Using Quantitative Methods to Understand Gender and Intersectionality in Urban Health Research: The MAIHDA approach

Qualitative approaches have been valuable in understanding how gender and other individual factors such as ethnicity, disability, and employment status work together to influence health outcomes, a field known as intersectionality research. However, quantitative approaches can also be used to identify and measure the extent to which gender and other social factors combine to produce unfair health outcomes especially in complex urban or conflict-affected areas.

The problem with single-level regression analysis

In the past, quantitative intersectionality relied on single level regression analysis, that looked at each factor such as gender, disability, and employment status on its own and how they might relate to each other. Yet, this statistical approach faces some practical problems (Evans et al., 2020). First, the more factors included in the analysis, the more complicated the results become, which makes them harder to understand. Second, each factor is usually looked at separately, instead of seeing how they might work together to affect health outcomes. Third, when different factors are combined, some groups end up with few individuals in them, which makes it difficult to get reliable results. Finally, the method picks one group as the standard (e.g., males) in gender, which means other groups such as females are always compared to that group. This can make it seem like the standard group is the dominant or most important.

Moving to MAIHDA

To deal with these problems, researchers have started using a method known as Multilevel Analysis of Individual Heterogeneity and Discriminatory Analysis (MAIHDA) (Evans et al., 2024; Evans et al., 2020; Merlo, 2018). MAIHDA looks at how different social factors such as gender, ethnicity, and employment status work together and affect each other in a more realistic way, instead of just adding up their influence on health outcome one by one. This approach helps show how overlapping social factors lead to different levels of inequality, without treating any one group as the most important for comparison.

Using MAIHDA to understand urban health

Normally, MAIHDA shows both the overall differences between group of individuals and how much individuals within each group vary from the group's average (Evans et al., 2024). For example, in the study of Nairobi's informal settlements, MAIHDA was able to measure how the risk of diarrhea, fever, and cough varied between intersectional groups defined by the child's age, the head of household ethnic group and education, length of stay in informal settlements, health insurance status, religion, food security, and housing tenure; and which groups were more vulnerable (Kibuchi et al., 2024). It also revealed how much the risk of these health outcomes for each child under five differed from their group's average.

In a similar study in Bangladesh's informal settlements, children under-five living in specific informal settlement locations, with unemployed mothers, younger heads of households, and poor waste disposal practices, were more likely to suffer from fever, cough, and acute respiratory infections (Barua et al., 2023). These findings suggest that certain groups of children face structural challenges that require interventions to be accessible to everyone in these settings but also delivered with an intensity that matches the level of vulnerability in each group.

When and how to use MAIHDA

MAIHDA can be applied to both continuous and categorical outcomes (Evans et al., 2024), using either cross-sectional or longitudinal data (Evans et al., 2024). To use MAIHDA effectively, researchers need to carefully choose which social factors to include when forming intersectional groups, based either on theory (Kibuchi et al., 2024) or on individuals' real-life experiences (Sesay et al., 2025). For example, in Kibuchi et al. (2024), the social factors used to form intersectional groups were selected through research and expert discussions. On other hand, the study by Sesay et al. (2025), which looked on healthcare utilisation in informal settlements in Freetown, Sierra Leone, based its groupings on social factors identified by community members as most important to them, supported by the literature. This helps ensure that the results consider the many factors that work together to create inequality. In summary, MAIHDA is particularly useful in situations where inequalities are not easy to spot and there is limited data on overlapping social factors, such as gender based analysis.

To show how MAIHDA can be used to study gendered-related issues, we can look at differences in healthcare utilisation; a binary outcome (utilised healthcare: 1= yes, 0 = no) across three informal settlements in an urban setting. For brevity, we describe individuals' social positions using four factors: gender (male, female, other), age (18-35, 36-45, 46 - 55, and over 55), social class (low, middle, or high), and which informal settlement they live in (A, B, C). When we combine all these categories, we get 108 unique groups (3 × 4 × 3 × 3). For example, one group might include all individuals who are male, aged 18-35, from a low social class, and living in settlement A. Each group is given a unique identification number, and everyone in the same group is assumed to face similar challenges because they share the same social characteristics.

To identify and measure inequalities in health utilisation among the 108 social groups, we use a two-step approach with logistic multilevel models. These models are designed for yes or no outcomes, such as whether someone utilised healthcare or not (Evans et al., 2024).

Let say y_{ij} shows whether an individual i in group j utilised healthcare (1 = yes, 0 = no). The first model, called the null model, does not include any social factors. It only includes a random intercept for each group and helps us understand how much healthcare utilisation varies across the 108 groups. The null model is statistically written as:

$$logit(\pi_j) = \log\left(\frac{\pi_j}{1-\pi_j}\right) = \beta_0 + \mu_j$$

Here, β_0 is the overall average risk of health utilisation (called the intercept), and μ_j represents random variation for each group .This variation is assumed to follow a normal distribution with an average of 0 and a certain amount of spread, denoted as σ_μ^2. The null model helps estimate overall inequality by predicting of healthcare utilisation for each group and checking how much variation exists both within and between the groups. The group-level variation σ_μ^2 tells us how different the groups are from one another.

