

PERSPECTIVE

Cross-national comparisons of later-life cognitive function using data from the Harmonized Cognitive Assessment Protocol (HCAP): Considerations and recommended best practices

Lindsay C. Kobayashi^{1,2}  | Richard N. Jones³ | Emily M. Briceño⁴ | Miguel Arce Rentería⁵ | Yuan Zhang^{6,7} | Erik Meijer⁸ | Kenneth M. Langa^{2,9,10} | Jinkook Lee^{8,11} | Alden L. Gross^{12,13}

¹Department of Epidemiology, Center for Social Epidemiology and Population Health, University of Michigan School of Public Health, Ann Arbor, Michigan, USA

²Survey Research Center, University of Michigan Institute for Social Research, Ann Arbor, Michigan, USA

³Department of Psychiatry and Human Behavior, Warren Alpert Medical School, Brown University, Providence, Rhode Island, USA

⁴Department of Physical Medicine & Rehabilitation, University of Michigan Medical School, Ann Arbor, Michigan, USA

⁵Taub Institute for Research on Alzheimer's Disease and the Aging Brain, Department of Neurology, Columbia University College of Physicians and Surgeons, New York City, New York, USA

⁶Department of Sociomedical Sciences, Columbia University, New York City, New York, USA

⁷Robert N. Butler Columbia Aging Center, Columbia University, New York City, New York, USA

⁸Center for Economic and Social Research, University of Southern California, Los Angeles, California, USA

⁹Department of Internal Medicine, University of Michigan Medical School, Ann Arbor, Michigan, USA

¹⁰Veterans Affairs Center for Clinical Management Research, Ann Arbor, Michigan, USA

¹¹Department of Economics, University of Southern California, Los Angeles, California, USA

¹²Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, USA

¹³Center on Aging and Health, Johns Hopkins University, Baltimore, Maryland, USA

Correspondence

Lindsay C. Kobayashi, Department of Epidemiology, Center for Social Epidemiology and Population Health, University of Michigan School of Public Health, 1415 Washington Heights, Ann Arbor, MI, 48109, USA.
Email: lkob@umich.edu

Funding information

National Institutes of Health, Grant/Award Numbers: R01AG070953, R01AG030153, U24AG065182, K23AG080035, R00AG070274

Abstract

The Harmonized Cognitive Assessment Protocol (HCAP) is a major innovation that provides, for the first time, harmonized data for cross-national comparisons of later-life cognitive functions that are sensitive to linguistic, cultural, and educational differences across countries. However, cognitive function does not lend itself to direct comparison across diverse populations without careful consideration of the best practices for such comparisons. This perspective discusses theoretical and methodological considerations and offers a set of recommended best practices for conducting cross-national comparisons of risk factor associations using HCAP data. Because existing and planned HCAP studies provide cognition data representing an estimated 75% of the global population ≥ 65 years of age, these recommended best practices will support high-quality comparative analyses of cognitive aging around the world. The principles

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial](https://creativecommons.org/licenses/by-nc/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2024 The Authors. *Alzheimer's & Dementia* published by Wiley Periodicals LLC on behalf of Alzheimer's Association.

described in this perspective are applicable to any researcher aiming to integrate or compare harmonized data on cognitive outcomes and their risk and protective factors across diverse populations.

KEYWORDS

cognitive function, cognitive impairment, cross-cultural research, cross-national comparisons, harmonization, measurement

1 | INTRODUCTION

By 2050, ≈75% of Alzheimer's disease and related dementias (ADRD) cases are projected to occur in low- and middle-income countries.¹ This projected shift in the ADRD case burden and associated global population aging has, in part, led to the rise of cross-nationally harmonized studies of aging around the world. The Harmonized Cognitive Assessment Protocol (HCAP) is a harmonized cognitive battery that aims to provide a comparable yet flexible instrument with which to measure the cognitive function of older adults around the world.² It was developed by investigators from the U.S. Health and Retirement Study (HRS) in collaboration with several of its International Partner Studies and has been fielded recently in sub-samples of these studies.^{2,3} The HCAP includes neuropsychological test items selected from existing, validated cognitive test batteries alongside an informant interview to provide data for the classification of mild cognitive impairment and dementia.² For more details on the HCAP battery content and its comparability across HCAP studies, see Langa et al. and Gross et al.^{2,3}

The HCAP and its implementation in the HRS International Partner Study network expand upon related initiatives, such as the 10/66 dementia studies in low- and middle-income countries⁴ and the post hoc data harmonization efforts by the Integrative Analysis of Longitudinal Studies on Aging (IALSA) network, the Cohort Studies of Memory in an International Consortium (COSMIC) collaboration, and others.^{5–7} The HCAP studies represent the novel implementation of a harmonized cognitive battery within an existing network of population-representative cohorts with harmonized designs and measures to further facilitate harmonization of data on cognitive function, impairment, and dementia and their risk factors across diverse high-, middle-, and low-income countries.² As of late 2023, the HCAP has been implemented in 18 countries around the world, with plans for future HCAP administration underway in at least 6 countries (Figure 1). These existing and planned HCAP studies represent ≈75% of the global population ≥65 years of age, such that the HCAP will be a major data resource for understanding the nature of cognitive aging globally.

