Association of Childhood Immunizations with Caregiver Beliefs and Experiences in Bugoye, Uganda

Introduction

In 2016, a national health survey in Uganda revealed that only 55% of children aged 12-23 months received recommended vaccinations (UBOS, 2016). The investigators hypothesize that areas with limited access to immunization services in Uganda show lower childhood vaccination coverage compared to regions with full access to these services (Boyce et al., 2019). For this reason, the target population for our study consists of households with children aged 12-23 months, who reside with a caregiver in Bugoye sub-county, a small rural town in western Uganda. This cross-sectional study aims to explore social determinants, such as caregiver attitudes and beliefs regarding vaccination, associated with childhood vaccination status. Additionally, we will compare vaccination uptake between households that provided vaccine cards with those that self-reported, and examine factors associated with both childhood vaccination status and card availability.

Children should receive the following vaccines and boosters to be considered fully vaccinated: tuberculosis, polio, pentadose 1-3, rotavirus 1-2, and measles. Based on the study aims, our primary hypothesis is that children are more likely to be fully vaccinated and receive specific vaccines, such as measles, if their caregivers hold more positive beliefs and experiences about vaccination than less positive beliefs and experiences. A secondary hypothesis is that sociodemographic factors—such as caregiver education, household characteristics, and clinic center—may act as confounders in the relationship between caregiver perceptions and child vaccination status.

Methods

The Gates Vaccine Access (Phase 3) dataset for the study was collected between January 20, 2021, and April 30, 2021. It is a cross-sectional survey of 1,689 children aged 12–23 months in Bugoye sub-county, capturing vaccination coverage data from vaccination cards or self-reports when cards were missing.

We began with an exploratory analysis, generating summary statistics for key variables, including the primary outcome of full vaccination status, and secondary outcomes such as vaccination card availability and, if unavailable, self-reported vaccination status. Main predictors—continuous or ordinal scores for positive beliefs and vaccination experience—were summarized alongside potential confounders, including caregiver characteristics, child anthropometric measures, vaccination location, and childbirth-related variables. The confounding variables were identified using prior knowledge, and the associations between each possible confounding variable and outcome or exposures were tested with Chi-square tests, Wilcoxon rank-sum tests, Kendall's tau tests, and Kruskal-Wallis tests. All statistical tests were performed at a two-sided level significance level of 0.05.

We used two approaches for measuring main exposures, beliefs and experiences. One of the approaches was to aggregate the questions related to beliefs and experiences into a composite score, and the other was to use multiple disaggregated belief and experience indicator variables. In our aggregated score, the more vaccine-positive response was assigned a score of 1, the less vaccine-positive response was assigned a score of 0, and responses which expressed uncertainty were assigned a score of 0.5. A weighted logistic regression was fitted for a primary outcome, fully vaccinated status, to obtain adjusted odds ratios and 95% confidence intervals. To account for recall bias among individuals without vaccine cards, sensitivity analysis was conducted by performing these analyses exclusively on data from those with vaccine cards (67% of the total sample size). However, because individuals with a vaccine card were more likely to be fully vaccinated, the data used for sensitivity analyses is significantly more imbalanced. After weighted logistic regression analysis, a weighted random forest was then implemented to compare performance against the weighted logistic regression after splitting the dataset into a train set (80%) and test set (20%). The evaluation metrics included receiver operating characteristic (ROC) curve, area under the curve (AUC), accuracy, and Kappa. For secondary outcomes, we fitted logistic regression models to identify factors associated with vaccine uptake of measles and card availability. However, for the model with measles vaccination status as the outcome, a Firth logistic regression model was used due to extreme imbalance between the majority class (vaccinated against measles, 96.8% of responses) and the minority class (not vaccinated against measles, 3.2% of responses). Due to the severity of this imbalance, we were unable to test the association of individual beliefs/experiences with measles vaccination status, as doing so would introduce too many covariates into our model.