This variance is used to calculate Variance Partition Coefficient (VPC), which ranges from 0 to 1. VPC shows how much of the total variation in healthcare utilisation is due to differences between the intersectional groups (Evans et al., 2024). A VPC above 5% suggests that healthcare utilisation varies a lot depending on which group someone belongs to, indicating meaningful intersectional inequalities. For example, Kibuchi et al. (2024) found a VPC of 11.0% in the first model, where the health outcome was diarrhea. This means that the risk of diarrhea among children under five living in Nairobi informal settlements differed significantly depending on their intersectional group.

The next step involves extending the first model by adding the same social factors used to create intersectional groups, in our case, gender, age, social class, and settlements, as explanatory variables in the analysis. This helps separate the individual effects of each factor (known as additive effects) on healthcare utilisation from the combined impact of overlapping effects of 108 groups (also known as intersectional effects). Model 2 is written as follows:

$$logit(\pi_j) = \beta_0 + \beta_1 x_{1j} + \cdots + \beta_k x_{kj} + \mu_j$$

Here, $x_{1j}, \ldots, x_{kj}$ are dummy variables representing social factors, and $\beta_0, \ldots., \beta_k$ are coefficients that show how each factor affects healthcare utilisation. In this example, there are 10 coefficients in total, including the intercept: 2 for gender, 3 for age, 2 for social class, and 2 for settlements. Each factor includes a reference category, which is why the number of coefficients for each variable equals the number of categories minus one

In second model, the VPC shows how much of the variation in healthcare utilisation is due to the combined effects of overlapping social factors after accounting for the individual impact of each factor on its own (Evans et al., 2024). By comparing group-level differences obtained from both the first and second models, we can calculate the Proportional Change in Variance (PCV).

Interpreting MAIHDA Results

PCV tells us how much of the inequality between groups (captured from VPC) can be explained by individual social factors included in the model (Evans et al., 2024). A high PCV, close to 100% implies that most of the differences between groups are explained by social factors. On the other hand, a low PCV means that much of the inequality is still unexplained, and likely due to unique combinations of overlapping social factors.

For example, Kibuchi et al. (2024) found a VPC of 9.0% and PCV of 19.9% for the diarrhea outcome. This means that only about 20% of the inequality between groups could be explained by individual social factors included in the model. The remaining 80% was likely due to the overlapping effects of social factors, highlighting the importance of intersectional group differences in the risk of diarrhea among children under five.

Finally, a third model can be fitted that includes both social factors used to define intersectional groups and other explanatory variables associated with health utilisation, such as marital status, disability status, health insurance, among others. The VPC from this model shows how much variation in healthcare utilisation is due to the combined effects of overlapping social factors, after accounting for both the individual social factors used to define the groups and other explanatory variables.

The PCV indicates how much of the inequality between groups (captured by the VPC) can be explained by individual social factors and other explanatory variables included in the model. The interpretation of both VPC and PCV is similar to that in the second model, except that this model accounts for additional explanatory variables. In Kibuchi et al. (2024), found a VPC of 9.7% and a PCV of 11.0% for the risk of diarrhea. This suggest that approximately 89.0% of the inequality between groups was likely due to intersectional groups.

To detect important patterns, results can be visualised by ranking stratum-level predictions and residuals from lowest to highest as shown in Figure 1. This offers deeper insights into the distribution of inequalities across continuum.