The HCAP data may be used by researchers for multiple purposes. An important use of the data is to provide comparable prevalence and incidence estimates of mild cognitive impairment and dementia across different countries. Another important purpose is to leverage cross-national variation in suspected risk and protective factors to gain comprehensive insights into dementia etiology around the world. Regardless of the research question, researchers wishing to use the HCAP or other harmonized data for cross-national compar-

isons of risk factor associations with later-life cognitive outcomes must carefully consider the theoretical grounding of their question and the methodological options available for answering it. Although there is a precedent for conducting cross-national comparisons for other health outcomes such as obesity,⁸ unique considerations are necessary for cross-national comparisons using continuous cognitive scores, such as those generated by the HCAP battery. Language, literacy, numeracy, and cultural differences across populations may bias assessments of cognitive function when test items are improperly selected for populations, improperly translated into different languages, or inappropriately adapted across educationally and culturally diverse contexts.^{9,10} There is a long and harmful history of intelligence testing that is inextricably tied to inaccurate and racist concepts of global “racial” differences and hierarchies.¹¹ Global cognitive aging researchers would do well to be aware of and to avoid reifying this type of approach when planning their studies and interpreting their findings. As population-based data on later-life cognitive function around the world become increasingly publicly available, it is imperative that researchers treat these data fairly, and that they avoid interpretations of results that attribute observed between-group differences in cognitive outcomes or their risk factors to innate differences between people or populations based on their assumed national or other origin.

The aim of this perspective is to discuss theoretical and methodological considerations and challenges and to describe a set of recommended best practices for researchers conducting cross-national comparisons of risk and protective factor associations with later-life cognitive function using data from the HCAP batteries. The principles described in this article are applicable not only to HCAP data users but to any researcher aiming to integrate or compare harmonized data on cognitive outcomes and their risk and protective factors across diverse populations.

2 | THEORETICAL CONSIDERATIONS

The theoretical considerations discussed in this section refer to the thought required in developing a scientific rationale, research question, and hypothesis for a cross-national comparison analysis. A set of recommended best practices regarding these considerations is outlined in Table 1 as a series of questions that researchers may ask themselves when designing a cross-national comparison analysis of later-life cognitive outcomes.

TABLE 1 Recommended best practices for cross-national comparisons of risk factor associations for later-life cognitive outcomes using harmonized cognitive function data.**Theoretical considerations:**

- **Carefully consider the purpose of the comparative research question.**
 - Is a cross-national comparison necessary to answer the research question or better contextualize the data or results at hand?
 - What is the hypothesis regarding cross-national differences?
- **Incorporate knowledge of relevant social, cultural, economic, political, and historical contextual factors of the countries under study.**
 - Is the research question appropriately informed by relevant country-specific contextual factors?
 - Is the interpretation of results appropriately informed by relevant country-specific contextual factors?
 - Are local collaborators needed to provide appropriate expertise and experience?

Methodological considerations:

- **Ensure that the primary exposure variable can be harmonized.**
 - Do the measured exposure variables conceptually represent the same construct across countries?
 - Do any measurement differences prohibit harmonization of the exposure variable?
 - If the exposure variable cannot be harmonized, what are the implications for the project?
- **Ensure that data on desired model covariates are available for all countries under study.**
 - Does the same confounding structure apply to the research question across all countries under study?
 - Are data available on all suspected confounders within each country?
 - If the suspected confounding structure varies across countries, will country-specific models with unique covariate sets be sufficient to answer the research question of interest?
- **Choose the most appropriate modeling strategy for the research question and data.**
 - Is a pooled analysis or parallel analyses the most appropriate?
 - If a pooled analysis is conducted, are fixed effects for country sufficient to answer the question?
 - If a pooled analysis with fixed effects for country is conducted, have interaction terms between confounders and country been tested or the outcome residualized for country-specific confounders?
 - If a pooled analysis is conducted and random effects preferred, is the group-level *N* sufficient?
 - If a pooled analysis is conducted, what is the desired approach to handling sampling weights?

and dementia risk,¹² yet unresolved concerns remain regarding the causality of this association due to the strong potential for confounding by early-life socioeconomic conditions or early-life cognitive ability as common causes of educational attainment and later-life dementia risk.^{16–18} Estimating the education–cognitive function association across countries where the early-life drivers of entry into and progression through the educational system vary would provide an opportunity to triangulate results and evaluate their consistency across settings with different confounding structures.

