nature portfolio

Peer Review File

Patterns of maternal and child health services utilization and associated socioeconomic disparities in sub-Saharan Africa.

Corresponding Author: Ms Claire Najjuuko

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
Overall Impression
I would like to commend the authors on this insightful and methodologically robust study addressing maternal, newborn, and child health (RMNCH) intervention disparities in sub-Saharan Africa. The use of machine learning clustering to identify unique coverage subgroups represents a novel and valuable approach, which I highly appreciate. This research has significant potential to inform targeted interventions and improve maternal and child health outcomes in resource-constrained settings. However, I have several concerns that I believe, if addressed, would strengthen the manuscript further.
List of Concerns:
1. While the use of machine learning for clustering is innovative, it is not clear what additional insights this method provides over traditional approaches like composite indices. I encourage the authors to emphasize the unique contributions and implications of their methodology more explicitly.
2. The statistical section lacks the inclusion of basic estimation equations, particularly for the multinomial regression and clustering metrics. Adding these would enhance clarity and reproducibility.
3. It is unclear how the survey weights and the complex study design of the DHS data were accounted for in the analysis. Clarification is needed on whether these were incorporated into the clustering and regression models.
4. The manuscript discusses socioeconomic disparities but does not quantify inequality measures. I strongly recommend performing additional analyses on absolute and relative inequalities for education, area (rural vs. urban), and wealth quintiles across countries.
5. The DHS data are hierarchically structured, yet the manuscript does not specify how within- and between-cluster heterogeneity was addressed. A discussion on whether multilevel modeling or clustering accounted for this structure would add robustness to the analysis.
6. The discussion section would benefit from a stronger alignment with recent literature. While some references are cited, integrating findings from other studies to support or contrast your results would enhance their contextual relevance.
7. The manuscript identifies several issues but lacks practical, evidence-based solutions to address these disparities. Including actionable recommendations, such as community health worker programs, poverty alleviation strategies, or culturally tailored interventions, would add significant value.
8. The implications for policy are only briefly mentioned. Expanding on how these findings can inform policy decisions, resource allocation, and targeted intervention strategies is critical.
9. While radar plots effectively summarize subgroup characteristics, additional visuals like heatmaps by country could help readers better understand cross-country disparities.
10. A sensitivity analysis examining the impact of imputation on clustering results would improve robustness.
11. The limitations section could be expanded to address potential biases from self-reported data and the exclusion of key RMNCH indicators like immunization and dietary diversity. Discussing these would provide a balanced view of the study's scope.
12. The manuscript uses some technical terms like "trustworthiness" and "stability scores" that may be unfamiliar to readers. Briefly defining these terms or providing a supplementary glossary would enhance accessibility.

This manuscript is an excellent and timely contribution to the field, but addressing these concerns would strengthen its impact and utility. I look forward to seeing the revised version and the potential improvements it will bring.

(Remarks on code availability)

Reviewer #2

(Remarks to the Author)
The authors use machine-learning clustering on Demographic and Health Surveys from sub-Saharan Africa to identify distinct patterns in the utilization of services related to reproductive, maternal, newborn, and child health and other factors. They identify three groups and estimate associations between socioeconomic characteristics and the odds of falling in either of the groups.

The authors should highlight more what their study adds to existing evidence. While the approach is novel, the authors themselves cite existing studies that show socioeconomic inequalities in service uptake (5,10,11, line 94, page 3). What seems to be the new evidence is the identification of joint prevalence of "interventions", e.g. that there is a considerable population share that utilizes antenatal and postnatal care services but does not deliver in health facilities.

Main comments:
- The authors use the term "RMNCH interventions" to describe standard care services and individual behavior. The term "interventions" is somewhat misleading, especially in the abstract, as it suggests external, additional programs or services. A different term could provide a better description.
- Similarly, the authors use the term "coverage" which is strongly linked to the supply side, describing where services are offered and to what extent. However, given the individual level self-reported data, they actually analyze uptake or utilization of services and individual behavior. The second paragraph in the introduction differentiates between utilization and coverage, but the remaining manuscript and description of analysis only refers to coverage. One statement in the introduction (line 107, page 3) and one in the discussion (line 326, page 11) refer to utilization/uptake, but the focus on the term "coverage" is misleading. The authors do not measure the supply side of health care. This should be clarified throughout the manuscript and discussed.
- The authors categorize the factors included in the cluster analysis into "intervention indicators", "protective environmental factors" and "traditional factors" (described in Methods, Coverage indicators). Both title and abstract only refer to the "intervention indicators". Moreover, the category "traditional factors" is not fittingly named: tobacco use does not seem to be "traditional", neither is "decision-making participation". The aspect "marriage at a mature age" is coded based on first birth instead of marriage (Table 1 in Appendix 2) and should be labeled as such.
- Methods, Unsupervised learning: The authors state that they use UMAP for dimensionality reduction. However, they also state this approach worked "significantly better" than conventional multiple correspondence analysis, indicating that this was also implemented but is not presented. This is inconsistent and it is not clear what "significantly better" means and what criteria the authors used to decide on the approach.
- Methods, Statistical analysis: The authors describe that they first conducted bivariate analysis for each covariate separately and included those with p-value <0.25 in the multivariable model. However, the results of the bivariate analysis are not presented. Instead, the results section mentions a "full model adjusting for all other covariates" and Table 3 presents "unadjusted and adjusted results". It is not clear what "adjusted" and "unadjusted" here means. This inconsistency should be removed or clarified.
- In the discussion, the authors speak of "SES-driven disparities in coverage" and a "pro-rich coverage gap". These phrases seem imprecise and potentially misleading. The authors identify that SES is associated with group membership, i.e. SES is associated with uptake of services and certain behaviors. The phrase "SES-driven", however, implies a causal relation which is not established in this study. "Pro-rich coverage gap" suggests that services are catering to the wealthier (driven by supply side), whereas it seems more likely that the disparity arises from the wealthier individuals being more likely to utilize the services.
- In the discussion, the authors suggest that the machine-learning approach of the analysis has potential to "facilitate improvement in intervention design and delivery". This is not obvious, and the authors should clarify how improvement is expected to be facilitated.
- In the discussion, the authors recommend individual and community-level strategies for the low coverage subgroup. It is not obvious how these recommendations are drawn from the results. Also, references suggesting the success of such strategies should be added.

Minor comments
- The research questions presented in the last paragraph of the introduction could be more precise. The research question of the clustering analysis is not clearly phrased. The hypothesis about the association between socioeconomic status (not disparities) and utilization should be presented as association, not indicating a causal relationship using the word "drive". The link between subgroup proportions and U5M rates is not assessed thoroughly in this study, so should not be presented as a research question.
- Methods, Data source: The abbreviation "LMICs" is used without prior definition. As this is the only instance where this abbreviation is used, it should be spelled out.
- Methods, Statistical analysis: the dependent variable is defined as "the presence of the cluster", whereas it seems that it should be "whether the individual is inside/part of the cluster" or "group membership".
- The authors should also discuss whether their findings on the associations between SES and utilization of services are in line with the existing evidence.
- Table 3 should contain the sample size in each specification (N).
- Figure 2 should specify the source of U5M rates. The sample used in this study only includes births that ended in U5M and is insufficient to determine U5M rates, so the source must be other data.

(Remarks on code availability)

Reviewer #3

(Remarks to the Author)
I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

(Remarks on code availability)

Reviewer #4

(Remarks to the Author)
Comments for authors, Nature Communications manuscript NCOMMS-24-67107 “Patterns of maternal, newborn, and child health interventions coverage, and associated socioeconomic disparities in sub-Saharan Africa: A machine learning cluster analysis.”

General comments:
1. Many thanks to the authors for an interesting analysis relating to an important global health challenge. The summary across countries is important and noteworthy. While interesting, I feel the description of the methodology needs to be more closely linked to the application to better allow reproducibility, to better support the conclusions, and to increase the significance to the field.
2. I suggest replacing the phrase “traditional factor” by “cultural factor” throughout since “traditional factor” makes for some awkward phrases, e.g., line 76, “traditional practices (such as marriage at a mature age)”. I believe the “traditional practice” would be marriage at a young age, so this is difficult to parse. As an option, one could say “traditional practices (such as age of marriage)”, but I prefer the phrase “cultural factors”).
3. I suggest clearly stating in the overall description and in the Methods description that the authors are using machine learning/clustering for the national level variables and multinomial regression for the individual level demographic and SES variables. On my initial read, I found it difficult to clearly define which variables were part of clustering and which were part of the multinomial regression. Once the reader takes a step back, you can notice the national/individual distinction, but it would be really helpful for readers to make this clear from the beginning.
4. I also think it would improve readability to swap the order of Tables 1 and 2 and present these in the same order in which the analysis flows: Summary of national level indicators of RMNCH/environmental/cultural interventions/practices, then grouping of these indicators into summary clusters, then assessment of individual U5M by cluster and individual mother/child measures. Following a consistent order will help readers keep track of the logic of the analysis.
5. The description of the methods is difficult to follow and would be difficult to replicate. In the main manuscript, it is difficult to grasp the distinction between the machine learning component and the clustering-based dimension reduction. That is, it is unclear (at least to me) where UMAP ends and the two clustering methods begin in the main body of the manuscript. The supplemental information provides some additional insight, but with only 16 RMNCH (national-level) variables, the clustering algorithms seem to take care of the dimension reduction (from 16 to 3), indeed, UMAP results do not seem to be mentioned in the Results section of the main manuscript and the description of the UMAP methodology in the supplemental information is very generic and not particularly specific to the data set for this study until Appendix 4 which is very brief. What did UMAP reveal, what were the 250, 500, 1000, 1500 neighbors mentioned? Table 2 in the Supplemental Information is presented without description. I feel a more careful discussion (perhaps in the supplement) would aid in others being able to understand and replicate the work.
6. The K-means and hierarchical cluster weights shown in Figures 1a and 1b are remarkably and surprisingly similar. Indeed, the only visible differences between the two figures are legend elements “Lowest coverage: 1708” and “Lowest coverage: 1752” and “Medium coverage: 1655” and “Medium coverage: 1611”. Typically, I would expect to see some minor variation in details when using different clustering mechanisms. Could the authors confirm that the results are this close (perhaps highlighting those observations that shifted between lowest and medium coverage in the supplemental information)? Are these countries geographically close together?