Results

We began by summarizing the key characteristics of the dataset (Table 1). Among the 1,689 children included, 81.8% were reported as fully vaccinated, while 13.6% were not, and 4.6% had missing vaccination status. Vaccination cards were available for 67.4% of the children, with self-reported vaccination status used when cards were unavailable—among those, 67.4% were recalled as fully vaccinated. Regarding measles vaccination specifically, 65.8% of children had received the vaccine, 1.7% had not, and 32.6% of abundance missing data for this outcome. Most caregivers were female (86.1%) and typically the child's mother (80.1%). Educational attainment among caregivers varied, with the majority having completed primary (60.6%) or secondary education (27.4%), and a small proportion (2.4%) having university-level education. Caregiver marital status was predominantly married (91.1%). The mean age of the children was 20 months, and the average age of the oldest child in the household was 10 years. Regarding healthcare access and practices, 91.6% of children slept under a bed net the night before the survey, and 20.2% stayed overnight in a hospital or clinic at some point. In terms of vaccination sites, most children received vaccines at government facilities (Health Centre II: 39.4%, Health Centre III: 25.6%, followed by private health centers or hospitals: 26.6%). Finally, the main predictor variables showed generally high scores, with a mean positive beliefs score of 8.09 (SD = 1.25) and a mean positive experiences score of 11.32 (SD = 1.2), indicating overall favorable attitudes and experiences related to vaccination.

Table 1. Descriptive Statistics of Key Variables in the Dataset (N=1689)			
Field Note	**Variables**	**Value**	**No. (%)**
Primary Outcome			
Fully Vaccinated	**vacc_full**	Yes	1381 (0.818)
		No	230 (0.136)
		Missing	78 (0.046)
Secondary Outcome			
Self-reported vaccination status when card=0	**vacc_recall**	Yes	1138 (0.674)
		No	541 (0.320)
		Missing	10 (0.006)
Child's Immunization Card Availability	**Card**	Yes	1138 (0.674)
		No	541 (0.320)
		Missing	10 (0.006)
Got the Measles Vaccine	**Measles**	Yes	1111 (0.658)
		No	28 (0.0166)
		Missing	550 (0.3256)
Potential Confounders			
Caregiver Sex	**caregiver_sex**	Male	231 (0.137)
		Female	1454 (0.861)
		Missing	4 (0.002)
Caregiver Relationship	**Relationship**	Mother	1353 (0.801)
		Father	193 (0.114)
		Grandparent	76 (0.045)
		Aunt or Uncle	40 (0.024)
		Other	18 (0.011)
		Missing	9 (0.005)
Caregiver Education	**Education**	No School	161 (0.095)
		Primary School	1024 (0.606)
		Secondary School	462 (0.274)
		University	40 (0.024)
		Missing	2 (0.001)
Caregiver Marital Status	**Marital**	Unmarried	84 (0.050)
		Married	1539 (0.911)
		Divorced	54 (0.032)
		Widowed	10 (0.006)
		Missing	2 (0.001)
Location the child receives most of his/her vaccination	**vacc_where**	Government Health Centre 3	432 (0.256)
		Government Health Centre 2	665 (0.394)
		Private Health Centre or Hospital	449 (0.266)
		Visit by Immunization Program	86 (0.051)
		Other	53 (0.031)
		Missing	4 (0.002)
Child slept under a bed net last night	**bednet**	Yes	1547 (0.916)
		No	140 (0.083)
		Missing	2 (0.001)
Child ever stay overnight in a hospital or clinic	**Inpatient**	Yes	342 (0.202)
		No	1344 (0.796)
		Missing	3 (0.0018)
Age of oldest child, Mean (SD, Median, Range)	**oldest_child**		10 (7.827, 9, [1,40])
		Missing	44 (0.026)
Child Age (Month), Mean (SD, Median, Range)	**age_months**		20 (5.053, 20, [1, 39])
		Missing	7 (0.004)
Main Predictors			
Beliefs positive scores, Mean (SD, Median, Range)	**beliefs1**		8.088 (1.253, 8, [2.5, 10])
		Missing	43 (0.0255)
Positive Experiences Score, Mean (SD, Median, Range)	**exp1**		11.32 (1.2,11, [7,14])
		Missing	72 (0.043)

We examined associations between sociodemographic and health-related factors and the primary outcome of full vaccination status, as well as the main exposures of the study—caregiver beliefs about vaccines and caregiver experiences with accessing vaccination services (Table 2).

