S. population. Enrollees with a non-English primary language are more likely to be Hispanic or non-Hispanic Asian/Pacific Islander (API) compared to enrollees whose primary language is English. They are also more likely to be over the age of 65, and to qualify for benefits through an eligibility category open only to older adults. This data brief excludes 16 states due to incomplete or unreliable primary language data in the 2020 TAF—a few of which are populous states with many non-English-speaking residents. The information in this brief, as well as updated counts that reflect data from later years, are available for download on data.medicaid.gov.

WHAT IS THE SOURCE OF INFORMATION?

This brief is based on data reported by states to CMS as part of the Transformed Medicaid Statistical Information System (T-MSIS). States report information each month via T-MSIS about their enrollees, Medicaid- and CHIP-covered services, payments to providers and managed care organizations, enrollees’ diagnoses and health conditions, and information on providers and managed care plans. These data are converted into the T-MSIS Analytic Files (TAF), which are optimized for research purposes. More information about T-MSIS and TAF is available at [Medicaid.gov](https://www.medicaid.gov).¹ Nearly all states have high-quality

and complete information on Medicaid and CHIP enrollment, scope of benefits, and age in the TAF, however some states have data quality issues related to their primary language data. This data brief excludes 16 states with incomplete or unreliable primary language information in the 2020 TAF. More information on TAF data quality can be found on the DQ Atlas.²

Figure 1 uses 2015–2019 American Community Survey (ACS) data on primary spoken language from the U.S. Census Bureau. More information is available at the Census Bureau website.³ Figure 3, which shows information on enrollee race and ethnicity, uses the 2020 TAF Race/Ethnicity Imputation (REI) Companion File, which includes state-reported information on enrollee race and ethnicity when it is reported and of good quality (74 percent of enrollees), and indirectly estimated race and ethnicity when the state-reported information is missing or unreliable (26 percent of enrollees). CMS based its indirect estimates of race and ethnicity on an enhanced version of a well-validated method that is widely used for this purpose: Bayesian Improved Surname and Geocoding (BISG). BISG⁴ draws on the racial and ethnic distribution associated with a person’s surname and geographic location to estimate the person’s probability of reporting being in each of six racial and ethnic groups. The distribution of race and ethnicity for surnames from the Census Bureau is not available separately for race and ethnicity, so the categories are combined here.⁵ CMS enhanced the standard BISG methodology for these data by supplementing with T-MSIS information about enrollees’ first names and their American Indian or Alaska Native certification. Lastly, Figure 6 uses data from the U.S. Department of Agriculture’s 2010 Rural-Urban Commuting Area code classification scheme.⁶

¹ <https://www.medicaid.gov/medicaid/data-systems/macbis/transformed-medicaid-statistical-information-system-t-msis/index.html>

² <https://www.medicaid.gov/dq-atlas/welcome>

³ <https://www.census.gov/data/developers/data-sets/acs-5year/2019.html>

⁴ <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6338295/>

⁵ <https://www.census.gov/data/developers/data-sets/surnames.html>

⁶ Enrollees are assigned to an urban/rural residence category based on the 2010 Rural-Urban Commuting Area (RUCA) code associated with their home or mailing address ZIP code reported in TAF. RUCA codes 4 through 10 correspond to rural areas. RUCA codes 1 through 3 are classified as urban areas, which include all towns and cities with a population of 50,000 or more, as well as surrounding suburban areas. More information on RUCA codes is available at: <https://www.ers.usda.gov/data-products/rural-urban-commuting-area-codes.aspx>.

About this brief

WHO IS INCLUDED IN THIS ANALYSIS?

All figures in this brief are limited to the 37 states and territories with low-concern or medium-concern data quality for the primary language variable in the 2020 TAF. Data from 16 states are excluded from all figures: Alaska, Arizona, California, Hawaii, Illinois, Louisiana, Michigan, Montana, Nebraska, Nevada, New Jersey, New Mexico, New York, Pennsylvania, Rhode Island, and Texas. Figures in this brief do not include enrollees in Guam, American Samoa, and the Northern Mariana Islands because T-MSIS data are not available for these territories. Medicaid and CHIP provide coverage to a diverse enrollee population, and since this brief is restricted to the 37 states and territories with quality primary language data in TAF, the results may not fully reflect the distribution of languages spoken in the national Medicaid and CHIP population.