Researchers aiming to use the HCAP or other cognitive function data that have been harmonized across countries or populations should consider the purpose of their research question. For example, is the purpose of the comparative research question to understand whether a suspected association consistently holds in direction and magnitude across different settings? If so, what does the researcher hypothesize about differences in the association under study across settings, and why? Setting up this kind of comparative research question and hypothesis requires substantial background knowledge of the social, political, economic, cultural, and historical contexts of the countries or populations under study, in addition to the design and methods of the data sources being used. For example, a researcher conducting a cross-nationally comparative analysis of later-life cognitive function utilizing data from the Longitudinal Study of Aging in India–Diagnostic Assessment of Dementia (LASI-DAD) study in India should be aware of social norms regarding gender and caste in India,

the diversity of language and religions in India, and the country's history of partition and independence from British colonial rule. Cognitive function and dementia risk are heavily socially patterned,^{19–25} and knowledge of a country's history and context should be drawn upon to inform the research question, help select and craft variables for analysis, and support a rich and thoughtful interpretation of results. If researchers do not have such knowledge themselves, they should consider involving local collaborators with relevant knowledge and experience.²⁶ Indeed, for etiological triangulation as described earlier to be successful, researchers must have an informed idea of the confounding structure for a given association of interest under study within each country of interest. Without a clear idea of the relevant context of each country as it relates to shaping cognitive function and dementia risk of the older adults represented by each HCAP study, a high-quality cross-national comparison study is not possible.

2.2 | A warning against essentialist interpretations of cross-national cognitive data

Because the HCAP data are intended to be a publicly available data resource, it is important to recognize that there are many possible interpretations of country-level differences in the means and distributions of cognitive function data. Although the HCAP is designed to be a harmonized cognitive battery, the items in the HCAP battery have

undergone translation into several different languages and adapted for cultural appropriateness and for non-literate and non-numerate populations. For example, an object naming (language) item that asks the respondent to name a cactus was adapted to refer to a coconut in Wave 1 of the LASI-DAD study in India, as cacti are not indigenous to India. These adaptations may alter the difficulty, meaning, interpretation, or scoring of items, which could lead to observed country-level differences in scores being due to incommensurate measurement, rather than real population differences in cognitive function.¹⁰

The HCAP battery data from six countries (United States, England, Mexico, China, India, and South Africa) have undergone a post hoc process of pre-statistical and statistical harmonization to account for these adaptations.^{3,9} In the pre-statistical harmonization step, each HCAP cognitive test item was determined by a panel of neuropsychologists and epidemiologists in collaboration with HCAP study fieldwork teams to be “comparable” or “non-comparable” across countries, based on their respective adaptations.^{3,9} In the statistical harmonization step, factor scores for general cognition and domains of memory, language, orientation, and executive function were constructed using item response theory-based models that were serially estimated across countries with item parameters for “comparable” items held constant and item parameters for “non-comparable” items freely estimated.³ Figure 2 presents the means and distributions of these statistically harmonized factor scores.³ These scores are in the process of becoming publicly available for data users through HRS International Partner Study websites and the Gateway to Global Aging Data website.

Although country-level differences in mean scores and their distributions are evident from Figure 2, it is important to guard against essentialist interpretations of these country-level differences that attribute them to inherent differences between people or populations. Rather, there are global differences in the totality of risk and protective factors for later-life cognitive function and decline that people experience across their lifetimes. The degrees to which the impacts of these factors are mitigated or exacerbated by contextual factors such as policies and social and physical environments is likely to vary across countries.^{25,27} Comparisons of risk-factor associations across diverse country contexts, rather than direct comparisons of means, are thus valuable for understanding how these factors play out across different settings to affect later-life cognitive function, decline, and dementia risk.²⁸

It is important to note that despite test item adaptations and subsequent statistical harmonization, as in the factor scores presented in Figure 2, there may be cultural, linguistic, and other assumptions “baked in” to many cognitive test items that cannot be corrected for, and which may favor English-speaking, highly educated, and Western populations. Such assumptions can never be fully empirically quantified after HCAP battery data collection, as the statistical procedures required to harmonize tests rely on untestable assumptions.⁹ HCAP data users should consider these possibilities as they interpret their findings. Additional sensitivity analyses can be helpful to identify potential areas of measurement error or bias, which could be achieved through approaches such as manipulating scores to have variable floor and ceiling effects, using alternative linking approaches such as

equipercentile equating, or using tobit regression where floor or ceiling effects are observed.^{29,30}