Table 2. Associations Between Potential Confounders and Vaccination Status, Positive Belief, and Experience Scores Among Children in Bugoye County (* $p < 0.05$, ** for $p < 0.01$, and *** for $p < 0.001$)

Variable	Type	Metrics for Vaccination Status	p-values	Metrics for Belief and Experience Scores	p-values for Belief Score	p-values for Experience Score
age_months	Continuous	Wilcoxon Rank Sum Test	0.008 **	Kendall's tau Test	0.228	0.279
oldest_child	Continuous	Wilcoxon Rank Sum Test	0.337	Kendall's tau Test	0.454	<0.001 ***
muac	Continuous	Wilcoxon Rank Sum Test	0.340	Kendall's tau Test	0.014 *	0.001 **
children	Continuous	Wilcoxon Rank Sum Test	0.172	Kendall's tau Test	0.946	<0.001 ***
birthplace	Continuous	Wilcoxon Rank Sum Test	0.567	Kendall's tau Test	0.859	0.098
hbv_res	Binary	Chi-square Test (χ^2)	0.409	Wilcoxon Rank Sum Test	0.141	0.161
dbs	Binary	Chi-square Test (χ^2)	0.820	Wilcoxon Rank Sum Test	0.341	0.924
bednet	Binary	Chi-square Test (χ^2)	<0.001 ***	Wilcoxon Rank Sum Test	<0.001 ***	0.808
inpatient	Binary	Chi-square Test (χ^2)	0.034 *	Wilcoxon Rank Sum Test	0.623	<0.001 ***
caregiver_sex	Binary	Chi-square Test (χ^2)	0.001 **	Wilcoxon Rank Sum Test	0.896	0.200
child_sex	Binary	Chi-square Test (χ^2)	0.292	Wilcoxon Rank Sum Test	0.206	0.133
vacc_where	Nominal	Chi-square Test (χ^2)	0.011 *	Kruskal-Wallis test	<0.001 ***	0.002 **
education	Nominal	Chi-square Test (χ^2)	0.002 **	Kruskal-Wallis test	0.006 **	0.019 *
marital	Nominal	Chi-square Test (χ^2)	0.005 **	Kruskal-Wallis test	0.396	0.002 **
birthplace2	Nominal	Chi-square Test (χ^2)	0.785	Kruskal-Wallis test	<0.001 ***	0.833
birth_status	Nominal	Chi-square Test (χ^2)	0.722	Kruskal-Wallis test	0.370	0.548
relationship	Nominal	Chi-square Test (χ^2)	<0.001 ***	Kruskal-Wallis test	0.278	0.350
muac_color	Nominal	Chi-square Test (χ^2)	0.555	Kruskal-Wallis test	0.043 *	0.121

Full vaccination was significantly associated with female (p = 0.001) or maternal caregivers (p < 0.001), caregiver marriage (p = 0.005), higher education (p = 0.002), older child age (p = 0.008), government or health center vaccination (p = 0.011), bed net use (p < 0.001), and inpatient care history (p = 0.034). Caregivers held more positive beliefs about vaccines if their child had better nutritional status (measured by MUAC, p = 0.014) or a green MUAC color (p = 0.043). Higher education (p = 0.006), institutional births (p < 0.001), and receiving vaccines at organized health facilities (p < 0.001) were also associated with stronger beliefs. Additionally, caregivers whose children slept under a bed net reported stronger vaccine beliefs (p < 0.001). Child sex (p = 0.527), birth status (p = 0.542), and card availability (p = 0.781) were not significantly associated, suggesting that beliefs are shaped more by caregiver characteristics and preventive health behaviors than by child-specific or procedural variables.

More favorable vaccination experiences were significantly associated with having an older oldest child (p < 0.001), a greater number of children (p < 0.001), better child nutritional status (p = 0.001), higher caregiver education (p = 0.019), being married (p = 0.002), receiving vaccines in structured healthcare settings (p = 0.002), and a child's prior inpatient care history (p < 0.001). Variables such as child sex (p =

0.293), birthplace ($p = 0.253$), and birth attendant ($p = 0.123$) were not significant, indicating that experiences are more influenced by parental and household factors than child-specific details.