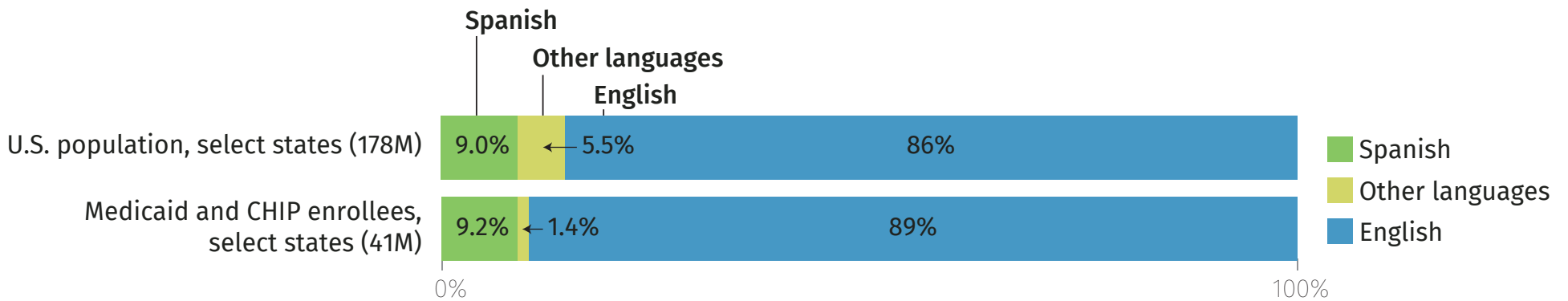
The figures include enrollees with any scope of benefits who were enrolled for at least one day in 2020. This includes enrollees who were dually eligible for Medicare and those with a limited benefit package. Figure 1 excludes enrollees younger than age 5 for consistency with the ACS data that are used as a comparison. Figures 1 and 3 exclude the U.S. Virgin Islands since this territory is not included in the ACS data or the REI Companion File.

HOW IS PRIMARY LANGUAGE MEASURED?

States report an enrollee's primary language in T-MSIS using the three-digit International Organization for Standardization (ISO) codes. For this brief, all languages other than "English" or "Spanish" are treated as a single category (that is, "other languages"), given the low number of enrollees with a reported primary language other than those two. Following the convention used by the Social Security Administration and the U.S. Census Bureau, many states report English as an enrollee's primary language by leaving the primary language data element missing. To accommodate this practice, all missing values are categorized as English in this brief. Because of the inability to differentiate between missing values that are intended to indicate English as a primary language and missing values that indicate the primary language is unknown or unrecorded, the number of enrollees whose primary language is English is likely overestimated.

What are the primary languages spoken by Medicaid and CHIP enrollees, and how does this compare to the broader U.S. population?

FIGURE 1. PRIMARY LANGUAGES SPOKEN BY INDIVIDUALS AGE 5 AND OLDER IN 2020



Complete information about the primary languages spoken by Medicaid and CHIP enrollees is available from 37 states and territories, which together covered 47 million enrollees in 2020. The distribution of primary languages spoken among the Medicaid and CHIP population is similar to that of the U.S. population in these states and territories. The share of Medicaid and CHIP enrollees with a primary language other than English and Spanish is smaller than the general U.S. population from the same areas.

The distribution of primary languages in these states may not be representative of the national Medicaid and CHIP population, as some populous states with many non-English-speaking residents—such as California and Texas—did not report complete information on primary language and are excluded from this figure.

SOURCES: 2020 T-MSIS Analytic Files Annual Demographic and Eligibility File, Release 1; 2015–2019 American Community Survey.

