

Physician Data Initiative Language Proficiency Standard

Development of the Language Proficiency Standard

To develop the proposed language proficiency standard recommendation, the Association of American Medical Colleges (AAMC), the Accreditation Council for Graduate Medical Education (ACGME) and the American Medical Association (AMA) reviewed both the AAMC language proficiency question and the Interagency Language Roundtable for Healthcare (ILR-H). The AAMC language proficiency question has been used on the ERAS and AMCAS applications since 2013 (<https://students-residents.aamc.org/media/9711/download>). It is worth noting that the AAMC question is a revision of the ILR-H. The ILR-H is a self-reporting tool used and validated in health care settings; based on the Federal Government's Interagency Language Roundtable with descriptions used by the Foreign Service Institute. The AAMC revision of the ILR-H preserved most elements of the ILR-H: both the AAMC question and the ILR-H are single-item questions with five categorical response choices and accompanying descriptors. The differences between the two scales consist of shortening of the descriptors in the AAMC question compared to the ILR-H's longer descriptors and some word choice edits for clarity. No substantive changes were made to the scale. The AAMC has been collecting language data through AMCAS and ERAS for 10 years, and some analyses using these data have been published (1,2). The ILR-H has been used to better characterize the skills of physicians and learners beyond solely identifying what languages they speak and to determine eligibility for medical language education or formal language testing (3,4,5).



The Physician Data Initiative's Common Data Standards group reviewed the AAMC language proficiency question and the ILR-H scale side-by-side and proposed a revised standard based on an updated review of the literature and language expert consultation. The revisions are intended to clarify the descriptors by applying inclusive language principles (e.g., avoiding words that may inadvertently exclude individuals), modernizing the wording (e.g., replacing outdated terms with clearer, updated terms), and enhancing the usability of the scale (e.g., removing jargon or idiomatic phrases that may not be understood by some users), while preserving the original meaning and keeping the descriptors as concise as possible. Furthermore, while the ILR-H has previously been validated only in spoken languages, the wording of the descriptors has been revised to be applicable to signed languages.

Language Proficiency

Indicate your languages and your proficiency level in each. Select all that apply. For each language that you select, including English, you will be asked to rate your proficiency in that language using the guidelines provided below. Note that the proficiency scale was developed for spoken languages. For reporting proficiency in a signed language, please select American Sign Language or Another signed language, and select the proficiency level that most closely matches your skill level.

- **Native/Near-Native:** I converse easily and accurately in all types of situations, including communicating health care concepts. Individuals at the native/near-native level may think that I have native/near-native skills, too.
- **Advanced:** I converse very accurately, and I understand others very accurately. Language ability only rarely hinders me in performing any task, including communicating health care concepts. Individuals at the native/near-native level have no problem understanding me, but they probably perceive that I do not have native/near-native skills.
- **Good:** I communicate well enough to participate in most conversations. Individuals at the native/near-native level notice some errors in my language production or my understanding, but I am generally able to repair the conversation if errors or misunderstandings occur. I have some difficulty communicating health care concepts.
- **Fair:** I communicate and understand well enough to have casual conversations about current events, work, family, or personal life and can get the general idea of most everyday conversations. Individuals at the native/near-native level notice many errors in my language production or my understanding. I have difficulty communicating about health care concepts.
- **Basic:** I can use the language at a level that permits me to understand and respond to 2 to 3 word entry level questions and meet minimum courtesy requirements. I have difficulty participating in or understanding conversations. I am unable to understand or communicate most health care concepts.

Recommendations for Applying the Language Proficiency Standard

The language proficiency standard is intended as a scale to report physician language proficiency across the continuum of medical education and career. Since language skills are dynamic, it is important that the instrument can be used to track language proficiency changes over time throughout a physician's education and career. For example, someone's language skills may be enhanced with increased use or through learning opportunities (e.g., taking a medical language course, practicing in an area where a language is used frequently in patient care) or, conversely, may decrease over time as a result of disuse or lack of practice.

It may be particularly useful for physicians to self-assess their language skills at major transition points in their education and career, such as starting clinical clerkships in medical school, starting residency or fellowship, entering medical practice, or changing practice sites or geographic location. Such transition points are times when physicians may have new responsibilities (graduated independence, new supervisory responsibilities) and may also result in a change in patient population with different language characteristics.