Some variables like MUAC, caregiver education, and vaccination location—were consistently associated with both positive beliefs and better experiences. However, notable differences emerged. Age of the oldest child and number of children were significant for experience scores (both $p < 0.001$) but not beliefs, implies parental familiarity may ease vaccination logistics without necessarily altering attitudes. While institutional birthplace ($p < 0.001$) and MUAC color ($p = 0.043$) were linked to stronger beliefs but not to experiences, suggests that caregivers' early interactions with health systems and visual cues about child health may shape vaccine attitudes more than practical navigation of services.

Thus, caregivers' education level, vaccination location, marital status, inpatient, and bed net use were significantly associated with caregiver beliefs/experiences and vaccination status. These variables may confound the relationship between caregiver perceptions and child immunization outcomes, highlighting the need for statistical adjustment to assess the independent effect of caregiver attitudes.

The normalized weight assigned to fully vaccinated was 0.583, and that for non-fully vaccinated was 3.502 in the weighted logistic regression. The variance inflation factors (VIF) were all below 2, suggesting no concerning multicollinearity. Also, the p-value 0.598, which is greater than 0.05, from the Hosmer and Lemeshow goodness of fit test confirmed the model fits the data well. The adjusted odds ratio for both the beliefs score (1.055) and experiences score (1.044) are greater than 1 (Figure 1). However, the associations are not found to be statistically significant, as the 95% confidence interval for beliefs [0.937, 1.186] and experiences [0.924, 1.180] includes 1, which is a null value. Children who got vaccinated from government health centre 3 (aOR = 0.097) or government health centre 2 (aOR = 0.131) had significantly lower odds of being fully vaccinated compared to those who did not, after adjusting for the other variables (Figure 1). Caregivers whose highest education level was secondary school had 1.816 times the odds of having their children fully vaccinated compared to those who had different education levels, after adjusting for the other variables. Caregivers who got married had 1.909 times the odds of having their children fully vaccinated. Children who used bed net had 3.016 times the odds of being fully vaccinated, and those who stayed overnight in a hospital had 1.562 times the odds of being fully vaccinated, after adjusting for the other variables (Figure 1).

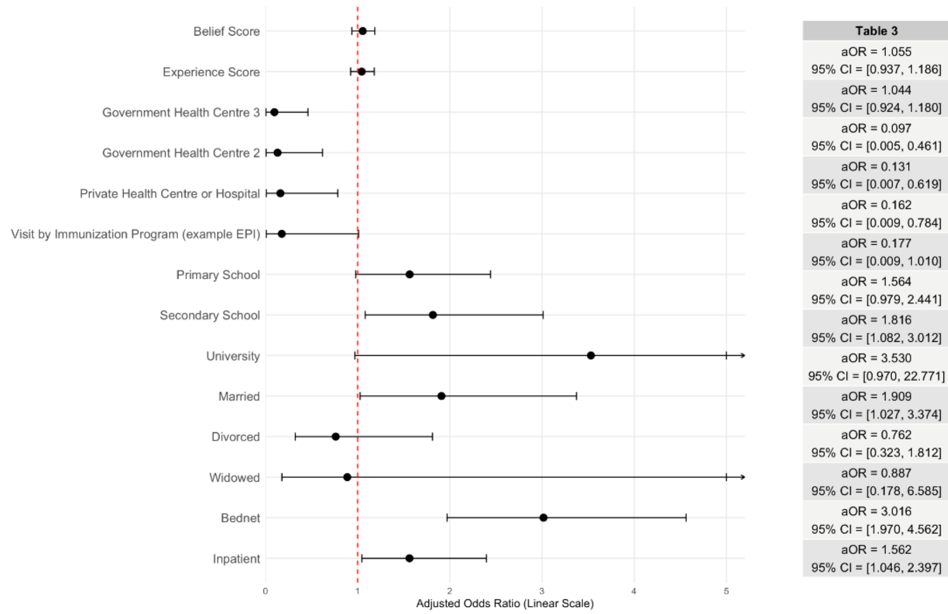

Figure 1. Forest Plot of Adjusted Odds Ratios with 95% Confidence Intervals of Fully Vaccinated Model