NOTES: Percentages may not sum to 100% due to rounding. This figure includes data from 36 states, including the District of Columbia and Puerto Rico, where complete and reliable information on primary language was available in T-MSIS. This figure excludes the following states from both the Medicaid and CHIP enrollee and U.S. population counts because of incomplete or unreliable T-MSIS data: Alaska, Arizona, California, Hawaii, Illinois, Louisiana, Michigan, Montana, Nebraska, Nevada, New Jersey, New Mexico, New York, Pennsylvania, Rhode Island, and Texas. The American Community Survey data on primary language only include people age 5 and older and do not include the U.S. Virgin Islands. For comparability, the Medicaid and CHIP population in this figure is restricted to people age 5 and older and does not include the U.S. Virgin Islands.

Which states have a higher-than-average proportion of enrollees whose primary language is not English?

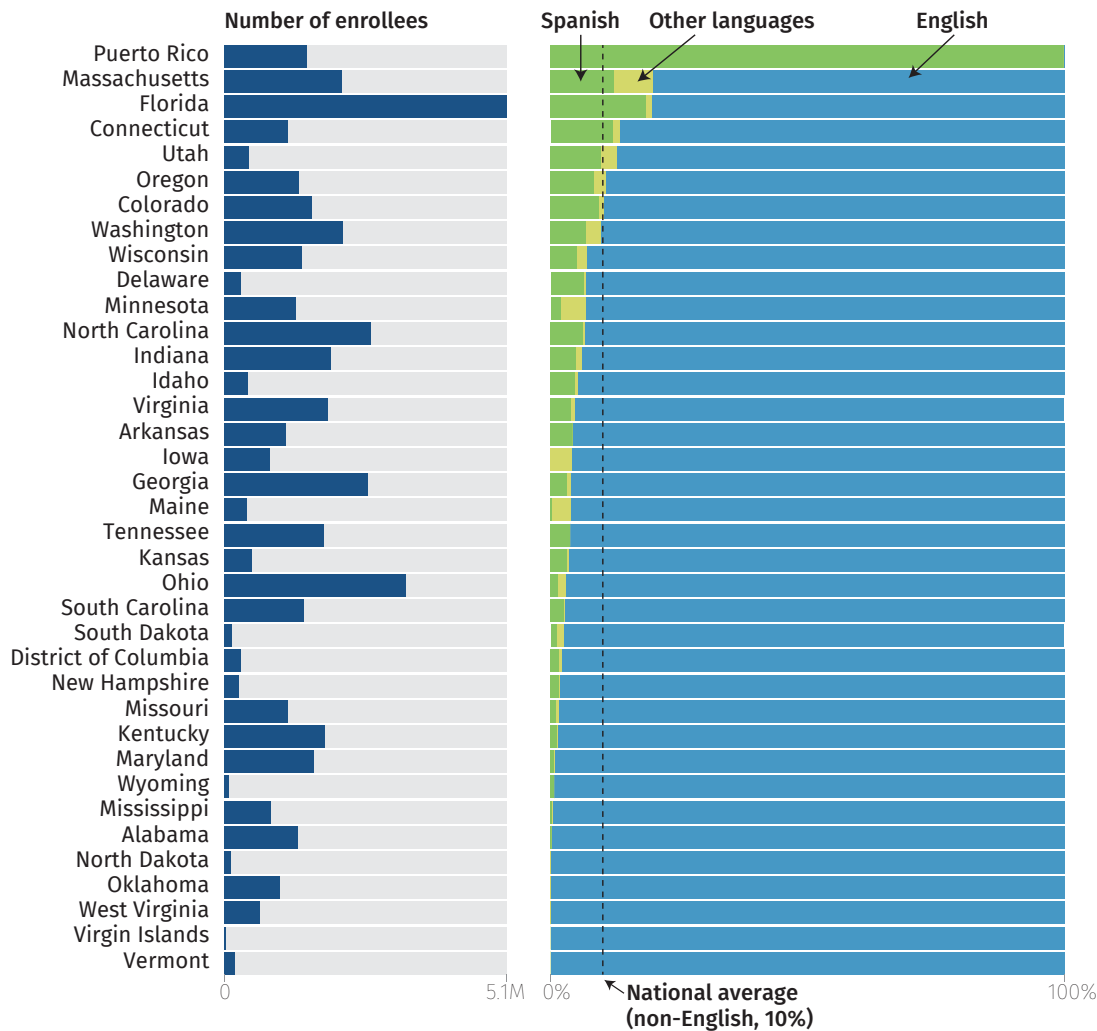


FIGURE 2. PROPORTION OF MEDICAID AND CHIP ENROLLEES WHOSE PRIMARY LANGUAGE IS NOT ENGLISH IN 2020, BY STATE

- Number of enrollees
- Spanish
- Other languages
- English

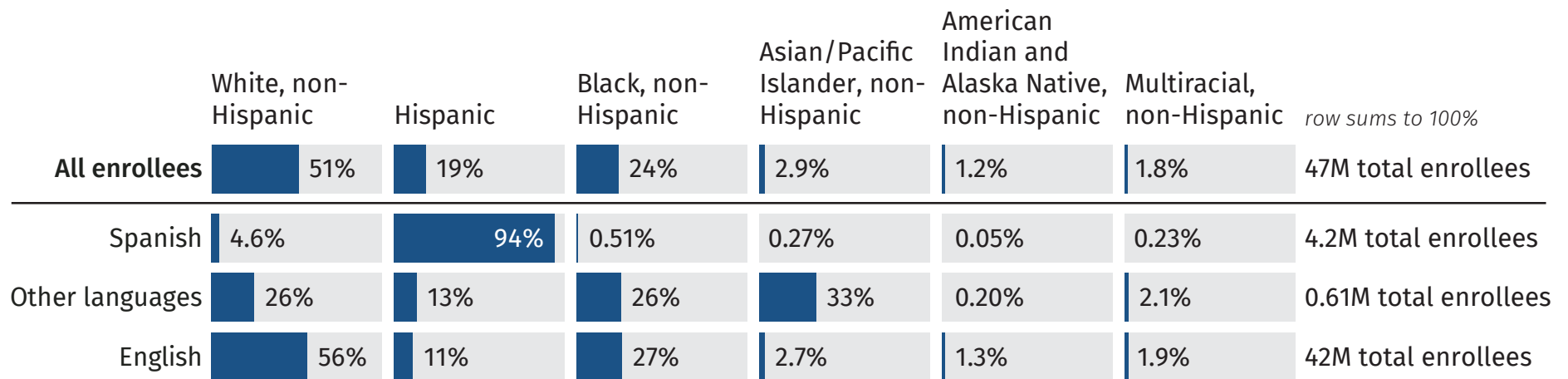
In six states and one U.S. territory, the proportion of Medicaid and CHIP enrollees whose primary language is not English exceeds the national average. More than 10 percent of enrollees in those states speak a non-English primary language. Spanish is the most common non-English primary language among enrollees residing in these areas. Puerto Rico has the highest proportion of enrollees whose primary language is Spanish (99.8 percent), followed by Florida (19 percent), and Massachusetts (12 percent). Compared with other states, Massachusetts has the largest share of enrollees (7.5 percent) who speak a language other than English or Spanish as their primary language.

SOURCE: 2020 T-MSIS Analytic Files Annual Demographic and Eligibility File, Release 1.

NOTES: This figure includes data from 37 states, including the District of Columbia, Puerto Rico, and the U.S. Virgin Islands, where complete and reliable information on primary language was available in T-MSIS. This figure excludes the following states because of incomplete or unreliable T-MSIS data: Alaska, Arizona, California, Hawaii, Illinois, Louisiana, Michigan, Montana, Nebraska, Nevada, New Jersey, New Mexico, New York, Pennsylvania, Rhode Island, and Texas.

What is the racial and ethnic composition of enrollees with a primary language other than English compared with enrollees whose primary language is English?