Self-Identification of Language Proficiency

The language proficiency standard is intended for self-identification of language skills and should be self-reported. Data must be collected and shared in compliance with applicable laws. Institutions should consider what these data will be used for and should disclose the purpose of data collection to respondents consistent with the ethical research, data privacy, accommodation, and human resources policies of the institutions. Examples of how language self-reported data may be used include the following:

- Screen for eligibility to enroll in available medical language courses
- Screen for eligibility to take a formal medical language proficiency assessment
- Identify the language skills represented among the workforce, relative to languages prevalent in the patient population
- Identify candidates with language skills that are aligned with the institution's patient population
- Enable analysis of quality of care and outcomes metrics and their potential association with clinician-patient language concordance
- Plan for institutional language assistance services resource allocation to address gaps in languages to areas of the hospital where there are greatest language-concordant clinician deficits

Importantly, experts recommend against the use of language proficiency self-assessment alone (i.e., without formal testing) for the purposes of medical language proficiency certification (4). Clinicians should use qualified medical interpreters when communicating with patients with a different language preference unless the clinician has been confirmed as a qualified multilingual clinician by their institution or another appropriate authority. Furthermore, all clinicians should partner with a medical interpreter whenever they encounter linguistic challenges in communicating with a patient, even if the clinician has been deemed qualified to use that language in direct patient care. Additional study is needed to confirm the validity of the language proficiency self-reporting question for signed languages.

List of Languages and Language Coding

When creating a list of answer choices for collecting responses to a language proficiency question, we recommend displaying as many languages as possible that are represented within an institution's population. For health care and medical education institutions, it is important to consider both the languages that are represented in the patient population as well as those used by learners, clinicians, and staff. Including patient languages will help identify potential gaps in clinicians with skills in those languages, known as language-concordant clinician deficits (6). Including the languages used by learners, clinicians, and staff, even those that are not common in the patient population, may provide visibility to multilingualism as a valued asset in the workforce (7).

For classification of spoken languages, we recommend using the International Organization for Standardization's ISO-639-3 standard, which is a widely used standard for listing and coding languages worldwide. Since 2016, the United States Census has been using the same ISO-639-3 language coding standard, which is publicly available as a [code list](#). For signed languages, American Sign Language (ASL) is the most common signed language in the US; it should be noted that substantial variations in sign language acquisition and use are prevalent even among individuals who identify as using ASL. Additionally, individuals of varied national origins may use other signed languages; hence, a separate write-in category should be included to enable the reporting of another signed language.

Following ACGME's review of the available residency applicant data (2013-2023) collected by the AAMC through the ERAS application, together with the most common languages spoken by individuals with limited English proficiency as collected by the 2022 American Community Survey and reported by the US Census (8), we recommend that health care workforce questions about language proficiency display the following list of languages (in alphabetical order):

- American Sign Language
- Amharic
- Arabic
- Armenian
- Bengali
- Burmese
- English
- French
- French Creole
- German
- Greek
- Gujarati
- Haitian
- Hawaiian
- Hebrew
- Hindi
- Hmong
- Igbo

- Ilocano
- Italian
- Japanese
- Kannada
- Khmer
- Korean
- Kru
- Lao
- Malayalam
- Mandarin Chinese
- Marathi
- Navajo
- Nepali
- Panjabi
- Pashto
- Persian (incl. Farsi, Dari)
- Polish
- Portuguese
- Punjabi
- Romanian
- Russian
- Samoan
- Serbo-Croatian
- Spanish
- Swahili
- Tagalog
- Tamil
- Telugu
- Thai
- Turkish
- Twi
- Ukrainian
- Urdu
- Vietnamese
- Yiddish
- Yoruba
- Yue Chinese (incl. Cantonese)
- Another spoken language [write-in]
- Another signed language [write-in]

A write-in option should be included to capture unlisted languages that respondents wish to report. Separate write-in options for spoken and signed languages are recommended for clarity. Individual countries have their own signed languages; hence, it is important to differentiate whether a person is reporting a spoken or signed language. Additionally, language lists should be periodically revisited for relevance to the populations of interest, as language prevalence is expected to change over time.

Language Reporting

If the number of individuals reporting a spoken language is too small to be reported as a single category, the following language groupings per the US Census can be used to guide aggregating groups. Other similar language family groupings can be ascertained from the ISO codes.