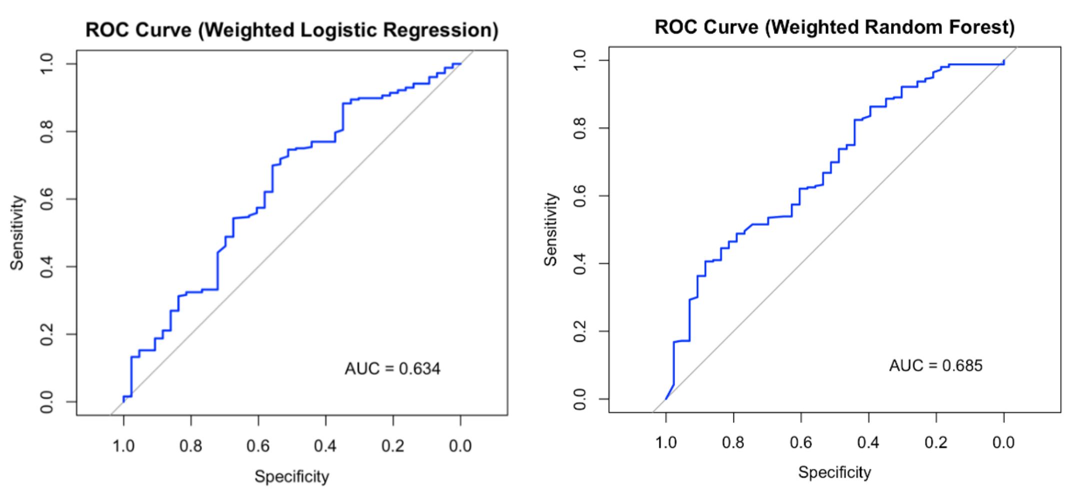

Figure 2. ROC Curve for Weighted Logistic Regression (Left) and

ROC Curve for Random Forest (Right)

The performance of the weighted logistic regression model was compared to that of the weighed random forest with mtry = 5 and number of trees = 50. Figure 2 shows that the random forest slightly outperforms the weighted logistic regression, having a curve closer to the top-left corner and a higher AUC (0.685 compared to 0.634). The accuracy and Kappa for the random forest are 0.863 and 0.158, respectively, whereas the weighted logistic regression yielded 0.632 and 0.120 for those. The most important variables identified by the random forest were belief score and experience score.

In the sensitivity analysis, the proportion of missing data increased to 39.4% compared to 11.3% in the main analysis, after excluding the recall question. The adjusted odds ratio for the beliefs score was 0.838, compared to 1.055 in the main analysis, and the association became statistically significant. That is, our results were sensitive to recall bias.

After examining the association between aggregated belief/experience scores and full vaccination status, we proceeded to examine the association between individual beliefs/experiences and full vaccination status. We fit two models: one in which individual beliefs were disaggregated and the aggregated experiences score was maintained as a control variable, and another in which individual experiences were disaggregated and the aggregated beliefs score was maintained as a control variable. The other control variables were identical to the ones used in the logistic model. For both models, all variance inflation factors (VIFs) were below 2, suggesting that correlation between individual beliefs/experiences (and any other variables) did not reach a level at which our model's assumptions would be violated. We found that three beliefs and three experiences were associated with full vaccination status at a statistically significant level. More specifically, agreeing with the statement that "Children get more vaccinations than are good for them." was associated with a lower probability of a caregiver's child being fully vaccinated (aOR = 0.585, 95% CI: [0.424, 0.808]), whereas agreeing that "Many of the illness which vaccinations prevent are severe." or that "When a parent refuses to vaccinate a child, it harms the entire community through risk of disease." was associated with a higher probability of a caregiver's child being fully vaccinated (aOR = 2.045, 95% CI: [1.229, 3.402]) and 1.804, 95% CI: [1.207, 2.697] respectively). Delaying a child's vaccination for reasons other than allergy or illness was associated with a lower probability of a caregiver's child being fully vaccinated (aOR = 0.384, 95% CI: [0.276, 0.534]), whereas being able to discuss any concerns with a child's healthcare provider or knowing a person who had either polio, pneumonia, measles, or whooping cough was associated with a higher probability of a caregiver's child being fully vaccinated (aOR = 2.584, 95% CI: [1.420, 4.705] and aOR = 1.682, 95% CI: [1.134, 2.315] respectively).