FIGURE 3. RACE AND ETHNICITY OF MEDICAID AND CHIP ENROLLEES WHOSE PRIMARY LANGUAGE IS NOT ENGLISH IN 2020



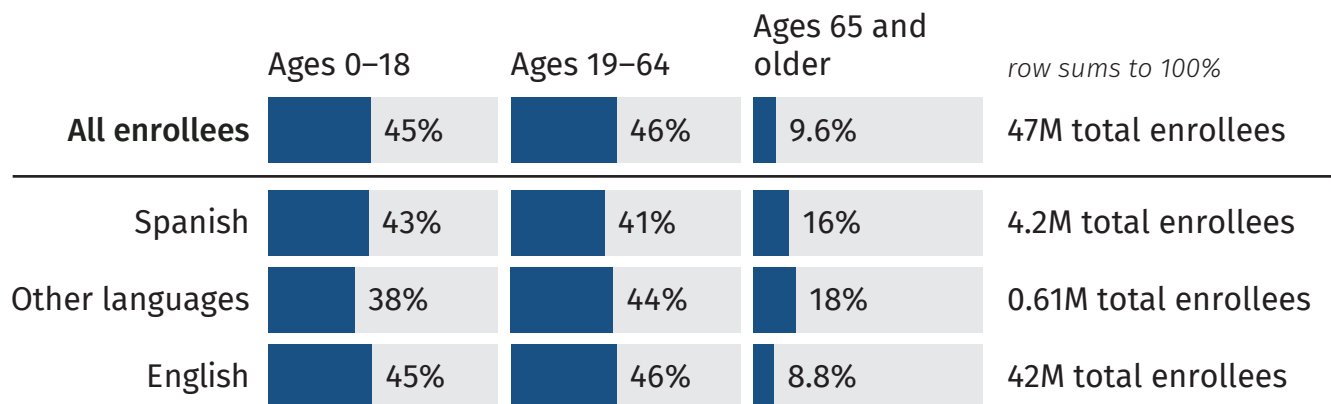
Medicaid and CHIP enrollees whose primary language is Spanish are much more likely to be Hispanic compared with enrollees who spoke English as their primary language. Enrollees whose primary language is neither English nor Spanish are much more likely to be non-Hispanic Asian/Pacific Islander compared with enrollees whose primary language is English.

SOURCES: 2020 T-MSIS Analytic Files Annual Demographic and Eligibility File, Release 1 and 2020 Race/Ethnicity Imputation (REI) Companion File.

NOTES: This figure includes data from 36 states, including the District of Columbia and Puerto Rico, where complete and reliable information on primary language was available in T-MSIS. This figure excludes the following states because of incomplete or unreliable T-MSIS data: Alaska, Arizona, California, Hawaii, Illinois, Louisiana, Michigan, Montana, Nebraska, Nevada, New Jersey, New Mexico, New York, Pennsylvania, Rhode Island, and Texas. The U.S. Virgin Islands are not included because data from that territory are not included in the 2020 REI Companion File.

What is the age distribution among enrollees whose primary language is not English compared with enrollees whose primary language is English?

FIGURE 4. AGE DISTRIBUTION OF MEDICAID AND CHIP ENROLLEES WHOSE PRIMARY LANGUAGE IS NOT ENGLISH IN 2020



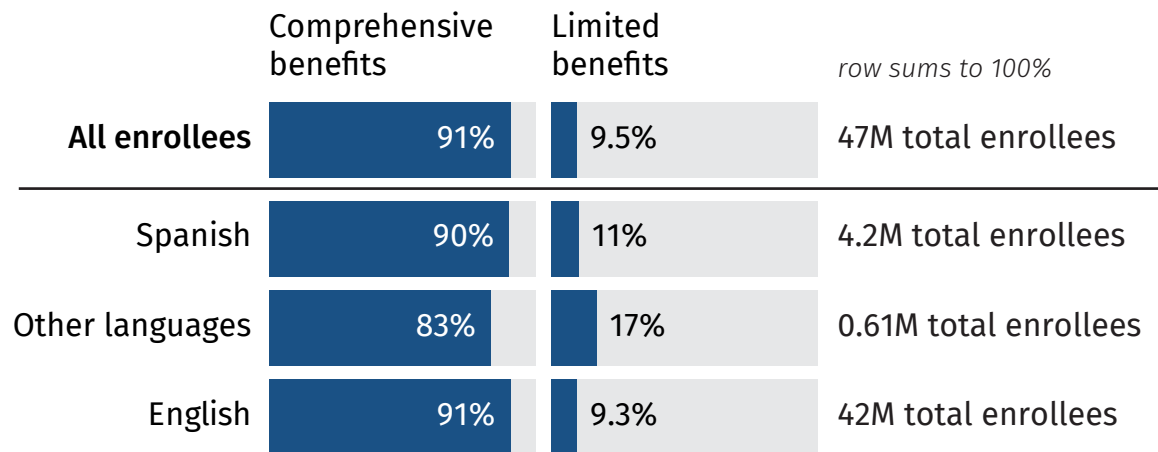
Compared with enrollees whose primary language is English, Medicaid and CHIP enrollees whose primary language is Spanish or another language are more likely to be age 65 and older.

