Growth-supporting food consumption in 6-23 months children in northern Benin and minimum diet diversity prediction with supervised machine learning approaches.

Abstract

Inadequate complementary feeding remains a major driver of child undernutrition in sub-Saharan Africa. Yet, the diet quality of infants and young children in northern Benin is poorly documented. This study monitored growth-supporting food consumption among children aged 6–23 months in the commune of Banikoara and assessed supervised machine learning (ML) approaches for predicting achievement of Minimum Dietary Diversity (MDD). A descriptive longitudinal study was conducted from November 2023 to July 2024 in five districts.Data were collected over five rounds using a non-quantitative 24-hour recall based on the Infant and Young Child Feeding Diet Quality Questionnaire applied to 149 children. Multivariate Generalized Estimating Equations identified determinants of MDD, Egg and Flesh Food (EFF) and Fruit and Vegetable (FV) consumption.Meanwhile, seven supervised ML algorithms predicted MDD from contextual variables (living district, socio-cultural group and month) under a 10-fold cross-validated and subject-level split. MDD prevalence declined significantly over time (OR = 0.83; 95% CI: 0.74–0.94) whereas EFF (OR = 1.32; 1.18–1.47) and FV (OR = 3.24; 2.16–4.85) consumption increased. Living in Soroko (OR = 6.0) and Somperekou (OR = 2.52) district favoured MDD whereas Gourmantche children (OR = 0.50) were disadvantaged. Decision Tree, Extreme Gradient Boosting and Naive Bayes performed best, with AUC of 0.74–0.75, accuracy of 0.74 and sensitivity of 0.89.ML models built on routine socio-demographic data can support affordable two-stage screening of at-risk children within community nutrition programmes such as the Positive Deviance/Hearth approach in northern Benin.

Keywords: Complementary feeding, Infant and young child feeding, Dietary diversity, Longitudinal study, Positive Deviance/Hearth.

Introduction

Dietis deeply linked tonutritional status. A balanced diet contributed to an optimal nutritional status. Infant malnutrition could be tackled by improving infant diet quality.Diet quality (DQ) is evaluated through its macronutrients balance, essential nutrientscontent to cover needs and diversity (Raru et al., 2023; Shang et al., 2020).DQ could be evaluated with (a) Acceptable Macronutrients Distribution Ranges (AMDRs) which measuresthe proportion of total energy derived from carbohydrates, proteins and fats, (b) Healthy Diet Score (HDS)which measures the consumption of healthy food groupsand (c)Diet Diversity Score (DDS) which is the count of food groups consumed(Raru et al., 2023; Van et al., 2022; Shang et al., 2020).

DDS remains a good proxy measurement to evaluate the nutrients intake adequacy in a study population (Diop et al., 2025). Its use has interest in evaluating the consumption of micronutrient-dense food intake (e.g. (1) legume and fruit and (2) egg and flesh foods) (del Valle M et al., 2024; Kennedy, 2009). A total of eight food groups (i.e. cereals/roots/tubers, legumes/nuts, dairy products, flesh foods (meat/fish/poultry), eggs, vitamin-A-rich fruits, other fruits, and vitamin-A-rich vegetables) are consideredto evaluate the DDS of children(Setegn& Dejene, 2025). Minimum Dietary Diversity (MDD) is achieved when a child consumes at least 5 of these eight food groups within a 24-hour period (Raru et al., 2023). As Diet quality index, MDD is anInfant and Young Child Feeding (IYCF) key indicators to evaluate of food intake in low incomes regions (Raru et al., 2023; Setegn& Dejene, 2025).In sub-Sahara, the prevalence of MDD among children aged 6-23 months rangesfrom 5.6% (Burkina Faso) to 43.9% (South Africa)(Raru et al., 2023). Most factors affecting MDD are food environment, household wealth,geographic context,maternal education and access to healthcare (Ebido et al., 2025; Hailu et al., 2025).

Moreover, recent indicators such as Egg and Flesh Food (EFF) consumption and Zero Vegetable or Fruit (ZVF) intake have been introduced to monitor the consumption of critical growth -supporting foods (Hailu et al., 2022). Only 30% of children (6-23 months) consumed egg and flesh foods in Benin (Zemene et al., 2024). Children protein intake is mainly provided by leguminous. Essential micronutrients (Iron, Zinc, Vitamine A) as well as essential fatty acids and high-quality proteins are provided by EFF (Hailu et al., 2022).