To account for recall bias among those without a vaccine card, we performed sensitivity analyses by fitting the two models exclusively on data from individuals with vaccine cards. In these analyses, we found that two beliefs and one experience were associated with full vaccination status at a statistically significant level. For both models, all variance inflation factors (VIFs) were below 2, suggesting that correlation between individual beliefs/experiences (and any other variables) did not reach a level at which our model's assumptions would be violated. As in our original analysis, agreeing with the statement "Many of the illness which vaccinations prevent are severe." was associated with a higher probability of a caregiver's child being fully vaccinated (aOR = 3.848, 95% CI: [1.412, 10.487]), while delaying a child's vaccination for reasons other than allergy or illness was associated with a lower probability of a caregiver's child being fully vaccinated (aOR = 0.345, 95% CI: [0.165, 0.720]). We also detected a statistically significant association that was not present in the original analysis: agreement with the statement "Healthy children do not need immunizations" was associated with a lower probability of a caregiver's child being fully vaccinated (aOR = 0.333, 95% CI: [0.133, 0.831]).

In our secondary model to assess the association between beliefs/experiences score and measles vaccination status, we used a Firth logistic regression due to the extreme imbalance in measles vaccination status data. Positive beliefs about vaccines (a higher "beliefs" score) were associated with a lower probability of vaccination (aOR = 0.54, 95% CI = [0.39, 0.73]). We also found that positive experiences with vaccines were associated with a higher probability of vaccination (aOR = 1.10), but this association was statistically insignificant (95% CI = [0.82, 1.48]). Even with a Firth regression model, extreme imbalance in the outcome data means that all results of this analysis should be interpreted with caution.

Table 4. Adjusted Odds Ratios and 95% Confidence Intervals of Vaccine Card Model

Variable	Adjusted odds ratio	95% CI	P-value
Belief score	0.979	(0.893, 1.074)	0.6584
Experience score	**1.151**	**(1.048, 1.265)**	**0.0033**
Government Health Centre 3 (vs. Other)	0.846	(0.411, 1.650)	0.634
Government Health Centre 2	0.788	(0.388, 1.513)	0.4884
Private Health Centre or Hospital	1.129	(0.547, 2.209)	0.7321
Visit by Immunization Program	0.837	(0.366, 1.856)	0.6656
Primary school (vs. no school)	**1.783**	**(1.226, 2.585)**	**0.0023**
Secondary school	**1.871**	**(1.244, 2.813)**	**0.0026**
University	**2.555**	**(1.073, 6.815)**	**0.0439**
Married (vs. Unmarried)	**2.031**	**(1.213, 3.389)**	**0.0066**
Divorced	0.951	(0.444, 2.035)	0.8958
Widowed	2.131	(0.502, 11.035)	0.3216
Bednet use	**2.294**	**(1.577, 3.339)**	**< 0.0001**
Inpatient	1.162	(0.876, 1.551)	0.3038

In the model for vaccine card availability, we tested the same covariates as the full vaccination model, that is: caregiver beliefs, caregiver experiences, vaccination location, caregiver education, caregiver marital status, child bednet use, and history of inpatient care. Like the main analysis, variables significantly associated with vaccine cards include caregivers at the secondary school level, married status, and child bednet use. However, vaccination location, history of child inpatient care, and caregiver education level at primary school or university were not significantly associated with vaccine card records. Importantly, we found that vaccine card availability was significantly associated with positive caregiver experiences as an aggregated score, after adjusting for vaccination location, caregiver education level, caregiver marital status, child bednet use, history of inpatient care, and caregiver beliefs (Table 4).