Specific comments (main text):
1. Line 54. “decreased of being” to “decreased odds of being”.
2. Line 64. Sustainable Development Goal be capitalized.
3. Lines 73-74. Note that more detail will be provided regarding the specific RMNCH interventions considered later in the manuscript.
4. Line 86. “coverage in” to “coverage of”.
5. Lines 94-95. “how specific interventions tend to cluster together of co-occur”. This point is not clear in the abstract but is central to the analysis. I suggest making the motivation for clustering clearer in the abstract.
6. Lines 102-103. “exceed preconceived associations” to “extend beyond a preconceived set of potential associations”.
7. Line 106. “This enables to test” to “This enables us to test”.
8. Line 108. “aimed to show” to “aimed to describe”.
9. Line 109. “explain geographical” to “explain observed geographical”.
10. Line 136. “water sources; while tobacco” to “water sources. Tobacco” (start a new sentence for the cultural factors.

11. Line 145. “wealth tier” Is this defined in the DHS data?
12. Line 155. Does “mode imputation” mean you imputed the mode for any missing values? Was a sensitivity analysis done to see if results change with multiple imputation (particularly for marriage at a mature age)?
13. Line 165. “significantly better”. Please describe in what ways the results are better (and I suggest avoiding the term “significantly”).
14. Line 183. How is “incorrect clustering” defined?
15. Line 226. “interventions” to “intervention coverage”.
16. Line 231. “majority” to “a majority”?
17. Line 251. “Being separated/divorced and having a partner was associated…”. Does this refer to one classification (separated/divorced with a partner) or two ((1) separate/divorced and single, (2) having a partner)?
18. Line 293. “One reason for this is probability the ability…”. Can this be surmised from the clusters or is this a supposition separate from the data analysis?
19. Line 312. “majority” to “the majority”. Also, no comma after “postnatal care”.
20. Line 377. “are critical” to “provide insight”.

Specific comments (supplemental information):
1. Page 4, Appendix 3. While interesting, since the authors only used one of the listed distance measures, it would be sufficient to list that method (Jaccard similarity) and provide a general reference to the others (which would not apply to the authors’ categorical data). Also, I believe “jaccard” should be capitalized throughout.

(Remarks on code availability)
While I did not run the code, it is well documented within a Docker setting with python code and simulated data to illustrate the approach.

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
Comments:
The authors have provided a detailed and thoughtful response to all reviewer comments. They successfully clarified methodological approaches, especially the transition from K-means/UMAP to multilevel latent class analysis (MLCA), appropriately addressing concerns about the hierarchical nature of DHS data. The revised manuscript now clearly defines estimation procedures, accounts for survey weights, and includes inequality metrics (SII and RCI), enhancing both rigor and policy relevance. Key technical terms were refined, new figures added, and language made more precise, especially in differentiating between "utilization" and "coverage."

Moreover, the manuscript now presents stronger alignment with existing literature and includes practical, evidence-based policy recommendations. The added sensitivity analyses and expanded limitations further strengthen its credibility.

(Remarks on code availability)

Reviewer #2

(Remarks to the Author)
All comments have been addressed.

(Remarks on code availability)

Reviewer #3

(Remarks to the Author)
I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

(Remarks on code availability)

Reviewer #4

(Remarks to the Author)
Many thanks to the authors for a careful and thorough revision.

All of my initial concerns have been addressed and I found the revision clearer and easier to follow.

(Remarks on code availability)

Responses to reviewers' comments: Nature Communications

Dear Editor, Reviewers,

We are grateful to you for your thoughtful comments and suggestions. Your feedback has been invaluable in helping us to improve our manuscript. We have carefully considered your feedback and revised the manuscript. We believe that these revisions have addressed the concerns that were raised.

Please see the point-by-point response below. In this document, we will utilize the following formatting conventions:

- **Comment:** Reviewers' comment
- **Response:** Our response to the comment
- ***"Verbatim text with double quotes that we modified in the manuscript based on the feedback"*** where necessary.

Sincerely,
Claire Najjuuko
CLAIRE NAJJUUKO
PhD Scholar
Division of Computational and Data Sciences, McKelvey School of Engineering,
Washington University in St. Louis
One Brookings Drive, St. Louis, MO 63130-4899
Email: c.najjuuko@wustl.edu

Reviewer #1 (Remarks to the Author):

Overall Impression

I would like to commend the authors on this insightful and methodologically robust study addressing maternal, newborn, and child health (RMNCH) intervention disparities in sub-Saharan Africa. The use of machine learning clustering to identify unique coverage subgroups represents a novel and valuable approach, which I highly appreciate. This research has significant potential to inform targeted interventions and improve maternal and child health outcomes in resource-constrained settings. However, I have several concerns that I believe, if addressed, would strengthen the manuscript further.

1. **Comment:** While the use of machine learning for clustering is innovative, it is not clear what additional insights this method provides over traditional approaches like composite indices. I encourage the authors to emphasize the unique contributions and implications of their methodology more explicitly.

   Response: Thank you for this comment. Composite indices constructed through weighted averages of select coverage indicators assess overall proportion of the population that has access to these essential services. However, composite indicators often miss key information on the relationship between different indicators[1]. For example, maternal and child health services are rarely utilized in isolation, and the services that are co-utilized are not always obvious. These co-utilization patterns may remain hidden in standard composite approaches. Therefore, a clustering / latent class analysis is necessary to reveal these relationships beyond analysis using composite indices. We have clarified this unique advantage in the Introduction section to better motivate our methodological choice. See below:

   "Previous studies have employed various methods to assess utilization of maternal and child health services. Approaches include the use of summary measures such as the composite utilization index (CCI) and co-utilization indicators, as well as analyses of individual utilization indicators[8,9]. The CCI, calculated as a weighted average of utilization across a set of maternal and child health services, provides a single measure of overall utilization at the population level, allowing for comparisons across countries and tracking progress over time. Co-utilization indicators, on the other hand, measure the proportion of the target population receiving all or most of a set of essential services. Studies using these methods have revealed substantial socioeconomic inequalities in maternal and child health services utilization[5,10,11]. However, these methods do not reveal how different services are co-utilized, and they often miss key information on the relationship between different indicators[12]. Evidence suggests that utilization or non-utilization of these life-saving services rarely takes place in isolation[13]. Instead, different patterns of utilization tend to co-occur, and yet it is not always an obvious relationship. For example, studies have found that mothers who attend recommended four antenatal care visits do not always deliver in a health facility with a skilled birth attendance[14,15]. In addition, existing composite indices rarely take into account other environmental and cultural practices that are protective of maternal and child health. Yet this information is crucial for understanding the interplay of risk factors and utilization gaps experienced by specific subgroups, and for designing targeted interventions to address these disparities effectively.

   Our study takes a novel approach by conducting multilevel latent class analysis (MLCA) using individual-level data of births that ended in U5M based on utilization/uptake of 16 key RMNCH services and protective environmental and cultural factors (collectively referred to as maternal and child health services hereafter). This study makes important contributions to the literature. First, it defines data-driven patterns of co-utilization of maternal and child health services. Second, it examines the association between these utilization patterns and SES. Third, it investigates the inequalities in utilization of these life-saving services along dimensions of wealth, education, place

of residence, and employment. We employed MLCA to account for the hierarchical structure of our data. At the individual level, we categorized births into distinct service utilization subgroups, while at the country level, we identified clusters of nations with similar service utilization profiles. This dual-level approach captures both within-country heterogeneity and between-country similarities in service use patterns. By exploring SES predictors of subgroup membership, we tested the hypothesis that SES is associated with health service utilization patterns. Additionally, we tested the hypothesis that there exist disparities in service utilization along SES dimensions by obtaining relative concentration indices (RCI) and slope indices of inequality (SII) related to membership in the optimal utilization subgroup."

2. **Comment:** The statistical section lacks the inclusion of basic estimation equations, particularly for the multinomial regression and clustering metrics. Adding these would enhance clarity and reproducibility.
Response: Thank you for this comment. We have revised the statistical methods section to explicitly include the basic estimation equations for our multinomial regression and multilevel latent class analysis (MLCA). The following section has been included to describe the MLCA analysis:

The model was initialized using the k-means algorithm to establish low-level latent classes, followed by the application of the Expectation Maximization (EM) algorithm to optimize the model's fit to the data by maximizing the log likelihood formally defined as:

$$L(\Phi, \Pi, \omega) = \sum_{j=1} \log P(Y_j) \qquad (1)$$

and parameterized by Φ, the class-specific item-response probabilities for the low-level latent classes; Π, the conditional low-level class-membership probabilities for individuals given their country's high-level class; and ω, the distribution of countries in high-level latent classes. $P(Y_j)$ represents the probability of observing a specific combination of responses within each country, and is formally defined as:

$$P(Y_j) = \sum_{m=1}^{M} \omega_m \prod_{i=1}^{n_j} \sum_{t=1}^{T} \pi_{t|m} \prod_{h=1}^{H} P(Y_{ijh} \mid X_{ij} = t) \qquad (2)$$

In this context, $Y_{ij} = (Y_{ij1}, \dots, Y_{ijH})'$ is the vector of observed responses, Y_{ijh} is the response of individual $i = (1, \dots, n_j)$ in country $j = (1, \dots, J)$ on the h_{th} utilization indicator variable, with $h = 1, \dots, H$. A multilevel latent class model specifies the probability $P(Y_j)$ of observing a particular response configuration for each country j. The probability expression ω_m represents the probability that country j belongs to class m, $\pi_{t|m}$ denotes the conditional probability that an individual belongs to latent class t given that their country is in latent class m, and $P(Y_{ijh} \mid X_{ij} = t)$ denotes the probability of the response Y_{ijh} given the latent class t. Details of this model have been published elsewhere[41].