3 | METHODOLOGICAL CONSIDERATIONS

Once a comparative research question and hypothesis are settled upon and well justified based on theory and prior evidence, the methodological considerations discussed in this section become relevant for conducting cross-national comparison analyses (Table 1). Table 1 presents these considerations as a series of questions that researchers may ask themselves when designing a cross-national comparison analysis of later-life cognitive outcomes. These considerations are discussed with the assumption that the cognitive outcome variable(s) of interest are appropriately measured and harmonized as best as possible according to modern methodology, as with the HCAP data or other cognitive function assessments that are harmonized across settings or populations. Further discussion of modeling approaches for cross-national comparisons of aging data broadly can be found in Kapteyn,³¹ with an emphasis on studying the causal effects of policies on aging-related health outcomes. Here, we complement his discussion, focusing on aspects that received less or no attention in Kapteyn's paper, with a focus on harmonized cognitive outcomes.

3.1 | Ensure that the primary exposure variable can be harmonized

This consideration is not simply a data feasibility issue. Harmonization of exposure variables requires an understanding of the construct of the exposure across the countries or populations under study. Several exposures of interest for understanding later-life cognitive outcomes may have different meanings across different settings.²⁸ For example, educational attainment is typically measured in research studies either as years of education or degree level obtained. The mechanisms through which education is thought to promote cognitive health include the conferral of knowledge, the conferral of skills such as literacy and numeracy, providing enhanced opportunities for employment and earnings, shaping social networks, and influencing health-related behaviors.^{32–34} Educational attainment may be differentially effective in conferring these outcomes across different settings where the quality, content, and access to educational curricula vary, and where post-educational opportunities for employment and social mobility may vary as well. Hence, the construct of educational attainment as typically measured in research studies may not be similarly associated with cognitive function across different contextual settings. Although some researchers may consider this a violation of the consistency assumption for causal inference, violations of the consistency assumption may in fact be informative.²⁸ In the case of cross-national cognitive aging research, understanding consistency violations can improve our understanding of global heterogeneity in commonly used research constructs, support the refinement of measurements of these constructs across diverse global settings,