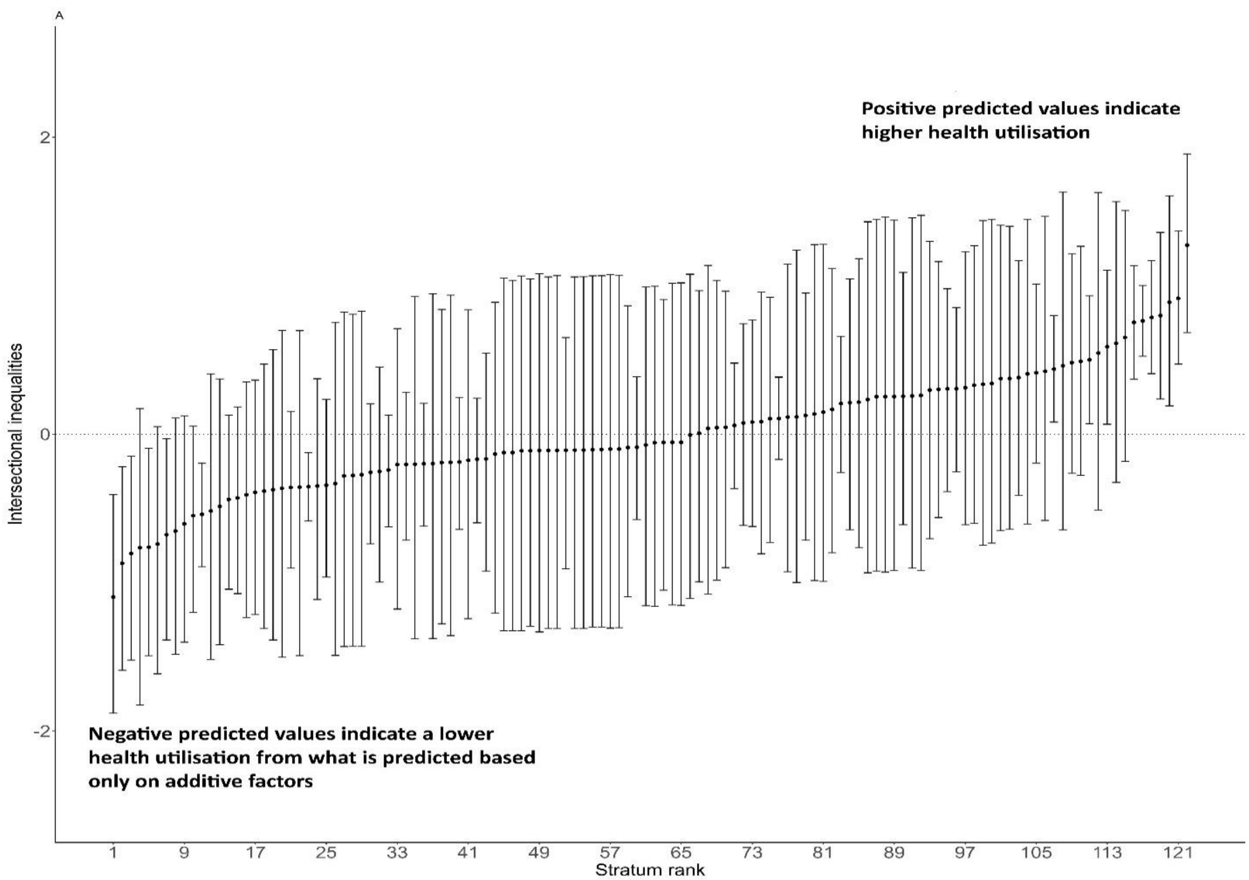

Figure 1: Estimated intersectional effect residual and their corresponding 95% credible intervals (CIs) for each intersectional group, ranked from the lowest to highest levels of health utilisation

Summary

In summary, MAIHDA is a powerful approach that comes from socio-epidemiology (Merlo, 2018) and intersectionality theory. It helps us quantify and measure how different social factors combine to create inequalities especially among the marginalised populations. This approach supports decision making for policies that are fairly in eliminating structural barriers, while also making sure extra support goes to those who are in most need.

Acknowledgement:

This introductory guide was written by Dr Eliud Kibuchi, Research Fellow, School of Health and Wellbeing, University of Glasgow.

References

Barua, P., Kibuchi, E., Aktar, B., Chowdhury, S. F., Mithu, I. H., Quayyum, Z., Filha, N. T. d. S., Leyland, A. H., Rashid, S. F., & Gray, L. (2023). The effects of social determinants on children's health outcomes in Bangladesh slums through an intersectionality lens: An application of multilevel analysis of individual heterogeneity and discriminatory accuracy (MAIHDA). *PLOS Global Public Health*, *3*(3), e0001588.

Evans, C. R., Leckie, G., & Merlo, J. (2020). Multilevel versus single-level regression for the analysis of multilevel information: The case of quantitative intersectional analysis. *Soc Sci Med*, *245*, 112499. https://doi.org/10.1016/j.socscimed.2019.112499

Evans, C. R., Leckie, G., Subramanian, S., Bell, A., & Merlo, J. (2024). A tutorial for conducting intersectional multilevel analysis of individual heterogeneity and discriminatory accuracy (MAIHDA). *SSM-population health*, 101664.

Kibuchi, E., Chumo, I., Kabaria, C., Elsey, H., Phillips-Howard, P., de Siqueira-Filha, N. T., Whittaker, L., Leyland, A. H., Mberu, B., & Gray, L. (2024). Health inequalities at the intersection of multiple social determinants among under five children residing Nairobi urban slums: An application of multilevel analysis of individual heterogeneity and discriminatory accuracy (MAIHDA). *PLOS Global Public Health*, *4*(2), e0002931.

Merlo, J. (2018). Multilevel analysis of individual heterogeneity and discriminatory accuracy (MAIHDA) within an intersectional framework. *Soc Sci Med*, *203*, 74-80. https://doi.org/10.1016/j.socscimed.2017.12.026

Sesay, I. J., Sesay, S., Wurie-Kamara, H., Tengbe, S. M., Fullah, S., Vangahun, D., Gandi, I., de Siqueira Filha, N. T., Lakshman, R., Conteh, A., Saidu, S., Koroma, B., Mansaray, B., Elsey, H., Whittaker, L., Dean, L., Georgi, N. W., Nganda, M., Refell, F., . . . Kibuchi, E. (2025). Intersectional inequalities in healthcare utilisation in informal settlements in Freetown, Sierra Leone: A multilevel analysis of individual heterogeneity and discriminatory accuracy (MAIHDA). *Under review: International Journal for Equity in Health*.