For the multinomial regression we have included the following in the manuscript:
"Following MLCA, a multivariable multinomial regression model was fitted, with low-level class membership as the dependent variable, and demographic and SES indicators as the independent variables. The multinomial logistic regression model for the three latent classes was formally defined as:

$$\log\left(\frac{P(Y_i = c)}{P(Y_i = 3)}\right) = \beta_{0c} + \beta_{1c}X_{i1} + \dots + \beta_{pc}X_{ip}, \qquad c = 1, 2 \qquad (6)$$

where Y_i is the low-level latent class for individual, i, X_{ip} are predictor variables, and β's represent regression coefficients."

3. **Comment:** It is unclear how the survey weights and the complex study design of the DHS data were accounted for in the analysis. Clarification is needed on whether these were incorporated into the clustering and regression models.
   Response: Thank you for this suggestion. We previously did not consider survey weights in the analysis. In this revision, we have modified the analysis to incorporate sample weights in our analysis. Given that we pool 31 countries' data, we rescale the sample weights within each country such that the total country weights is proportional to that country's population during the year of the survey[11]. We obtain population estimates from the United Nations World Population Prospects[12]. In general, this modification does not alter our clustering solution and conclusions made. Please see Comment 5 for details. We have modified our methods section to show these changes:

   "The final dataset contained 9,307 births meeting these criteria. In this pooled dataset, we rescaled the sampling weights (provided by DHS), such that each country's total weight is proportional to the country's population size during the year of survey. We obtained population estimates from the United Nations World Population Prospects*[12]*."

   "To properly account for the complex survey design of the DHS datasets, we implemented the analysis using the svydesign framework, incorporating sampling weights, cluster identifiers (primary sampling units), and stratification variables as specified in the DHS sampling methodology. This approach ensures that our estimates correctly reflect the hierarchical sampling structure and provides appropriate standard errors accounting for the design effects."

4. **Comment:** The manuscript discusses socioeconomic disparities but does not quantify inequality measures. I strongly recommend performing additional analyses on absolute and relative inequalities for education, area (rural vs. urban), and wealth quintiles across countries.
   Response: Thank you for this recommendation. We have included additional analyses to include inequality measures along wealth, education, place of residence and employment. We have made the following modifications to the Methods section and the Results sections respectively:

   "Measures of Inequality
   To quantify utilization inequalities across wealth, education, place of residence, and employment dimensions, we employed two key inequality measures: the SII and the RCI. SII is an absolute measure of inequality that shows the difference in estimated indicator values between the most-advantaged and most-disadvantaged subgroups, while accounting for all other subgroups - through appropriate regression modeling[46]. RCI on the other hand is a relative measure showing the gradient across population subgroups on a relative scale[46]. It indicates the extent to which an indicator is concentrated among disadvantaged or advantaged subgroups. Subgroups are weighted by their population share in both measures. For both measures, a value of zero indicates no inequality, positive values indicate a concentration of the indicator among the advantaged and negative values indicate a concentration of the indicator among the disadvantaged. Since the highest utilization subgroup represents the most desirable health-related characteristics, we applied these inequality measures to evaluate its distribution through the population across various SES groups. We implemented this analysis using the healthequal library in R. To properly account for the complex survey design of the DHS datasets, we specified the sampling weights, cluster identifier (primary sampling units), and stratification variable as specified in the DHS sampling methodology."

   "Inequality measures
   We examined utilization inequality using the SII and RCI across key SES indicators (see Table 5). Our findings revealed disparities, particularly pronounced in education, wealth, and place of residence where higher positive values of SII and RCI indicate that better health outcomes are

disproportionately concentrated among more affluent, educated, and urban populations. Employment-related disparities, although less marked, also highlight utilization inequalities.

Table 5: Inequality measures

Ranking variable	*Slope Index of Inequality*	*Relative Concentration Index*
Wealth quintile	***0.525 (0.474 - 0.576)***	***0.155 (0.138 - 0.171)***
Mother's education level	***0.679 (0.634 - 0.720)***	***0.197 (0.182 - 0.211)***
Place of residence	***0.599 (0.545 - 0.653)***	***0.120 (0.108 - 0.133)***
Mother's employment status	***0.128 (0.059 - 0.197)***	***0.028 (0.013 - 0.043)***

"

5. **Comment:** The DHS data are hierarchically structured, yet the manuscript does not specify how within- and between-cluster heterogeneity was addressed. A discussion on whether multilevel modeling or clustering accounted for this structure would add robustness to the analysis.
   Response: Thank you for this comment and suggestion. To account for the hierarchical structure and complex survey design of DHS data, we have included the survey design details in our multinomial analysis and inequality measures analysis. (Please see comment 3 and 4 above for changes in manuscript) We have also modified our clustering analysis from simply K-means to a multilevel latent class analysis which allows for specification of low-level and higher-level units. We note that even with the use of multi-level latent class analysis, the clustering solution we obtain is very similar to what we had previously. We calculated an adjusted rand index to compare the two clustering solutions, and we obtain a value of 0.8658 which indicates that both approaches have largely identified similar structures within our data. We have modified our Methods section to indicate these changes:

   "To identify data-driven patterns of maternal and child health service utilization, we used latent class analysis (LCA). LCA is a "person-centered" approach that groups individuals into discrete classes or groups based on a set of responses to a set of observable variables[40]. The LCA model assumes that observations are independent of each other. However, this assumption is often violated when the data have a multilevel structure, namely when lower-level units are nested in higher-level ones[41], as is our case (individuals are nested within primary sampling units nested within countries). For this reason, we employed the multilevel latent class analysis method[41] to account for the hierarchical data structure. To maintain analytical tractability while still accounting for the hierarchical structure in the data, we focused exclusively on country-level clustering in our multilevel framework. We fitted a multilevel latent class model estimated with a two-step estimator using the multiLCA[42] function in R. The model was specified with individual births as the lower-level units and countries as the higher-level units."

6. **Comment:** The discussion section would benefit from a stronger alignment with recent literature. While some references are cited, integrating findings from other studies to support or contrast your results would enhance their contextual relevance.
   Response: Thank you for this comment. We have strengthened the discussion by integrating recent literature more explicitly to contextualize our results and highlight alignment or contrast with previous findings. Specifically, we have added relevant citations and comparisons throughout the discussion to situate our findings within the broader context of existing research. See below some examples:

   "Multinomial regression analysis revealed associations between maternal and child health service utilization and SES. Mothers with higher education had lower odds of belonging to the lowest utilization subgroup (OR: 0.00, 95% CI: 0.00-0.00). Mothers in higher wealth quintiles were less likely to be in the lowest utilization group (0.23, 0.12-0.43). Urban residence was associated with decreased odds of belonging in the lowest utilization subgroup (0.37, 0.27-0.51), and mothers' employment showed a protective effect against belonging to the lowest utilization subgroup (0.60,

0.47-0.78). Furthermore, SII and RCI inequality measures indicated disparities in service utilization along dimensions of education (SII: 0.679, RCI: 0.197), wealth (SII: 0.525, RCI: 0.155), place of residence (SII: 0.599, RCI: 0.120), and employment (SII: 0.128, RCI: 0.028). Our findings corroborate a well-established body of literature indicating a SES gradient in maternal and child healthcare utilization and access[5,6,16,17]. However, co-utilization patterns of maternal and child health services along the continuum of care remains underexplored, and our study adds novel insight into this area.

At the country level we observed notable heterogeneity in the prevalence of the low-level utilization subgroups across countries. Countries classified as HL1, representing nearly half of our sample, demonstrate high rates of "high utilization" (79.08%), suggesting these nations have likely developed more accessible maternal and child health services. Meanwhile, HL3 countries exhibit a more fragmented utilization landscape, with substantially higher proportions of "medium" and "low" utilization individuals. To contextualize our findings, we compare them with a recent multi-country study on wealth-based inequalities in the continuum of maternal health service utilization across 16 SSA countries[27]. This study found that completion of the maternal continuum of care was below 50% in nine of the 16 countries they analyzed, - four of which (Angola, Ethiopia, Guinea, and Nigeria) have been classified as HL3 countries in our study. This further supports our findings. Additionally, our latent classes, reveal co-utilization patterns of health services, within these countries, which provides insights into gaps and opportunities for intervention delivery.