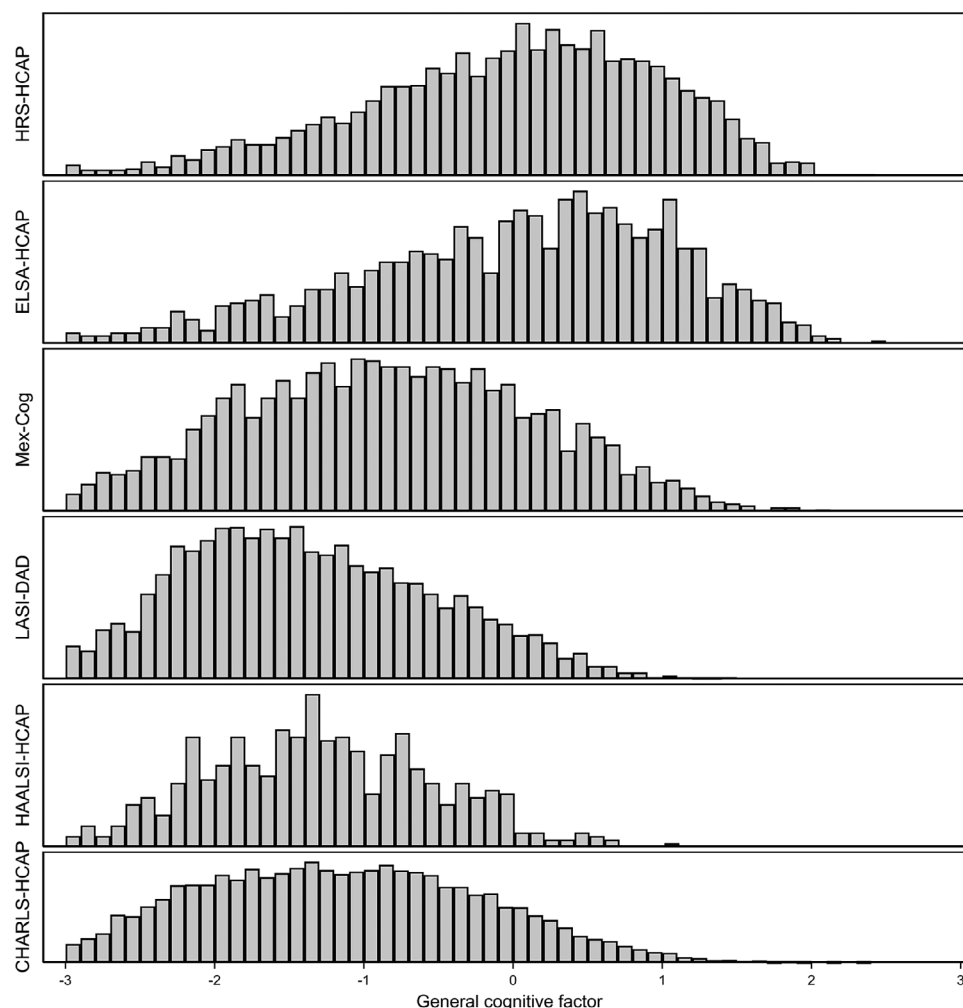


FIGURE 2 Distributions of general cognitive factor scores for HCAP data in six studies: Health and Retirement Study Harmonized Cognitive Assessment Protocol (HRS-HCAP, United States), English Longitudinal Study of Aging Healthy Cognitive Aging Project (ELSA-HCAP, England), Mexican Cognitive Aging Ancillary Study to the Mexican Health and Aging Study (Mex-Cog, Mexico), Longitudinal Study of Aging in India—Diagnostic Assessment of Dementia (LASI-DAD, India), Health and Aging in Africa: A Longitudinal Study of an INDEPTH Community in South Africa Dementia Study (HAALSI-HCAP), and China Health and Retirement Longitudinal Study (CHARLS-HCAP, China). The general cognitive factor scores are standardized to the HRS-HCAP distribution, which has a mean of 0 and standard deviation of 1, and for each country they incorporate all cognitive test items included in its HCAP protocol.

and allow the generation of new hypotheses regarding the various mechanisms influencing cognitive outcomes in varying populations and places.

3.2 | Ensure that data on desired covariates are available for all countries under study

The selection of model covariates should be informed by the suspected confounding structure of the exposure–outcome relationship for each country under study, the availability of data on desired confounding variables, and the choice of modeling approach. Confounder selection based on a priori theory and empirical evidence and guided by the use of directed acyclic graphs (DAGs) is the preferred approach in many health science disciplines.³⁵ Researchers conducting cross-national

comparison analyses should consider whether the same confounding structure applies to their research question of interest across all countries under study, or whether there should be country-specific confounding variables. A good example is that of race and/or ethnicity as a confounding variable for many research questions of interest regarding cognitive function and dementia risk.²⁵ Because race and ethnicity are socially constructed and reflect experiences of different forms of racism,³⁶ variables representing race and/or ethnicity across different countries and cultures have different sociological meanings as well as different categories that cannot easily be harmonized. In addition, differential response propensities across countries on scale-based self-reported measures, such as self-rated health, may complicate the harmonization of such measures across countries.³⁷

The suspected presence or absence of different confounding structures across countries has implications for the choice of modeling

approach. If there is cross-country variation in the desired set of confounding variables and their ideal operationalizations, the primary outcome can be residualized for the country-specific set of confounding variables. Otherwise, pooled analyses may leave results subject to residual confounding that varies in magnitude across countries. If adjustment is infeasible, parallel analyses of country-specific stratified models may be preferred. Of note, even if the same confounding structure is thought to apply across country settings, the magnitudes and directions of confounding relationships may vary across countries. In pooled models with fixed effects for country, allowing harmonized confounding variables to interact with an indicator variable for country would allow them to be differentially related with the outcome across countries, a strategy which could help to improve confounder control and model fit.

3.3 | Choose the most appropriate modeling strategy for the research question and data

3.3.1 | Pooled analysis versus parallel analysis?

A primary decision point for modeling cross-national comparisons, aside from the type of model to be used, is whether to do a pooled analysis of data from multiple countries in a single model or a parallel analysis of such data in stratified country-specific models.⁷ The choice of pooled or parallel analyses depends on the research question at hand, the researcher's hypotheses about the consistency of associations across countries, and the feasibility of pooling data based on the nature of the exposure of interest and desired model covariates. Parallel analyses of stratified models allow for greater flexibility in model specification and the selection of model covariates across samples, at the potential cost of comparability. Parallel analyses allow the qualitative comparison of associational point estimates and can be used to assess the consistency of associations across countries and might be preferred during initial stages of model building prior to pooled analyses. This approach has been used by the IALSA network, for the integration of data more broadly from longitudinal aging studies around the world.⁷ Pooled analyses allow for statistical significance testing of interaction terms between exposure variables and country, but they have stricter variable harmonization requirements and may be subject to differential degrees of residual confounding and model misspecification across countries.

3.3.2 | Fixed effects or random effects for country?