FV intake remains low in Benin although there is a variety of leafy vegetable (African eggplant, amaranth, vernonia leaves, cowpea leaves and cassava leaves, baobab leaves and kenaf leaves) and some fruits (e.g. tomato, mango and okra fruit) are the most consumed plant parts valorised as vegetable and locally found in Benin in infant nutrition (Mitchodigni-

Houndoloa et al., 2024; Koukou et al., 2022; Honfo et al., 2010). FV consumption is limited by (a) accessibility and affordability, (b) seasonality and environments constraints and (c) consumer food habits and food preferences (Diatta et al., 2025; Mitchodigni-Houndolo et al., 2024). Essential micronutrients, beneficial phytochemical (carotenoid) and dietary fiber are the benefit provided by FV (Mitchodigni-Houndolo et al., 2024; Hailu et al., 2022).

The use of machine learning (ML) is a toolkit that helps tackling global challenge related to malnutrition (e.g. undernutrition, micronutrients deficiency, obesity and dietary diversity and food intake) (Setegn& Dejene, 2025; Qasrawi et al., 2024; Ndagijimana et al., 2023; Bitew et al., 2022; Khan et al., 2019). Compared to traditional statistical models (e.g. Logistic regression), ML demonstrates higher prediction accuracy allowing (1) hierarchisation of risk factors (2) earlier identification of subjects at risk and (3) conception and execution of earlier interventions (Qasrawi et al., 2024; Bitew et al., 2022; Fenta et al., 2021; Khan et al., 2019).

ML are usually evaluated based on their (1) discriminating power (area under the receiver operating characteristic curve), (2) overall classification (or prediction) accuracy and (2) ability to misclassify (sensitivity and specificity) (Ndagijimana et al., 2023; Van et al., 2022; Fenta et al., 2021). Random forest, gradient boosting, logistic regression, support vector machine, KNN and decision tree are the most popular ML techniques used on nutrition data (Qasrawi et al., 2024; Fenta et al., 2021; Khan et al., 2019).

The integration of modelling techniques (i.e. supervised ML) with the assessment of diet quality is a growing field in public health, used to predict nutritional risks, identify leading dietary determinants of disease, and guide policy interventions (Bitew et al., 2022). Unfortunately, its use in African countries (e.g. Benin) is still new and should be encouraged. This paper aims to monitor diet quality in northern Benin and to predict MDD. In alongitudinal study on the diet quality on 6-23 months children in northern Benin, two growth-supporting food groups (i.e. EFF and FV) were monitored. In addition, the MDD was predicted using supervised ML approaches.

Materials and methods

A descriptive longitudinal study was conducted from November 2023 to July 2024 in the commune of Banikoara in northern Benin (Figure 1). The objective was tomonitor the diet quality of children aged 6 to 23 months. The target population was households with at least one child aged from 6 to 13 months at the time of the first round of the surveyin Novembre 2023.

The criteria for household inclusion were as follows: (1) residing in commune of Banikoara during the study period, (2) having at least one child aged from 6 to 13 months at the beginning of the study, (3) giving informed consent to participate in the study.

The criteria for household exclusion were as follows: (1) residing outside the commune of Banikoara, (2) having children whose age was outside 6 to 13 months at the beginning of the study, (3) children whose parents did not give informed consent to participate in the study, (4) children whose parents did not participate in all the interviews planned for the study and (5) children with medical conditions that may influence their eating habits the day before the interview.

The districts concerned by thestudy were Banikoara, Founougo, Kokiborou, Sompérékou and Soroko (Figure 1).In each district, simple random sampling was used to select the houses for the door-to-door visits.Households eligible in the visited houses were selected and interviewed. When no household was eligible in the selected house, the next eligible house on the right was selected as substitute.