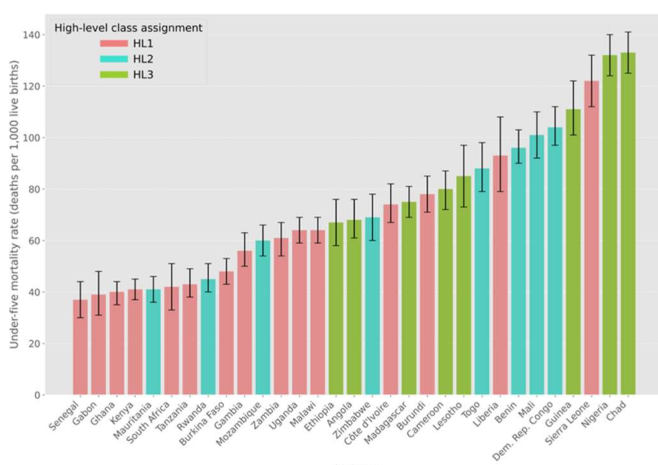
By demonstrating clear subgroup differences, the policy implications for our findings is a call for targeted interventions to address specific subgroup needs. As such, identification of mothers, belonging to a given subgroup is crucial, to enable suitable recommendations and follow ups. This can be done using demographics and SES characteristics. For the low utilization subgroup, comprehensive individual and community-level strategies, aimed at improving SES, and access of maternal and child health services are necessary. At the individual level, poverty alleviation programs such as conditional cash transfers have been shown to increase antenatal care attendance and institutional deliveries by reducing financial barriers[28]. At the community level, community health worker programs have been shown to effectively increase antenatal care, and reduce infant mortality rates, by providing home-based care and referrals[29,30]. Similarly, women empowerment programs have been shown to enhance education, financial independence, and decision-making power to improve maternal and child healthcare utilization[31,32]. For the medium utilization subgroup, interventions should focus on addressing the specific barriers that prevent women from accessing institutional delivery services, while leveraging the relatively higher uptake of antenatal and postnatal care services. For example, community-based antenatal counseling. A study in Burkina Faso found that behavior change communication during antenatal care significantly improved facility-based births[33]. Additionally, a study in Nigeria showed that engaging traditional birth attendants has been successful in increasing institutional deliveries[34]. These evidence-based interventions can effectively increase institutional delivery rates in this subgroup, improving outcomes. It is worth noting that the patterns derived from our findings are still preliminary. Confirmatory studies are imperative to assess subgroup differences in relevant outcomes, that is maternal and child morbidity and mortality."

7. **Comment:** The manuscript identifies several issues but lacks practical, evidence-based solutions to address these disparities. Including actionable recommendations, such as community health worker programs, poverty alleviation strategies, or culturally tailored interventions, would add significant value.
   Response: Thank you for this comment. We have modified our discussion to include practical evidence-based solutions to address the issues from our findings. The current Discussion section includes the following statements now:

"By demonstrating clear subgroup differences, the policy implications for our findings is a call for targeted interventions to address specific subgroup needs. As such, identification of mothers, belonging to a given subgroup is crucial, to enable suitable recommendations and follow ups. This can be done using demographics and SES characteristics. For the low utilization subgroup, comprehensive individual and community-level strategies, aimed at improving SES, and access of maternal and child health services are necessary. At the individual level, poverty alleviation programs such as conditional cash transfers have been shown to increase antenatal care attendance and institutional deliveries by reducing financial barriers[28]. At the community level, community health worker programs have been shown to effectively increase antenatal care, and reduce infant mortality rates, by providing home-based care and referrals[29,30]. Similarly, women empowerment programs have been shown to enhance education, financial independence, and decision-making power to improve maternal and child healthcare utilization[31,32]. For the medium utilization subgroup, interventions should focus on addressing the specific barriers that prevent women from accessing institutional delivery services, while leveraging the relatively higher uptake of antenatal and postnatal care services. For example, community-based antenatal counseling. A study in Burkina Faso found that behavior change communication during antenatal care significantly improved facility-based births[33]. Additionally, a study in Nigeria showed that engaging traditional birth attendants has been successful in increasing institutional deliveries[34]. These evidence-based interventions can effectively increase institutional delivery rates in this subgroup, improving outcomes. It is worth noting that the patterns derived from our findings are still preliminary. Confirmatory studies are imperative to assess subgroup differences in relevant outcomes, that is maternal and child morbidity and mortality."

8. **Comment:** The implications for policy are only briefly mentioned. Expanding on how these findings can inform policy decisions, resource allocation, and targeted intervention strategies is critical.
   Response: Thank you for this comment. We have expanded on the policy implications for our main findings and included these in our discussion section. (Please see Comment 7)

9. **Comment:** While radar plots effectively summarize subgroup characteristics, additional visuals like heatmaps by country could help readers better understand cross-country disparities.
   Response: Thank you for this comment. We have included an additional bar plot to show countries' latent class assignment and their corresponding under five mortality rates. Please see below the added plot:



"Figure 3: Assignment of countries to high-level latent classes, based on maximum a posteriori classification rule. The height of the bars indicate each countries' corresponding under-five mortality rates obtained from the Demographic and Health Surveys. The error bars represent 95% confidence intervals of the under-five mortality rates.

10. **Comment:** A sensitivity analysis examining the impact of imputation on clustering results would improve robustness.
Response: We have added a sensitivity analysis using only complete cases and compared the results to our main model. We do not observe significant difference in the created latent classes. These sensitivity analysis results are presented in Appendix 3, as shown below:

"Appendix 3: Sensitivity Analysis

In the main analysis, mode imputation was applied to address missing values across 14 variables, with missingness percentages ranging from 0% to 7.84%. To assess the robustness of our latent class results to the use of imputed data, we conducted a sensitivity analysis comparing latent classes derived from the primary analysis (including imputed data) against those obtained from a complete-case analysis (excluding imputed cases). Figures 1a and 1b below illustrate the latent class outcomes derived from the complete-case analysis and the analysis including imputed data, respectively. Visual comparison of these radar plots shows consistency of latent class assignments, with minimal observable differences across the two methods.

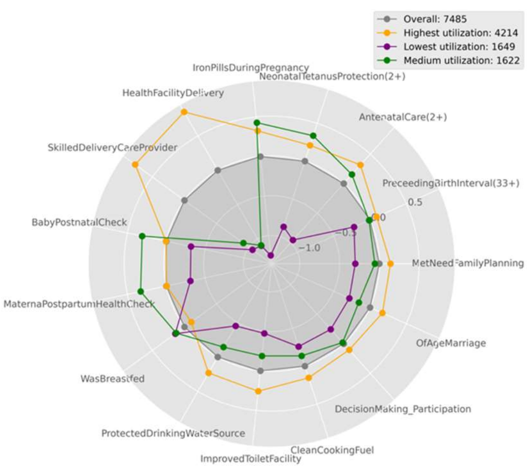


Figure 1a: Latent classes with complete case analysis

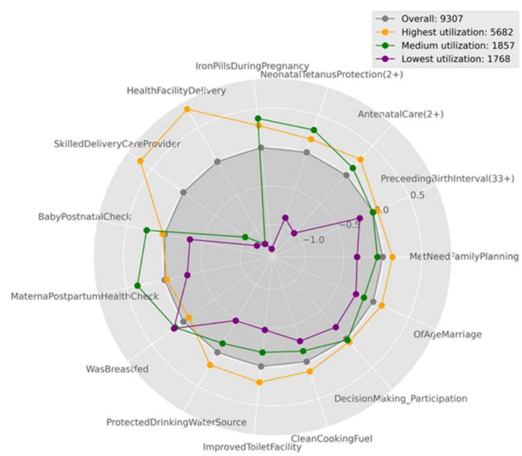


Figure 1b: Latent classes with imputation"

11. **Comment:** The limitations section could be expanded to address potential biases from self-reported data and the exclusion of key RMNCH indicators like immunization and dietary diversity. Discussing these would provide a balanced view of the study's scope.
Response: Thank you. These limitations have been included. See below:

"First, the data are self-reported, which may introduce response biases, and the cross-sectional nature of the surveys limits our ability to establish causality. Furthermore, the indicators used did not include certain key variables along the continuum of maternal and child health services—such as immunizations, minimum dietary diversity, and health-seeking practices—because these data are unavailable at the individual level for births that resulted in U5M."

12. **Comment:** The manuscript uses some technical terms like "trustworthiness" and "stability scores" that may be unfamiliar to readers. Briefly defining these terms or providing a supplementary glossary would enhance accessibility.
Response: Thank you for this suggestion. Given the changes we have made to our clustering analysis from using K-means to using multilevel latent class analysis, we are no longer using these measures for selecting the optimal number of classes. We are however using goodness of fit measures that is the

bayesian information criterion, akaike information criterion, and the integrated complete likelihood bayesian information criterion. We have included these definitions in methods:

"To determine the optimal number of high-level and low-level latent classes, we systematically explored various model configurations by estimating latent class models for one to nine classes to ascertain the optimal structure for latent classes. We first examined low-level latent classes using the 16 indicators of utilization of maternal and child health services for each individual. The selection criteria were based on evaluating the AIC[43], BIC[44], and ICL-BIC[45] for each model. AIC is defined as:

$$\boldsymbol{AIC = -2\ln L + 2k} \tag{3}$$

where: L is the maximized likelihood of the model given the data; k is the number of parameters in the model; and n is the number of data points. BIC on the other hand is defined as:

$$\boldsymbol{BIC = -2\ln L + k\ln n} \tag{4}$$

The ICL-BIC is an extension of BIC for latent class and mixture models, by adding an entropy term. It is defined as:

$$\boldsymbol{ICL - BIC = BIC + Entropy} \tag{5}$$

Entropy measures how well the model separates data into distinct latent classes. Therefore, ICL-BIC accounts for both model fit and the quality of the latent class solution, preferring models that produce clear, well-separated groups. This approach enabled us to assess the trade-off between model complexity and fit, guiding our selection of the most parsimonious and theoretically sound model. Upon establishing the optimal low-level latent class model, we extended our analysis to include high-level latent classes, focusing on the nested nature of the data with countries as the high-level indicators. We similarly evaluated models ranging from one to nine high-level latent classes based on AIC, BIC, and ICL-BIC, aiming to identify the configuration that best balanced detailed representation of the data with model simplicity."