If a pooled analysis is conducted, researchers may run a standard least squares regression model with fixed effects for country to adjust for unmeasured country-level differences and interact these country fixed effects with other variables, as appropriate, or they may run a random effects model to model heterogeneity between countries while accounting for the within-country clustering of cognitive outcome data. The fixed-effects approach may be most appropriate if

the researcher is interested in the specific countries in the sample, if the number of countries is small, or if the set of countries cannot be viewed as representative of a population of countries. Dependence of observations at the country level can be handled using a cluster-robust variance estimator, although appropriate corrections to variance estimation are needed because model standard errors will be biased downwards with this approach if the number of countries is small (e.g., less than 30).³⁸

Random-effects models are most suitable when there is a sufficient sample size of countries, and they can be viewed as a representative sample of some population of countries. However, random-effects models are generally less robust to confounding than fixed-effects models. Multilevel models are an extension of random-effects models that can more flexibly model heterogeneity between countries. Some researchers may be interested in decomposing variance in later-life cognitive function to between-person and between-population variance or identifying potential ecological country-level influences on later-life cognitive function, such as gender inequality, income inequality, or other macro-contextual or compositional factors. Multilevel models with individuals clustered within countries are a possible approach for answering these kinds of research questions, but the number of available countries may be insufficient, especially for examining country-level ecological variables.³⁹ An alternative approach for decomposing variance and examining contextual variables at the group level would be to use sub-national geographies such as states, provinces, or counties in countries where such divisions are theoretically and empirically feasible for a given research question.

3.3.3 | Handling HCAP study sampling weights

Sampling weights are desirable or necessary for inclusion in most substantive analyses of risk and protective factors using HCAP data as there are disparate approaches to sampling across parent studies of the HRS International Partner Study network, and because several HCAP study samples are not random samples of their parent study.² In pooled analyses of HCAP battery data or any other harmonized cognitive outcome data that require the application of sampling weights, such weights need to be standardized consistently across samples so that appropriate inferences can be drawn from weighted models. In general, there are three options for handling HCAP study sampling weights in pooled cross-country models: (1) Use unstandardized weights that sum to the size of the underlying general population of each country (i.e., the sum of the weights for each sample is equal to the target population of each country); (2) Standardize weights so that each HCAP study sample is weighted equally in the analysis (i.e., the sum of the weights for each sample is equal to an arbitrary constant such as 1 or 3.14); and (3) Standardize weights so that the contribution of each HCAP study sample is proportionate to its sample size (i.e., the sum of the weights for each sample is equal to the sample size). The desired approach for handling sampling weights will depend on the researcher's question and goal of modeling. In the first option, countries with very large populations such as China or India will have domineering effects on the resulting

pooled regression coefficient, but this option may be appropriate for questions aiming to describe characteristics of such populations. In the second option, the sizes of individual HCAP samples are ignored in the weighting, which may not be preferable due to the influence of sample size on precision. In the third option, sample sizes may differ enough to make within-country standard errors meaningfully different, but not so much as to minimize the influence of small countries on pooled regression coefficients. We believe that the third option for handling sampling weights in pooled analyses of HCAP data is the most appropriate for most research questions, but researchers should consider their goals of modeling, hypotheses about the consistency of associations across countries, the potential influence of sample sizes or population sizes on results, and the desirability of such influences.

4 | CONCLUSION

The increasing public availability of harmonized data on later-life cognitive function brings about powerful opportunities for expanding the global dementia evidence base. However, cognitive function is an outcome that is highly sensitive to linguistic, cultural, social, and educational factors in its measurement, and differences in such factors across diverse countries around the world have the potential to introduce bias into measurement and interpretation. Researchers should carefully consider the purpose of their cross-national comparison analyses of later-life cognitive function. Careful consideration includes being thoughtful about the context and history of the countries included in analyses, even for research questions that are not explicitly social in nature. Researchers should also be informed about the methods and measurements of the data sources being used. Resources such as the Gateway to Global Aging are valuable for gaining insights into cross-nationally harmonized data,⁴⁰ and local collaborators should also be involved, where appropriate. We caution against attributing observed differences in cognitive function and other cognitive health outcomes to innate differences between people or populations based on their assumed national or other origins. Despite these challenges, the HCAP battery brings incredible potential for new insights into the etiology of dementia and expansion of the evidence base to represent some of the most globally disadvantaged populations that will bear a heavy share of global dementia prevalence in the coming decades. The research community has the important privilege and power of being able to use these data for the public good in reducing the global dementia burden in an equitable fashion.

ACKNOWLEDGMENTS

The authors have nothing to report. This work was made possible by funding from the National Institute on Aging at the U. S. National Institutes of Health, under grant numbers: R01AG070953 (Kobayashi, Gross), R01AG030153 (Lee, Gross, Adar), U24AG065182 (Langa, Weir), K23AG080035 (Briceño), and R00AG070274 (Zhang).