Discussion

We found that more positive beliefs and experiences were associated with full vaccination status, although they were not statistically significant, after adjusting for vaccination location, caregiver education, caregiver marital status, bednet use, and child history of inpatient care from the weighted logistic regression. However, we found that responses to three individual belief questions and three individual experience questions were associated with full vaccination status. For all six of these questions, holding the more vaccine-positive belief or having a more vaccine positive-experience was associated with a greater likelihood of a caregiver's child being fully vaccinated. When restricting analysis to those with vaccine cards, two of these beliefs/experiences retained a statistically significant association. Specifically, believing that “Many of the illness which vaccinations prevent are severe.” was associated with a greater probability of a caregiver’s child being fully vaccinated, while delaying a child’s vaccination for reasons other than allergy or illness was associated with a lower probability of a caregiver’s child being fully vaccinated. Because these two associations were observed at a statistically significant level in both analyses, we believe they are the most likely ones to reflect real-world phenomena.

Previous research also suggests an association between certain vaccine-positive beliefs/experiences and vaccination status, with one study about influenza vaccination (Shahrabani and Benzion, 2012) showing that individuals with positive beliefs about the benefits of vaccines were more likely to be vaccinated. Collectively, this evidence may suggest that survey design plays a critical role in determining if an association between an overall vaccine-positive belief/experience score and vaccination status is detected. If suitable questions are chosen (such as the six for which we detected a statistically significant association), an association may be observed between an aggregated beliefs/experiences variable and vaccination status. However, if questions are chosen poorly, potential associations between a few individual experiences/beliefs and vaccination status, but such associations may remain undetected when examining the association between overall experience/belief score and vaccination status. Because determining suitable survey questions about experiences/beliefs is challenging, it may be preferable to examine associations between individual responses and vaccination status; however, such an approach may be challenging due to the number of covariates involved and possible violations of the collinearity assumption in regression models.

In terms of the confounding variables in our study, the main vaccination locations, government health centres and private health centre, were negatively associated with fully vaccinated status, compared to other locations. However, caregiver’s highest education level as secondary school, married status, use of a bed net, and history of the child staying overnight were all positively associated with fully vaccinated status compared to no school, unmarried status, no use of bednet, and no inpatient history respectively. Similarly, vaccine cards were positively associated with caregiver secondary school education, married status, and bed net use, strengthening the relationship that sociodemographic factors and healthy practices are associated with childhood vaccination uptake in Bugoye, Uganda.

For model comparison, the weighted random forest demonstrated better performance than the weighted logistic regression based on the ROC curve, accuracy, and Kappa. A study that examined the association between low birth weight and some maternal and fetal factors using an imbalanced dataset where only 9.5% of cases were low birth weight similarly found that random forest outperformed logistic regression (Ahmadi et al., 2017). In addition, we conducted the sensitivity analysis which excluded the recall question, and the adjusted odds ratio for beliefs became smaller than 1, suggesting that more negative beliefs were associated with fully vaccinated status. This indicates that recall bias may have masked a true negative association between fully vaccinated status and beliefs in the main analysis.

There are some limitations in this study. A significant percentage of participants (32.03%) did not have a vaccination card available, suggesting the accuracy of our analysis was limited. Among participants without a vaccination card, just 57.33% recalled receiving all vaccines, followed by 39.33% of those who recalled receiving most vaccines and 3.34% of those who recalled non vaccinated. That likely contributes to the observed difference in the results between the main analysis and sensitivity analysis. Moreover, another challenge was handling the imbalance in the data. We used a weighted logistic regression, weighted random forest, or Firth regression model, but it should be noted that the results may still be affected by the predominance of the majority class.

References

Ahmadi, P., Alavimajd, H., Khodakarim, S., Tapak, L., Kariman, N., Amini, P., & Pazhuheian, F. (2017). Prediction of low birth weight using Random Forest: A comparison with Logistic Regression. Archives of Advances in Biosciences, 8(3), 36-43.

Boyce RM, Delamater P, Muhindo R et al. Accessible metrics of access: Novel tools to measure immunization coverage in rural sub-Saharan Africa [version 1; peer review: 1 approved, 1 approved with reservations]. Gates Open Res 2019, 3:1540 (https://doi.org/10.12688/gatesopenres.13066.1)

Shahrabani, S., & Benzion, U. (2012). How experience shapes health beliefs: the case of influenza vaccination. Health education & behavior, 39(5), 612-619.

Uganda Bureau of Statistics (UBOS) and ICF. (2018). Uganda Demographic and Health Survey 2016. Kampala, Uganda and Rockville, Maryland, USA: UBOS and ICF.