This manuscript is an excellent and timely contribution to the field but addressing these concerns would strengthen its impact and utility. I look forward to seeing the revised version and the potential improvements it will bring.

Reviewer #2 (Remarks to the Author):

Comment: The authors use machine-learning clustering on Demographic and Health Surveys from sub-Saharan Africa to identify distinct patterns in the utilization of services related to reproductive, maternal, newborn, and child health and other factors. They identify three groups and estimate associations between socioeconomic characteristics and the odds of falling in either of the groups.

13. **Comment:** The authors should highlight more what their study adds to existing evidence. While the approach is novel, the authors themselves cite existing studies that show socioeconomic inequalities in service uptake (5,10,11, line 94, page 3). What seems to be the new evidence is the identification of joint prevalence of "interventions", e.g. that there is a considerable population share that utilizes antenatal and postnatal care services but does not deliver in health facilities.

Response: Thank you for this important comment. It is true, the main contribution of our study is in identifying co-utilization patterns in maternal and child health services and practices. For example, we identify a medium utilization subgroup**,** where mothers attend antenatal care and postnatal/postpartum care yet have non-institutional deliveries. Conversely, we observe a highest utilization subgroup**,** where mothers mostly deliver in health facilities and attend antenatal care but do not engage in postnatal or postpartum care. This subgroup also has the highest access to key protective environmental factors, such as clean cooking fuel, improved toilet facilities, and protected drinking water sources. Our study is the first to extract such data-driven utilization patterns offering new insights into how maternal and child health interventions cluster together. We further discover country-level insights; most importantly countries in HL3 demonstrate significant proportions of their populations belonging to the lowest utilization and medium utilization groups. This suggests that despite some engagement with specific interventions, a large segment of the population in these countries does not receive the full continuum of care. An additional contribution of our study lies in examining the associations between socio-economic status and specific utilization patterns and assessing inequalities in health service utilization across key dimensions, including wealth, education, place of residence, and employment. By doing so, our findings contribute to the literature on disparities in maternal and child health service access. We have modified our Discussion to articulate these points more clearly. Here are selected paragraphs specifying these contributions:

Using MLCA, we revealed patterns of co-utilization of RMNCH services and protective environmental and cultural practices, in 31 SSA countries. This method signaled three distinct utilization subgroups that is highest, medium, and lowest utilization – highlighting ways in which services are co-utilized. Notably, mothers in the medium utilization subgroup attended antenatal and postnatal care yet had home deliveries, while mothers in the highest utilization subgroup primarily had institutional deliveries but did not attend postnatal and postpartum care. These findings indicate unique gaps in the continuum of care. At the country level, we uncovered three high-level latent classes (HL1, HL2, HL3), revealing country similarities and differences in service utilization. HL3 countries had the highest share of their populations belonging in the lowest and medium utilization subgroups.

Multinomial regression analysis revealed associations between maternal and child health service utilization and SES. Mothers with higher education had lower odds of belonging to the lowest utilization subgroup (OR: 0.00, 95% CI: 0.00-0.00). Mothers in higher wealth quintiles were less likely to be in the lowest utilization group (0.23, 0.12-0.43). Urban residence was associated with decreased odds of belonging in the lowest utilization subgroup (0.37, 0.27-0.51), and mothers' employment showed a protective effect against belonging to the lowest utilization subgroup (0.60, 0.47-0.78). Furthermore, SII and RCI inequality measures indicated disparities in service utilization along dimensions of education (SII: 0.679, RCI: 0.197), wealth (SII: 0.525, RCI: 0.155), place of residence (SII: 0.599, RCI: 0.120), and employment (SII: 0.128, RCI: 0.028). Our findings corroborate a well-established body of literature indicating a SES gradient in maternal and child healthcare utilization and access[5,6,16,17]. However, co-utilization patterns of maternal and child health services along the continuum of care remains underexplored, and our study adds novel insight into this area.

At the country level we observed notable heterogeneity in the prevalence of low-level utilization subgroups across countries. Countries classified as HL1, representing nearly half of our sample, demonstrate high rates of "high utilization" (79.08%), suggesting these nations have likely developed more accessible maternal and child health services. Meanwhile, HL3 countries exhibit a more fragmented utilization landscape, with substantially higher proportions of "medium" and "low" utilization individuals. To contextualize our findings, we compare them with a recent multi-country study on wealth-based inequalities in the continuum of maternal health service utilization across 16 SSA countries27. This study found that completion of the maternal continuum of care was below 50% in nine of the 16 countries they analyzed, - four of which (Angola, Ethiopia, Guinea, and Nigeria) have been classified as HL3 countries in our study. This further supports our findings. Additionally, our latent classes, reveal co-utilization patterns of health services, within these countries, which provides insights into gaps and opportunities for intervention delivery."

14. **Comment:** The authors use the term “RMNCH interventions” to describe standard care services and individual behavior. The term “interventions” is somewhat misleading, especially in the abstract, as it suggests external, additional programs or services. A different term could provide a better description. **Response:** Thank you for this suggestion. We have modified this term to use “services and practices”.

15. **Comment:** Similarly, the authors use the term “coverage” which is strongly linked to the supply side, describing where services are offered and to what extent. However, given the individual level self-reported data, they actually analyze uptake or utilization of services and individual behavior. The second paragraph in the introduction differentiates between utilization and coverage, but the remaining manuscript and description of analysis only refers to coverage. One statement in the introduction (line 107, page 3) and one in the discussion (line 326, page 11) refer to utilization/uptake, but the focus on the term “coverage” is misleading. The authors do not measure the supply side of health care. This should be clarified throughout the manuscript and discussed.

    Response: Thank you for this suggestion. Indeed, we do not measure the supply side but rather the uptake/utilization side. We have modified our manuscript to use the term “utilization”.

16. **Comment:** The authors categorize the factors included in the cluster analysis into “intervention indicators”, “protective environmental factors” and “traditional factors” (described in Methods, Coverage indicators). Both title and abstract only refer to the “intervention indicators”. Moreover, the category “traditional factors” is not fittingly named: tobacco use does not seem to be “traditional”, neither is “decision-making participation”. The aspect “marriage at a mature age” is coded based on first birth instead of marriage (Table 1 in Appendix 2) and should be labeled as such.

    Response: Thank you for this comment. We have renamed the “traditional factors” category to “cultural factors” as also suggested by reviewer 3 (See comment 30). We have additionally corrected the description of “marriage at a mature age” to “Whether the mother was 18 years or older at first marriage” as it was coded based on first marriage or consensual union (see Table 1 in Appendix 2).

17. **Comment:** Methods, Unsupervised learning: The authors state that they use UMAP for dimensionality reduction. However, they also state this approach worked “significantly better” than conventional multiple correspondence analysis, indicating that this was also implemented but is not presented. This is inconsistent and it is not clear what “significantly better” means and what criteria the authors used to decide on the approach.

    Response: Thank you for this observation. It’s true we implemented multiple correspondence analysis, but we only showed its results in the supplementary material. Following suggestions raised by reviewer 1 however, we have modified our analysis to use multilevel latent class analysis instead of UMAP and K-means. This approach allows us to account for the hierarchical nature in our data. (see comment 5).

18. **Comment:** Methods, Statistical analysis: The authors describe that they first conducted bivariate analysis for each covariate separately and included those with p-value <0.25 in the multivariable model. However, the results of the bivariate analysis are not presented. Instead, the results section mentions a “full model adjusting for all other covariates” and Table 3 presents “unadjusted and adjusted results”. It is not clear what “adjusted” and “unadjusted” here means. This inconsistency should be removed or clarified.

Response: Thank you for this comment. We have modified our methods and results section which talk about the multinomial regression analysis. Since we were interested in including all the demographic and socio-economic status predictors of subgroup membership, there is no need for the bivariate analysis. The modified methods and results section is presented below:

In the methods section;

"Multinomial analysis

Following MLCA, a multivariable multinomial regression model was fitted, with low-level class membership as the dependent variable, and demographic and SES indicators as the independent variables. The multinomial logistic regression model for the three latent classes was formally defined as:

$$\boldsymbol{log\left(\frac{P(Y_i = c)}{P(Y_i = 3)}\right) = \beta_{0c} + \beta_{1c}X_{i1} + \cdots + \beta_{pc}X_{ip}\,, \qquad c = 1, 2} \qquad (6)$$

where Y_i is the low-level latent class for individual, i, X_{ip} are predictor variables, and β's represent regression coefficients. To properly account for the complex survey design of the DHS datasets, we implemented the analysis using the svydesign framework, incorporating sampling weights, cluster identifiers (primary sampling units), and stratification variables as specified in the DHS sampling methodology. This approach ensures that our estimates correctly reflect the hierarchical sampling structure and provides appropriate standard errors accounting for the design effects. The highest utilization subgroup was used as the reference subgroup for class membership. Odds ratios (OR) with 95% confidence interval (CI) of belonging to a given class were reported, and variables with p-value < 0.05 in the multivariable analysis were considered significant predictors of utilization subgroups. All statistical tests were two-tailed to account for both positive and negative associations.