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

ORCID

Lindsay C. Kobayashi  <https://orcid.org/0000-0003-2725-3107>

REFERENCES

1. Prince MJ, Wimo A, Guerchet MM, Ali GC, Wu Y-T, Prina M. World Alzheimer Report 2015 - The Global Impact of Dementia: An analysis of prevalence, incidence, cost and trends. London: Alzheimer's Disease International, 2015. 84 p.
2. Langa KM, Ryan LH, McCammon RJ, et al. The health and retirement study harmonized cognitive assessment protocol project: study design and methods. *Neuroepidemiology*. 2020;54(1):64-74.
3. Gross AL, Li C, Briceño EM, et al. Harmonisation of later-life cognitive function across national contexts: results from the harmonized cognitive assessment protocols. *Lancet Healthy Longev*. 2023;4(10):e573-e583.
4. Prince M, Ferri CP, Acosta D, et al. The protocols for the 10/66 dementia research group population-based research programme. *BMC Public Health*. 2007;7:165.
5. Sachdev PS, Lipnicki DM, Kochan NA, et al. COSMIC (Cohort Studies of Memory in an International Consortium): an international consortium to identify risk and protective factors and biomarkers of cognitive ageing and dementia in diverse ethnic and sociocultural groups. *BMC Neurol*. 2013;13(1):165.
6. Giorgio J, Tanna A, Malpetti M, et al. A robust harmonization approach for cognitive data from multiple aging and dementia cohorts. *Alzheimers Dement Amst Neth*. 2023;15(3):e12453.
7. Hofer SM, Piccinin AM. Integrative data analysis through coordination of measurement and analysis protocol across independent longitudinal studies. *Psychol Methods*. 2009;14(2):150-164.
8. Kim R, Kawachi I, Coull BA, Subramanian SV. Contribution of socioeconomic factors to the variation in body-mass index in 58 low-income and middle-income countries: an econometric analysis of multilevel data. *Lancet Glob Health*. 2018;6(7):e777-e786.
9. Briceño EM, Arce Rentería M, Gross AL, et al. A cultural neuropsychological approach to harmonization of cognitive data across culturally and linguistically diverse older adult populations. *Neuropsychology*. 2023;37(3):247-257.
10. Kobayashi LC, Gross AL, Gibbons LE, et al. You say tomato, I say radish: can brief cognitive assessments in the U.S. Health Retirement Study be harmonized with its international partner studies? *J Gerontol B Psychol Sci Soc Sci*. 2021;76(9):1767-1776.
11. Gould SJ. *The Mismeasure of Man*. Norton; 1981.
12. Livingston G, Huntley J, Sommerlad A, et al. Dementia prevention, intervention, and care: 2020 report of the Lancet Commission. *Lancet Lond Engl*. 2020;396(10248):413-446.
13. Glymour MM, Whitmer RA. Using cross-cultural studies to improve evidence on dementia prevention: lessons from the special issue sponsored by the International Research Network on Dementia Prevention (IRNDP). *J Alzheimers Dis*. 2019;70(Suppl 1):S5-S10.
14. Rose G. Sick individuals and sick populations. *Int J Epidemiol*. 2001;30(3):427-432. discussion 433-434.
15. Lawlor DA, Tilling K, Davey Smith G. Triangulation in aetiological epidemiology. *Int J Epidemiol*. 2016;45(6):1866-1886.
16. Glymour MM, Manly JJ. Compulsory Schooling Laws as quasi-experiments for the health effects of education: reconsidering mechanisms to understand inconsistent results. *Soc Sci Med*. 2018;214:67-69.
17. Seblova D, Fischer M, Fors S, et al. Does prolonged education causally affect dementia risk when adult socioeconomic status is not altered? A Swedish natural experiment in 1.3 million individuals. *Am J Epidemiol*. 2021;190(5):817-826.
18. Thomas MD, Calmasini C, Seblova D, et al. Postsecondary education and late-life cognitive outcomes among Black and White participants in the project talent aging study: can early-life cognitive skills account