- Amharic, Somali, or other Afro-Asiatic languages
- Chinese (incl. Mandarin and Cantonese)
- Ilocano, Samoan, Hawaiian, or other Austronesian languages
- Nepali, Marathi, or other Indic languages
- Swahili or other languages of Central, Eastern, and Southern Africa
- Thai, Lao, or other Tai-Kadai languages
- Ukrainian or other Slavic languages
- Yiddish, Pennsylvania Dutch or other West Germanic languages
- Yoruba, Twi, Igbo, or other languages of Western Africa

If the number of individuals reporting a signed language is too small to be reported as a single category, a grouping of “Signed Languages” may be appropriate.

References

1. Diamond L, Grbic D, Genoff M, et al. Non-English-language proficiency of applicants to US residency programs. *JAMA*. 2014;312(22):2405-2407. doi:10.1001/jama.2014.15444
2. Diamond LC, Mujawar I, Vickstrom E, Garzon MG, Gany F. Supply and Demand: Association Between Non-English Language-Speaking First Year Resident Physicians and Areas of Need in the USA. *J Gen Intern Med*. 2020;35(8):2289-2295.
3. Diamond LC, Luft HS, Chung S, Jacobs EA. "Does this doctor speak my language?" Improving the characterization of physician non-English language skills. *Health Serv Res*. 2012;47(1 Pt 2):556-569. doi:10.1111/j.1475-6773.2011.01338.x
4. Diamond L, Toro Bejarano M, Chung S, et al. Factors Associated With Accuracy of Self-Assessment Compared With Tested Non-English Language Proficiency Among Primary Care Providers. *Med Care*. 2019;57(5):385-390. doi:10.1097/MLR.0000000000001105
5. Ortega P, Diamond L, Alemán MA, et al. Medical Spanish Standardization in U.S. Medical Schools: Consensus Statement From a Multidisciplinary Expert Panel. *Acad Med*. 2020;95(1):22-31. doi:10.1097/ACM.0000000000002917
6. Hsu P, Balderas-Medina Anaya Y, Anglin L, Hayes-Bautista D. California's Language Concordance Mismatch: Clear Evidence for Increasing Physician Diversity. <https://latino.ucla.edu/wp-content/uploads/2019/06/AltaMed-Policy-Brief-1.pdf>. September 2018.
7. Ortega P, Shin TM, Martínez GA. Rethinking the Term "Limited English Proficiency" to Improve Language-Appropriate Healthcare for All. *J Immigr Minor Health*. 2022;24(3):799-805. doi:10.1007/s10903-021-01257-w
8. U.S. Census Bureau. "Language Spoken at Home by Ability to Speak English for the Population 5 Years and Over." American Community Survey, ACS 5-Year Estimates Detailed Tables, Table B16001, 2022, <https://data.census.gov/table/ACSDT5Y2022.B16001?q=B16001>: LANGUAGE SPOKEN AT HOME BY ABILITY TO SPEAK ENGLISH FOR THE POPULATION 5 YEARS AND OVER. Accessed on April 2, 2024.

Copyright Notice, Permission and Disclaimer and Attribution Statement

This content may be modified at any time upon written agreement of the Parties (email is sufficient).

Unless the instructions indicate otherwise, each use or disclosure of the Standards by any Party or third-party shall include the following placed in close proximity to the Standards:

Copyright Notice:

[*Original Standards – Physician Data Initiative*] © 2024. Accreditation Council for Graduate Medical Education (ACGME), American Medical Association (AMA), and Association of American Medical Colleges (AAMC). May be reproduced and distributed with attribution to the AAMC, ACGME, and AMA. Modifications are permitted with an acknowledgement of the modifications made.

Disclaimer:

DISCLAIMER. The Standards are not guidelines and have not been tested for all potential applications. The Standards, in their original or a modified form, are provided “as is” without express or implied warranties of any kind, including warranties of merchantability, fitness for a particular purpose, or non-infringement. The AAMC, ACGME, and AMA expressly disclaim and assume no liability for use of or reliance upon the Standards whether in their original or a modified form.

Attribution Statements:

Any reproduction or distribution of the Standards without modification shall include the following attribution statement:

“This [reproduction, etc.] was prepared using [Physician Data Initiative Standards] developed by the Accreditation Council for Graduate Medical Education (ACGME), American Medical Association (AMA), and Association of American Medical Colleges (AAMC). The content reflects the views of [name of user/author].”

Any reproduction or distribution of the Standards with modification shall include the following attribution statement:

“This [reproduction, etc.] was prepared using [Physician Data Initiative Standards] originally developed by the Accreditation Council for Graduate Medical Education (ACGME), American Medical Association (AMA), and Association of American Medical Colleges (AAMC) and modified as follows: [____]. The content reflects the views of [name of user/author].”