SOURCE: 2020 T-MSIS Analytic Files Annual Demographic and Eligibility File, Release 1.

NOTES: Percentages may not sum to 100% due to rounding. This figure includes data from 37 states, including the District of Columbia, Puerto Rico, and the U.S. Virgin Islands, where complete and reliable information on primary language was available in T-MSIS. This figure excludes the following states because of incomplete or unreliable T-MSIS data: Alaska, Arizona, California, Hawaii, Illinois, Louisiana, Michigan, Montana, Nebraska, Nevada, New Jersey, New Mexico, New York, Pennsylvania, Rhode Island, and Texas.

Are enrollees whose primary language is not English equally likely to be enrolled with the same scope of benefits?

FIGURE 5. SCOPE OF BENEFITS AMONG MEDICAID AND CHIP ENROLLEES WHOSE PRIMARY LANGUAGE IS NOT ENGLISH IN 2020



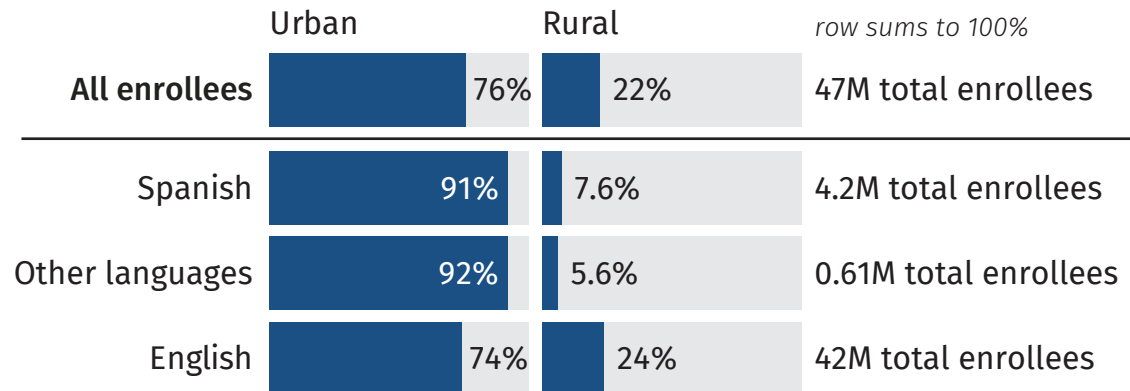
The majority of Medicaid and CHIP enrollees qualify for a comprehensive benefit package that meets the minimum essential coverage standard under the Affordable Care Act. A much smaller group qualifies for limited benefits only. The most common limited benefit packages in Medicaid cover only family-planning services, emergency services, or assistance with premiums and cost-sharing in Medicare. Medicaid and CHIP enrollees who speak other languages as their primary language are much more likely to have a limited benefit package than enrollees whose primary language is English. This may reflect that enrollees in the “other languages” group are more likely to be older than 65. In most cases, people older than 65 receive their primary health coverage through Medicare, and Medicaid may only provide partial benefits to help cover the costs of premiums and cost-sharing for that program.

SOURCE: 2020 T-MSIS Analytic Files Annual Demographic and Eligibility File, Release 1.