- for educational differences in late-life cognition? *Alzheimer Dis Assoc Disord*. 2022;36(3):215-221.
19. Glymour MM, Manly JJ. Lifecourse social conditions and racial and ethnic patterns of cognitive aging. *Neuropsychol Rev*. 2008;18(3):223-254.
 20. Mayeda ER, Glymour MM, Quesenberry CP, Whitmer RA. Inequalities in dementia incidence between six racial and ethnic groups over 14 years. *Alzheimers Dement J Alzheimers Assoc*. 2016;12(3):216-224.
 21. Gilsanz P, Mayeda ER, Glymour MM, Quesenberry CP, Whitmer RA. Association between birth in a high stroke mortality state, race, and risk of dementia. *JAMA Neurol*. 2017;74(9):1056-1062.
 22. Lee J, Meijer E, Langa KM, et al. Prevalence of dementia in India: national and state estimates from a nationwide study. *Alzheimers Dement J Alzheimers Assoc*. 2023;19(7):2898-2912.
 23. Saenz JL, Adar SD, Zhang YS, et al. Household use of polluting cooking fuels and late-life cognitive function: a harmonized analysis of India, Mexico, and China. *Environ Int*. 2021;156:106722.
 24. Kobayashi LC, Farrell MT, Langa KM, Mahlalela N, Wagner RG, Berkman LF. Incidence of cognitive impairment during aging in rural South Africa: evidence from HAALSI, 2014 to 2019. *Neuroepidemiology*. 2021;55(2):100-108.
 25. Adkins-Jackson PB, George KM, Besser LM, et al. The structural and social determinants of Alzheimer's disease related dementias. *Alzheimers Dement J Alzheimers Assoc*. 2023;19(7):3171-3185.
 26. Khan M, Abimbola S, Aloudat T, Capobianco E, Hawkes S, Rahman-Shepherd A. Decolonising global health in 2021: a roadmap to move from rhetoric to reform. *BMJ Glob Health*. 2021;6(3):e005604.
 27. Tollman SM, Norris SA, Berkman LF. Commentary: the value of life course epidemiology in low- and middle-income countries: an ageing perspective. *Int J Epidemiol*. 2016;45(4):997-999.
 28. Rehkopf DH, Glymour MM, Osypuk TL. The consistency assumption for causal inference in social epidemiology: when a rose is not a rose. *Curr Epidemiol Rep*. 2016;3(1):63-71.
 29. Gross AL, Kueider-Paisley AM, Sullivan C, Schretlen D, International Neuropsychological Normative Database Initiative. Comparison of approaches for equating different versions of the mini-mental state examination administered in 22 studies. *Am J Epidemiol*. 2019;188(12):2202-2212.
 30. Nichols EL, Cadar D, Lee J, Jones RN, Gross AL. Linear linking for related traits (LLRT): a novel method for the harmonization of cognitive domains with no or few common items. *Methods San Diego Calif*. 2022;204:179-188.
 31. Kapteyn A. What can we learn from (and about) global aging? *Demography*. 2010;47(Suppl 1):S191-S209.
 32. Ross CE, Wu C-I. The links between education and health. *Am Sociol Rev*. 1995;60(5):719-745.
 33. Zimmerman E, Woolf SH. Understanding the relationship between education and health. *NAM Perspect*. Discussion Paper, National Academy of Medicine, Washington, DC. 2014. doi:<https://doi.org/10.31478/201406a>
 34. Richards M, James SN, Sizer A, et al. Identifying the lifetime cognitive and socioeconomic antecedents of cognitive state: seven decades of follow-up in a British Birth Cohort Study. *BMJ Open*. 2019;9(4):e024404.
 35. Tennant PWG, Murray EJ, Arnold KF, et al. Use of directed acyclic graphs (DAGs) to identify confounders in applied health research: review and recommendations. *Int J Epidemiol*. 2021;50(2):620-632.
 36. Bailey ZD, Krieger N, Agénor M, Graves J, Linos N, Bassett MT. Structural racism and health inequities in the USA: evidence and interventions. *Lancet Lond Engl*. 2017;389(10077):1453-1463.
 37. Jürges H. True health vs response styles: exploring cross-country differences in self-reported health. *Health Econ*. 2007;16(2):163-178.
 38. Cameron AC, Gelbach JB, Miller DL. Bootstrap-based improvements for inference with clustered errors. *Rev Econ Stat*. 2008;90(3):414-427.
 39. McNeish DM, Stapleton LM. The effect of small sample size on two-level model estimates: a review and illustration. *Educ Psychol Rev*. 2016;28(2):295-314.
 40. Lee J, Phillips D, Wilkens J, Gateway to Global Aging Data Team. Gateway to global aging data: resources for cross-national comparisons of family, social environment, and healthy aging. *J Gerontol B Psychol Sci Soc Sci*. 2021;76(Suppl 1):S5-S16.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Kobayashi LC, Jones RN, Briceño EM, et al. Cross-national comparisons of later-life cognitive function using data from the Harmonized Cognitive Assessment Protocol (HCAP): Considerations and recommended best practices. *Alzheimer's Dement*. 2024;20:2273-2281. <https://doi.org/10.1002/alz.13694>