In the results section;

Associations between utilization patterns and SES indicators

When examining predictors of the lower utilization subgroups versus the highest utilization subgroup (Table 4), younger maternal age and high parity were associated with a significant increase in odds of membership to the lowest and medium utilization low-level subgroups. Being separated/divorced and having a partner were associated with higher odds of membership to the lowest utilization low-level subgroup. In terms of SES, mothers having attained some level of education, being in the middle, high, or highest wealth quintile, and living in an urban area was associated with decreased odds of being in the lowest and medium utilization subgroups. Additionally, mothers being employed was associated with decreased odds of being in the lowest utilization subgroup.

Table 4: Multinomial subgroup analysis

#	*Characteristics*	**Medium utilization (n = 1857) vs Highest utilization (n = 5682) subgroup OR (95% CI)**	**Lowest utilization (n = 1768) vs Highest utilization (n = 5682) subgroup OR (95% CI)**
1	*Child gender*		
	Female	*Ref*	*Ref*
	Male	*1·06 (0·85-1·33)*	*1·12 (0·88-1·41)*
2	*Mother's current age*	*0·96 (0·94-0·98)****	*0·92 (0·90-0·95)****
3	*Mother's Marital status*		
	Never in union	*Ref*	*Ref*
	Married / living with partner	*1·36 (0·74-2·49)*	*1·64 (1·00-2·67)**
	Widowed/ divorced / separated	*1·06 (0·54-2·07)*	*2·02 (1·07-3·83)**

4	***Maternal parity***	***1·18 (1·10-1·27)****	***1·30 (1·21-1·41)****
5	***Mother's employment status***		
	No	***Ref***	***Ref***
	Yes	***1·08 (0·84-1·39)***	***0·60 (0·47-0·78)****
6	***Mother's education level***		
	No Education	***Ref***	***Ref***
	Primary	***0·69 (0·54-0·87)***	***0·31 (0·24-0·41)****
	Secondary	***0·42 (0·30-0·59)****	***0·09 (0·06-0·14)****
	Higher	***0·19 (0·06-0·61)***	***0·00 (0·00-0·00)****
7	***Wealth quintile***		
	Lowest	***Ref***	***Ref***
	Low	***0·79 (0·58-1·07)***	***0·79 (0·56-1·10)***
	Middle	***0·74 (0·56-1·10)****	***0·51 (0·37-0·70)****
	High	***0·49 (0·34-0·69)****	***0·48 (0·32-0·72)****
	Highest	***0·52 (0·30-0·88)****	***0·23 (0·12-0·43)****
8	***Place of residence***		
	Rural	***Ref***	***Ref***
	Urban	***0·68 (0·51-0·92)****	***0·37 (0·27-0·51)****

**** p-value < 0.05, **p-value < 0.01, ***p-value < 0.001; Ref – reference group for the given variable; OR – Odds Ratio; and 95% CI – 95% confidence interval."***

19. **Comment:** In the discussion, the authors speak of "SES-driven disparities in coverage" and a "pro-rich coverage gap". These phrases seem imprecise and potentially misleading. The authors identify that SES is associated with group membership, i.e. SES is associated with uptake of services and certain behaviors. The phrase "SES-driven", however, implies a causal relation which is not established in this study. "Pro-rich coverage gap" suggests that services are catering to the wealthier (driven by supply side), whereas it seems more likely that the disparity arises from the wealthier individuals being more likely to utilize the services.

    Response: Thank you for raising this concern. We have modified the language used in our manuscript, to show that we simply tested for associations and not causality. We have additionally included analysis to assess inequality measures along dimensions of wealth, education, place of residence, using slope index of inequality and the relative concentration index of inequality. (see comment 4). The new wording in the discussion section is shown below:

    "Multinomial regression analysis revealed associations between maternal and child health service utilization and SES. Mothers with higher education had lower odds of belonging to the lowest utilization subgroup (OR: 0.00, 95% CI: 0.00-0.00). Mothers in higher wealth quintiles were less likely to be in the lowest utilization group (0.23, 0.12-0.43). Urban residence was associated with decreased odds of belonging in the lowest utilization subgroup (0.37, 0.27-0.51), and mothers' employment showed a protective effect against belonging to the lowest utilization subgroup (0.60, 0.47-0.78). Furthermore, SII and RCI inequality measures indicated disparities in service utilization along dimensions of education (SII: 0.679, RCI: 0.197), wealth (SII: 0.525, RCI: 0.155), place of residence (SII: 0.599, RCI: 0.120), and employment (SII: 0.128, RCI: 0.028)."

20. **Comment:** In the discussion, the authors suggest that the machine-learning approach of the analysis has potential to "facilitate improvement in intervention design and delivery". This is not obvious, and the authors should clarify how improvement is expected to be facilitated.

Response: Thank you for this comment. By performing the latent class analysis, we demonstrate co-utilization patterns rather than isolated service use. For example, the medium utilization subgroup showed an unexpected pattern of low institutional births coupled with high postnatal care and antenatal care. This pattern would not be apparent in traditional single-variable or composite indices analyses, and yet it presents unique opportunities for intervention design - such as using antenatal care visits as strategic entry points to educate mothers about the importance of institutional delivery. We have modified our discussion section to articulate this better.

"By demonstrating clear subgroup differences, the policy implications for our findings is a call for targeted interventions to address specific subgroup needs. As such, identification of mothers, belonging to a given subgroup is crucial, to enable suitable recommendations and follow ups. This can be done using demographics and SES characteristics. For the low utilization subgroup, comprehensive individual and community-level strategies, aimed at improving SES, and access of maternal and child health services are necessary. At the individual level, poverty alleviation programs such as conditional cash transfers have been shown to increase antenatal care attendance and institutional deliveries by reducing financial barriers[28]. At the community level, community health worker programs have been shown to effectively increase antenatal care, and reduce infant mortality rates, by providing home-based care and referrals[29,30]. Similarly, women empowerment programs have been shown to enhance education, financial independence, and decision-making power to improve maternal and child healthcare utilization[31,32]. For the medium utilization subgroup, interventions should focus on addressing the specific barriers that prevent women from accessing institutional delivery services, while leveraging the relatively higher uptake of antenatal and postnatal care services. For example, community-based antenatal counseling. A study in Burkina Faso found that behavior change communication during antenatal care significantly improved facility-based births[33]. Additionally, a study in Nigeria showed that engaging traditional birth attendants has been successful in increasing institutional deliveries[34]. These evidence-based interventions can effectively increase institutional delivery rates in this subgroup, improving outcomes. It is worth noting that the patterns derived from our findings are still preliminary. Confirmatory studies are imperative to assess subgroup differences in relevant outcomes, that is maternal and child morbidity and mortality."

21. **Comment:** In the discussion, the authors recommend individual and community-level strategies for the low coverage subgroup. It is not obvious how these recommendations are drawn from the results. Also, references suggesting the success of such strategies should be added.

    Response: Thank you for this comment. We clarify that our recommendations for individual-and community-level strategies for the low-utilization subgroup are directly informed by our latent class results, which reveal that this subgroup has minimal engagement with maternal and child health services and have least utilization of protective environmental and cultural practices in their households. We recommend poverty alleviation programs which are mainly delivered at the individual level to enhance mothers' ability to afford protective environmental practices in their homes, as well as increase their attendance of health-facility services, by removing financial barriers such as transportation costs. We also recommend community-level interventions such as community health worker outreach and maternal health education programs to increase awareness of maternal and child health services and their availability at community health facilities. We have modified this paragraph in the Discussion section and provided examples of similar strategies that have worked:

    For the low utilization subgroup, comprehensive individual and community-level strategies, aimed at improving SES, and access of maternal and child health services are necessary. At the individual level, poverty alleviation programs such as conditional cash transfers have been shown to increase antenatal care attendance and institutional deliveries by reducing financial barriers[28]. At the

community level, community health worker programs have been shown to effectively increase antenatal care, and reduce infant mortality rates, by providing home-based care and referrals[29,30]. Similarly, women empowerment programs have been shown to enhance education, financial independence, and decision-making power to improve maternal and child healthcare utilization[31,32].

Minor comments

22. **Comment:** The research questions presented in the last paragraph of the introduction could be more precise. The research question of the clustering analysis is not clearly phrased. The hypothesis about the association between socioeconomic status (not disparities) and utilization should be presented as association, not indicating a causal relationship using the word "drive".

    Response: Thank you. We have modified the research questions to be more clearly written:

    "Our study takes a novel approach by conducting multilevel latent class analysis (MLCA) using individual-level data of births that ended in U5M based on utilization/uptake of 16 key RMNCH services and protective environmental and cultural factors (collectively referred to as maternal and child health services hereafter). This study makes important contributions to the literature. First, it defines data-driven patterns of co-utilization of maternal and child health services. Second, it examines the association between these utilization patterns and SES. Third, it investigates the inequalities in utilization of these life-saving services along dimensions of wealth, education, place of residence, and employment. We employed MLCA to account for the hierarchical structure of our data. At the individual level, we categorized births into distinct service utilization subgroups, while at the country level, we identified clusters of nations with similar service utilization profiles. This dual-level approach captures both within-country heterogeneity and between-country similarities in service use patterns. By exploring SES predictors of subgroup membership, we tested the hypothesis that SES is associated with health service utilization patterns. Additionally, we tested the hypothesis that there exist disparities in service utilization along SES dimensions by obtaining relative concentration indices (RCI) and slope indices of inequality (SII) related to membership in the optimal utilization subgroup.

23. **Comment:** The link between subgroup proportions and U5M rates is not assessed thoroughly in this study, so should not be presented as a research question.
    Response: Thank you for this suggestion. We have removed this as a main aim of the study. (Please see comment 22 above)

24. **Comment:** Methods, Data source: The abbreviation "LMICs" is used without prior definition. As this is the only instance where this abbreviation is used, it should be spelled out.
    Response: Thank you, this has been modified.