NOTES: Percentages may not sum to 100% due to rounding. This figure includes data from 37 states, including the District of Columbia, Puerto Rico, and the U.S. Virgin Islands, where complete and reliable information on primary language was available in T-MSIS. This figure excludes the following states because of incomplete or unreliable T-MSIS data: Alaska, Arizona, California, Hawaii, Illinois, Louisiana, Michigan, Montana, Nebraska, Nevada, New Jersey, New Mexico, New York, Pennsylvania, Rhode Island, and Texas.

Are enrollees whose primary language is not English more or less likely to live in rural areas?

FIGURE 6. RURAL RESIDENCE AMONG MEDICAID AND CHIP ENROLLEES WHOSE PRIMARY LANGUAGE IS NOT ENGLISH IN 2020



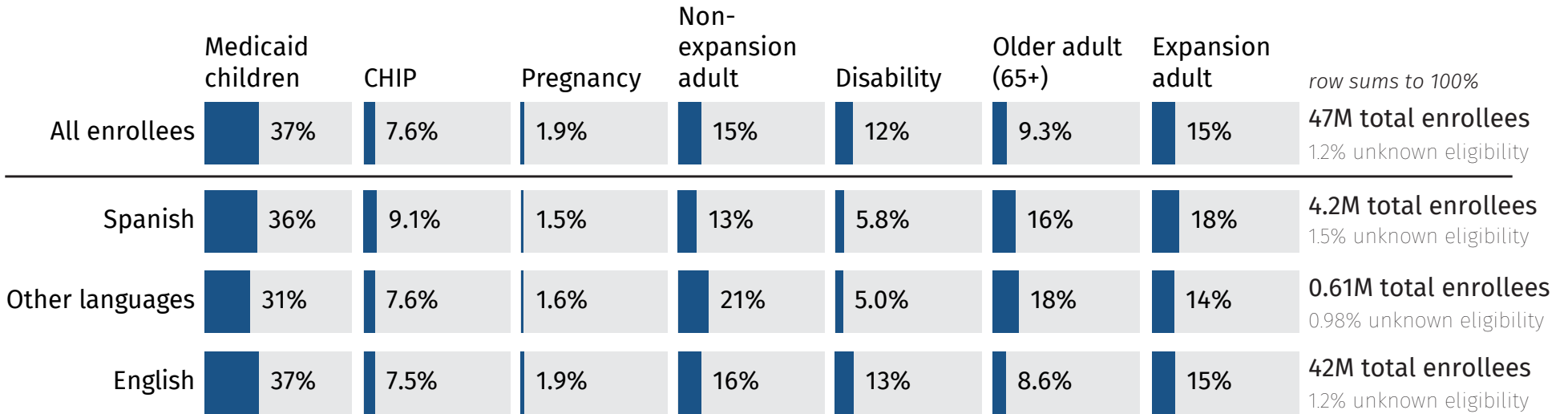
Medicaid and CHIP enrollees whose primary language is not English are much less likely to reside in a rural area than enrollees whose primary language is English. More than 90 percent of enrollees whose primary language is not English live in an area designated as urban, which encompass all cities and towns with a population exceeding 50,000, as well as their surrounding suburbs. The share of enrollees that reside in a rural area is largest among enrollees whose primary language is English.

SOURCE: 2020 T-MSIS Analytic Files Annual Demographic and Eligibility File, Release 1.

NOTES: Percentages may not sum to 100% due to rounding. This figure includes data from 37 states, including the District of Columbia, Puerto Rico, and the U.S. Virgin Islands, where complete and reliable information on primary language was available in T-MSIS. This figure excludes the following states because of incomplete or unreliable T-MSIS data: Alaska, Arizona, California, Hawaii, Illinois, Louisiana, Michigan, Montana, Nebraska, Nevada, New Jersey, New Mexico, New York, Pennsylvania, Rhode Island, and Texas.

What are the common eligibility categories among enrollees whose primary language is not English?