25. **Comment:** Methods, Statistical analysis: the dependent variable is defined as "the presence of the cluster", whereas it seems that it should be "whether the individual is inside/part of the cluster" or "group membership".
    Response: Thank you. This has been modified as suggested. (Please see comment 18 above)

26. **Comment:** The authors should also discuss whether their findings on the associations between SES and utilization of services are in line with the existing evidence.
    Response: Thank you. This discussion has been included. Please see below:

    Multinomial regression analysis revealed associations between maternal and child health service utilization and SES. Mothers with higher education had lower odds of belonging to the lowest

utilization subgroup (OR: 0.00, 95% CI: 0.00-0.00). Mothers in higher wealth quintiles were less likely to be in the lowest utilization group (0.23, 0.12-0.43). Urban residence was associated with decreased odds of belonging in the lowest utilization subgroup (0.37, 0.27-0.51), and mothers' employment showed a protective effect against belonging to the lowest utilization subgroup (0.60, 0.47-0.78). Furthermore, SII and RCI inequality measures indicated disparities in service utilization along dimensions of education (SII: 0.679, RCI: 0.197), wealth (SII: 0.525, RCI: 0.155), place of residence (SII: 0.599, RCI: 0.120), and employment (SII: 0.128, RCI: 0.028). Our findings corroborate a well-established body of literature indicating a SES gradient in maternal and child healthcare utilization and access[5,6,16,17]. However, co-utilization patterns of maternal and child health services along the continuum of care remains underexplored, and our study adds novel insight into this area."

27. **Comment:** Table 3 should contain the sample size in each specification (N).
    Response: Thank you. We have included this information. (Please see Comment 18 above)

28. **Comment:** Figure 2 should specify the source of U5M rates. The sample used in this study only includes births that ended in U5M and is insufficient to determine U5M rates, so the source must be other data.
    Response: Thank you. We have modified this. (Please see Comment 9)

Reviewer #3 (Remarks to the Author):

Comment: I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.
Response: Thank you for your time.

Reviewer #4 (Remarks to the Author):

Comments for authors, Nature Communications manuscript NCOMMS-24-67107 "Patterns of maternal, newborn, and child health interventions coverage, and associated socioeconomic disparities in sub-Saharan Africa: A machine learning cluster analysis."

General comments:

29. **Comment:** Many thanks to the authors for an interesting analysis relating to an important global health challenge. The summary across countries is important and noteworthy. While interesting, I feel the description of the methodology needs to be more closely linked to the application to better allow reproducibility, to better support the conclusions, and to increase the significance to the field.
    Response: Thank you for this comment. We have improved the methodology, to better suit the structure of our dataset, and linked it to the application. We have also presented estimation equations, and functions used to obtain these results, to allow for reproducibility. (Please see the methods section)

30. **Comment:** I suggest replacing the phrase "traditional factor" by "cultural factor" throughout since "traditional factor" makes for some awkward phrases, e.g., line 76, "traditional practices (such as marriage at a mature age)". I believe the "traditional practice" would be marriage at a young age, so this is difficult to parse. As an option, one could say "traditional practices (such as age of marriage)", but I prefer the phrase "cultural factors").
    Response: Thank you for this suggestion. We have modified this naming to "cultural factors".

31. **Comment:** I suggest clearly stating in the overall description and in the Methods description that the authors are using machine learning/clustering for the national level variables and multinomial regression for the individual level demographic and SES variables. On my initial read, I found it difficult to clearly define which variables were part of clustering and which were part of the multinomial regression. Once the reader takes a step back, you can notice the national/individual distinction, but it would be really helpful for readers to make this clear from the beginning.

    Response: Thank you for this comment. We apologize for the confusion. In this revised manuscript, we performed a multilevel latent class analysis. This analysis involves firstly, estimating low-level latent classes using individual data, on utilization of maternal and child health services. Secondly, high-level latent class analysis is performed by clustering countries with similar service utilization distributions. This allowed us to obtain latent class assignments for both the individual units (births) and the countries. We later performed multinomial regression analysis using individual data with the dependent variable being the low-level latent class membership and the independent variables being demographic and socio-economic status indicators. (Please see Methods section)

32. **Comment:** I also think it would improve readability to swap the order of Tables 1 and 2 and present these in the same order in which the analysis flows: Summary of national level indicators of RMNCH/environmental/cultural interventions/practices, then grouping of these indicators into summary clusters, then assessment of individual U5M by cluster and individual mother/child measures. Following a consistent order will help readers keep track of the logic of the analysis.

    Response: Thank you for this comment. Table 1 indicates the descriptive statistics of the sample demographic and socio-economic characteristics and since it is conventional practice in papers for this table to come first as Table 1, we feel it is appropriate to leave it as Table 1.

33. **Comment:** The description of the methods is difficult to follow and would be difficult to replicate. In the main manuscript, it is difficult to grasp the distinction between the machine learning component and the clustering-based dimension reduction. That is, it is unclear (at least to me) where UMAP ends and the two clustering methods begin in the main body of the manuscript. The supplemental information provides some additional insight, but with only 16 RMNCH (national-level) variables, the clustering algorithms seem to take care of the dimension reduction (from 16 to 3), indeed, UMAP results do not seem to be mentioned in the Results section of the main manuscript and the description of the UMAP methodology in the supplemental information is very generic and not particularly specific to the data set for this study until Appendix 4 which is very brief. What did UMAP reveal, what were the 250, 500, 1000, 1500 neighbors mentioned? Table 2 in the Supplemental Information is presented without description. I feel a more careful discussion (perhaps in the supplement) would aid in others being able to understand and replicate the work.

    Response: Thank you for this detailed comment. We have made various improvements in the methods section. In response to Reviewer One's comments (see comments 3 and 5), we have revised our analysis to use multilevel latent class analysis instead of the UMAP and K-means methods, as the latter do not account for the hierarchical structure of our dataset in the clustering process. We have clarified that the latent class analysis was performed at both the individual and country levels, with both individuals and countries assigned to groups based on service utilization patterns (see Methods and Results sections).

34. **Comment:** The K-means and hierarchical cluster weights shown in Figures 1a and 1b are remarkably and surprisingly similar. Indeed, the only visible differences between the two figures are legend elements "Lowest coverage: 1708" and "Lowest coverage: 1752" and "Medium coverage: 1655" and "Medium coverage: 1611". Typically, I would expect to see some minor variation in details when using different clustering mechanisms. Could the authors confirm that the results are this close (perhaps

highlighting those observations that shifted between lowest and medium coverage in the supplemental information)? Are these countries geographically close together?

Response: Thank you for this comment. We confirmed that these results were indeed this close, and this was confirmed by the adjusted rand index of 0.99, which indicates near perfect agreement in the clustering solutions of the two methods. However, we have revised our analysis to account for the hierarchical structure in the data and observed similar latent class solutions as previously reported in the original submission. The adjusted rand index between the current latent class solution and the previous solution is 0.8658 which indicates that both approaches have largely identified similar structures within our data.

Specific comments (main text):

35. **Comment:** Line 54. “decreased of being” to “decreased odds of being”.
    Response: Thank you for this comment. We have made the edit.

36. **Comment:** Line 64. Sustainable Development Goal be capitalized.
    Response: Thank you for this comment. We have made the edit.

37. **Comment:** Lines 73-74. Note that more detail will be provided regarding the specific RMNCH interventions considered later in the manuscript.
    Response: Thank you, indeed we provide further details about the descriptions of each of the maternal and child health services indicators in the Supplementary material. (Table 1, Appendix 2)

38. **Comment:** Line 86. “coverage in” to “coverage of”.
    Response: Thank you for this comment. We have adopted the use of “utilization” instead of coverage throughout the manuscript, as requested by reviewer 2. (see comment 15).

39. **Comment:** Lines 94-95. “how specific interventions tend to cluster together of co-occur”. This point is not clear in the abstract but is central to the analysis. I suggest making the motivation for clustering clearer in the abstract.
    Response: Thank you for this comment. We have modifed out introduction to better motivate the reason for clustering/latent class analysis. (Please see Comment 1)

40. **Comment:** Lines 102-103. “exceed preconceived associations” to “extend beyond a preconceived set of potential associations”.
    Response: Thank you. We have removed this statement from the manuscript.

41. **Comment:** Line 106. “This enables to test” to “This enables us to test”.
    Response: Thank you for this suggestion. This statement is no longer part of the manuscript.

42. **Comment:** Line 108. “aimed to show” to “aimed to describe”.
    Response: Thank you for this suggestion. This statement is no longer part of the manuscript.

43. **Comment:** Line 109. “explain geographical” to “explain observed geographical”.
    Response: Thank you for this suggestion. This statement is no longer part of the manuscript.

44. **Comment:** Line 136. “water sources; while tobacco” to “water sources. Tobacco” (start a new sentence for the cultural factors.
    Response: Thank you for this suggestion. We have made the edit.

45. **Comment:** Line 145. “wealth tier” Is this defined in the DHS data?

Response: Thank you for this comment. We have used the wealth quintile variable as defined in the DHS data.