FIGURE 7. MEDICAID AND CHIP ELIGIBILITY CATEGORIES AMONG ENROLLEES WHOSE PRIMARY LANGUAGE IS NOT ENGLISH IN 2020



The most common eligibility category for obtaining coverage among Medicaid and CHIP enrollees, regardless of their primary language, is through the Medicaid children category. Compared with enrollees whose primary language is English, enrollees whose primary language is Spanish or another language are more likely to obtain coverage through the older adult eligibility category. These enrollees are much less likely to qualify for benefits on the basis of a disability compared with enrollees whose primary language is English.

SOURCE: 2020 T-MSIS Analytic Files Annual Demographic and Eligibility File, Release 1

NOTES: Percentages may not sum to 100% due to rounding. This figure includes data from 37 states, including the District of Columbia, Puerto Rico, and the U.S. Virgin Islands, where complete and reliable information on primary language was available in T-MSIS. This figure excludes the following states because of incomplete or unreliable T-MSIS data: Alaska, Arizona, California, Hawaii, Illinois, Louisiana, Michigan, Montana, Nebraska, Nevada, New Jersey, New Mexico, New York, Pennsylvania, Rhode Island, and Texas.

Number of Medicaid and CHIP enrollees in each primary language group, by race and ethnicity, age, benefits package, rural residency, and eligibility category

The data in these tables are available for download at data.medicaid.gov. Updated counts that reflect data from later years will be released as they become available on data.medicaid.gov.

Table 1. Primary language among people age 5 and older

	U.S. population, select states (178M)	Medicaid and CHIP enrollees, select states (41M)
Spanish	16M	3.8M
Other languages	9.8M	0.56M
English	152M	37M

Table 2. Primary language among Medicaid and CHIP enrollees, by race and ethnicity

	White, non-Hispanic (24M)	Hispanic (8.7M)	Black, non-Hispanic (11M)	API, non-Hispanic (1.4M)	AIAN, non-Hispanic (0.54M)	Multiracial, non-Hispanic (0.83M)
Spanish (4.2M)	0.19M	3.9M	0.02M	0.01M	0.002M	0.01M
Other languages (0.61M)	0.16M	0.08M	0.16M	0.20M	0.001M	0.01M
English (42M)	24M	4.7M	11M	1.1M	0.54M	0.81M

API = Asian/Pacific Islander, AIAN = American Indian and Alaska Native

Table 3. Primary language among Medicaid and CHIP enrollees, by age, scope of benefits, rural residency, and eligibility category

	Ages 0–18 (21M)	Ages 19–64 (21M)	Ages 65+ (4.5M)	Comprehensive benefits (42M)	Limited benefits (4.4M)	Urban (35M)	Rural (10M)	Medicaid children (17M)	CHIP (3.6M)	Pregnancy (0.88M)	Non-expansion adult (7.2M)	Disability (5.6M)	Older adult, 65+ (4.4M)	Expansion adult (7.2M)
Spanish (4.2M)	1.8M	1.7M	0.7M	3.7M	0.44M	3.8M	0.31M	1.5M	0.38M	0.06M	0.52M	0.24M	0.64M	0.76M
Other languages (0.61M)	0.23M	0.27M	0.11M	0.50M	0.10M	0.56M	0.03M	0.19M	0.05M	0.01M	0.13M	0.03M	0.11M	0.09M
English (42M)	19M	19M	3.7M	38M	3.9M	31M	10M	16M	3.1M	0.81M	6.5M	5.3M	3.6M	6.4M

SOURCES: 2020 T-MSIS Analytic Files Annual Demographic and Eligibility File, Release 1; 2020 Race/Ethnicity Imputation (REI) Companion File; 2015–2019 American Community Survey.

NOTES: These tables include data from 37 states, including the District of Columbia and Puerto Rico, where complete and reliable information on primary language was available in T-MSIS. Data from the U.S. Virgin Islands are included in Table 3 but not in Tables 1 and 2 since this territory is not included in the American Community Survey or the REI Companion File. Table 1 is further restricted to individuals age 5 and older. All tables exclude the following states because of incomplete or unreliable T-MSIS data: Alaska, Arizona, California, Hawaii, Illinois, Louisiana, Michigan, Montana, Nebraska, Nevada, New Jersey, New Mexico, New York, Pennsylvania, Rhode Island, and Texas.