46. **Comment:** Line 155. Does "mode imputation" mean you imputed the mode for any missing values? Was a sensitivity analysis done to see if results change with multiple imputation (particularly for marriage at a mature age)?
Response: Thank you for this comment. Yes we performed imputation for all values with missing values that is mother's employment status (0·16%), health facility delivery (0·17%), skilled delivery provider (0·20%), breastfeeding (0·59%), postnatal check (0·73%), iron pills during pregnancy (0·89%), family planning demand satisfaction (1.34%), neonatal tetanus protection (1·62%), antenatal care visits (4+) (1·71%), protected drinking water source (2·25%), clean cooking fuel (2·25%), improved sanitation facility (2·29%), postpartum check (5·78), and marriage at mature age (7·84%). We have conducted a sensitivity analysis to see if the results change with complete case analysis. The results did not change, and we still make the same conclusions about the formed latent classes. (see Comment 10)

47. **Comment:** Line 165. "Significantly better". Please describe in what ways the results are better (and I suggest avoiding the term "significantly").
Response: Thank you for this suggestion. We have removed the use of the term significant.

48. **Comment:** Line 183. How is "incorrect clustering" defined?
Response: Thank you for this question. Given that clustering is often done in an unsupervised way, where we have no ground truths, the quality of clustering solutions is often assessed by metrics such as stability[30] and silhouette scores[31]. To assess stability of clustering results, the clustering is repeated several times on random samples of the dataset and the cluster assignment of individual units is compared across different runs. A higher stability score means individual units have consistently been assigned to the same clusters across multiple runs. Silhouette scores on the other hand measure the quality of formed clusters by measuring how similar an individual is to its own cluster compared to other clusters. The silhouette score ranges from -1 to 1, where a high value indicates that the object is well matched to its own cluster and poorly matched to neighboring clusters.

49. **Comment:** Line 226. "interventions" to "intervention coverage".
Response: Thank you. We have adopted the term "services" instead of interventions as suggested by Reviewer 2. (See comment 14)

50. **Comment:** Line 231. "majority" to "a majority"?
Response: Thank you for this comment. We have made the edit.

51. **Comment:** Line 251. "Being separated/divorced and having a partner was associated…". Does this refer to one classification (separated/divorced with a partner) or two ((1) separate/divorced and single, (2) having a partner)?
Response: Thank you for this question. These are two separate classes, that is 1) Widowed/ divorced / separated and 2) Married / living with partner. We have clarified this statement.

52. **Comment:** Line 293. "One reason for this is probability the ability…". Can this be surmised from the clusters or is this a supposition separate from the data analysis?
Response: Thank you for this question. This cannot directly be surmised from the clusters, but it is a supposition separate from the analysis, as our way of explaining probable reasons behind our results.

53. **Comment:** Line 312. "majority" to "the majority". Also, no comma after "postnatal care".
Response: Thank you for this suggestion, we have made these modifications.

54. **Comment:** Line 377. “are critical” to “provide insight”.
Response: Thank you for this suggestion. We have removed the use of the term significant.

Specific comments (supplemental information):

55. **Comment:** 1. Page 4, Appendix 3. While interesting, since the authors only used one of the listed distance measures, it would be sufficient to list that method (Jaccard similarity) and provide a general reference to the others (which would not apply to the authors’ categorical data). Also, I believe “jaccard” should be capitalized throughout.
Response: Thank you for this comment. Since we are no longer using these metrics, we have removed this section from the supplementary material.

Reviewer #4 (Remarks on code availability):
While I did not run the code, it is well documented within a Docker setting with python code and simulated data to illustrate the approach.

References

1. Verma, A., Angelini, O. & Di Matteo, T. A new set of cluster driven composite development indicators. *EPJ Data Sci.* **9**, 1–21 (2020).

2. Barros, A. J. *et al.* Equity in maternal, newborn, and child health interventions in Countdown to 2015: a retrospective review of survey data from 54 countries. *The Lancet* **379**, 1225–1233 (2012).

3. Barros, A. J. D. & Victora, C. G. Measuring Coverage in MNCH: Determining and Interpreting Inequalities in Coverage of Maternal, Newborn, and Child Health Interventions. *PLOS Med.* **10**, e1001390 (2013).

4. Leventhal, D. G. P. *et al.* Delivery channels and socioeconomic inequalities in coverage of reproductive, maternal, newborn, and child health interventions: analysis of 36 cross-sectional surveys in low-income and middle-income countries. *Lancet Glob. Health* **9**, e1101–e1109 (2021).

5. Boerma, T. *et al.* Countdown to 2030: tracking progress towards universal coverage for reproductive, maternal, newborn, and child health. *The Lancet* **391**, 1538–1548 (2018).

6. Wehrmeister, F. C., Barros, A. J. D., Hosseinpoor, A. R., Boerma, T. & Victora, C. G. Measuring universal health coverage in reproductive, maternal, newborn and child health: An update of the composite coverage index. *PLoS ONE* **15**, e0232350 (2020).

7. Singh, K., Story, W. T. & Moran, A. C. Assessing the Continuum of Care Pathway for Maternal Health in South Asia and Sub-Saharan Africa. *Matern. Child Health J.* **20**, 281–289 (2016).

8. Anastasi, E. *et al.* Losing women along the path to safe motherhood: why is there such a gap between women's use of antenatal care and skilled birth attendance? A mixed methods study in northern Uganda. *BMC Pregnancy Childbirth* **15**, 287 (2015).

9. Wang, W. & Hong, R. Levels and determinants of continuum of care for maternal and newborn health in Cambodia-evidence from a population-based survey. *BMC Pregnancy Childbirth* **15**, 62 (2015).

10. Mari, R. D., Bakk, Z., Oser, J. & Kuha, J. A two-step estimator for multilevel latent class analysis with covariates. *Psychometrika* **88**, 1144–1170 (2023).

11. The DHS Program User Forum: Weighting data » Weighting and pooling multicountry datasets. https://userforum.dhsprogram.com/index.php?t=msg&th=11763&start=0&S=Google.
12. World Population Prospects. https://population.un.org/wpp/.
13. Schlotheuber, A. & Hosseinpoor, A. R. Summary Measures of Health Inequality: A Review of Existing Measures and Their Application. *Int. J. Environ. Res. Public. Health* **19**, 3697 (2022).
14. Hagenaars, J. A. & McCutcheon, A. L. *Applied Latent Class Analysis*. (Cambridge University Press, 2002).
15. Lyrvall, J., Mari, R. D., Bakk, Z., Oser, J. & Kuha, J. multilevLCA: An R Package for Single-Level and Multilevel Latent Class Analysis with Covariates. Preprint at https://doi.org/10.48550/arXiv.2305.07276 (2024).
16. Alam, N., Hajizadeh, M., Dumont, A. & Fournier, P. Inequalities in Maternal Health Care Utilization in Sub-Saharan African Countries: A Multiyear and Multi-Country Analysis. *PLOS ONE* **10**, e0120922 (2015).
17. Langlois, É. V. *et al.* Inequities in postnatal care in low- and middle-income countries: a systematic review and meta-analysis. *Bull. World Health Organ.* **93**, 259-270G (2015).
18. Atake, E.-H. Socio-economic inequality in maternal health care utilization in Sub-Saharan Africa: Evidence from Togo. *Int. J. Health Plann. Manage.* **36**, 288–301 (2021).
19. Asefa, A. *et al.* Wealth-based inequality in the continuum of maternal health service utilisation in 16 sub-Saharan African countries. *Int. J. Equity Health* **22**, 203 (2023).
20. Glassman, A. *et al.* Impact of conditional cash transfers on maternal and newborn health. *J. Health Popul. Nutr.* **31**, 48–66 (2013).
21. Kayentao, K. *et al.* Effect of community health worker home visits on antenatal care and institutional delivery: an analysis of secondary outcomes from a cluster randomised trial in Mali. *BMJ Glob. Health* **8**, (2023).

22. Nadella, P., Subramanian, S. V. & Roman-Urrestarazu, A. The impact of community health workers on antenatal and infant health in India: A cross-sectional study. *SSM - Popul. Health* **15**, 100872 (2021).

23. Prost, A. *et al.* Women's groups practising participatory learning and action to improve maternal and newborn health in low-resource settings: a systematic review and meta-analysis. *Lancet Lond. Engl.* **381**, 1736–1746 (2013).

24. Fottrell, E. *et al.* The Effect of Increased Coverage of Participatory Women's Groups on Neonatal Mortality in Bangladesh: A Cluster Randomized Trial. *JAMA Pediatr.* **167**, 816–825 (2013).

25. Soubeiga, D., Sia, D. & Gauvin, L. Increasing institutional deliveries among antenatal clients: effect of birth preparedness counselling. *Health Policy Plan.* **29**, 1061–1070 (2014).

26. Oyeneyin, L., Osunmakinwa, O. & Olagbuji, Y. Incorporating traditional birth attendants into the mainstream maternal health system in Nigeria - An evaluation of the Ondo State Agbebiye program. *Afr. J. Reprod. Health Rev. Afr. Santé Reprod.* **25**, 82–88 (2021).

27. Cavanaugh, J. E. & Neath, A. A. The Akaike information criterion: Background, derivation, properties, application, interpretation, and refinements. *WIREs Comput. Stat.* **11**, e1460 (2019).

28. Penny, W. D., Mattout, J. & Trujillo-Barreto, N. CHAPTER 35 - Bayesian model selection and averaging. in *Statistical Parametric Mapping* (eds. Friston, K., Ashburner, J., Kiebel, S., Nichols, T. & Penny, W.) 454–467 (Academic Press, London, 2007). doi:10.1016/B978-012372560-8/50035-8.

29. Assessing a Mixture Model for Clustering with the Integrated Completed Likelihood | IEEE Transactions on Pattern Analysis and Machine Intelligence. https://dl.acm.org/doi/10.1109/34.865189.

30. Yu, H. *et al.* Bootstrapping estimates of stability for clusters, observations and model selection. *Comput. Stat.* **34**, 349–372 (2019).

31. Rousseeuw, P. J. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *J. Comput. Appl. Math.* **20**, 53–65 (1987).