A nonquantitative 24 h recall using a list-based questions were used to collect data on the complementary feeding practices of children aged 6 to 23 months. The list-based questionscollected information on the 29 foods groups according to the Infant and Young Child Feeding (IYCF) Diet Quality Questionnaire (DQQ) (Herforth et al., 2024). The country adapted IYCF DQQ for Benin was downloaded on the Global Diet Quality Project(https://www.dietquality.org/countries/ben). Additional data on the social cultural group, matrimonial status of the head of the household, living district and gender of the children were collected as well only on the first round of the survey since they remained unchanged during the study. Data was collected in the last week of November 2023, January 2024, March 2024, May 2024 and July 2024. The main local language spoken during the survey is Baatonou.

A total of 153 children aged from 6 to 13 months were randomly chosen in the five (05) districts of the commune of Banikoara and 4 children were excluded based on the criteria for household exclusion. Data of 149 children were processed and used for data analysis.

Ethical consideration

The ethical clearance was obtained from the Ethics Committee of the University of Parakou. The children's parents were informed about the conduct of the study, the objectives of the study, and the data collection procedures. Participation was voluntary, and participants could withdraw at any time without penalty. Verbal consent was obtained, followed by written informed consent. No financial incentive was offered for participation in this study.

Data analysis

Data were compiled and computed in Excel according to the instructions provided by WHO & UNICEF (2021). Monitored indicators were (1) Egg and Flesh Food (EFF) consumption and (2) Fruit and Vegetable (FV) consumption and (3) Minimum Dietary Diversity (MDD). MDD was defined as the consumption of at least five groups of food out of the eight defined food groups (i.e. breast milk, starchy based food, pulses, dairy products, flesh foods, eggs, vitamin-A rich fruits and vegetables, others fruit and vegetables) during the 24 h recall.EFF, FV and MDD were defined as binary dietary diversity indicators and were analysed with generalised estimated equations (GEE) using a binomial family with logit link function and an autoregressive AR(1) working correlation structure. For association inference, a single multivariate GEE logistic model was fitted for MDD including all the predictors (i.e.children gender, living district, socio cultural group, marital status of head of household and time point). Regression coefficients (β) and odds ratio (OR) with 95% confidence intervals (CI) and two-sided p-values were reported. Statistical significance was set at 5%. Reference categories were Female, Bariba, Banikoara, Monogamous for (a) children gender, (b) socio cultural group, (c) living districtand (c) marital status of head of household, respectively. GEE analysis was performed in R 4.6.1 with the package geepack.

Classification training and prediction of MDD was performed using the software Orange Data Mining 3.39.0.Using one-vs-one approach, variables (i.e. socio cultural groups, living district and data collection month) were used predict the achievement of MDD. To prevent information leakage, from repeated measurements on the same subject, data were split at subject level. Data of 104 children (54 female and 50 male) were used to train models. The data of the remaining 45 children (23 female and 22 male) were used to predict MDD outcome.A 10-fold cross validation was used on the train set to train the models. Four groups of classificationmodels were tested on the data. Tree-based model (i.e. decision treeand extreme gradient boosting), distance-based model (K-nearest neighbours), margin-basedmodel (support vector machine), probabilistic classifier (naive bayes),linear model (logistic regression) and deep learning model (artificial neural network) were tested.

The models were evaluated based on the precision, sensitivity (or recall), F1- score, specificity, accuracyand Area Under ROC Curve (AUC).

Precision is the percentage of the Positive Samples Accurately Identified (PSAI) by the model out of the Predicted Positive Samples (PPS). High precision means good ability to predictpositive samplesaccurately.

$$Precision = \frac{PSAI}{PPS} \times 100$$

Sensitivity (or recall) is the percentage of Positive Samples AccuratelyIdentified (PSAI) by the model out of the Total Positive Samples (TPS). High sensitivity means low chance to misclassify positive samples.

$$Sensitivity = \frac{PSAI}{TPS} \times 100$$

F1-score is meant for minimizing misclassified negative samples whilemaximizing Positive Samples Accurately Identified (PSAI). High F1-score (F1-score >0.7) means good performance and good balance between precision and sensitivity.

$$F1 - score = \frac{2 \times precision \times sensitivity}{precision\ + sensitivity} \times 100$$

Specificity is the percentage of Negative Samples Accurately Identified (NSAI) by the model out of the Total Negative Samples TNS. High specificity means low chance to misclassify negative samples.

$$Specificity = \frac{NSAI}{TNS} \times 100$$

Accuracy is the percentage of samples correctly identified by the models out of the total samples.

$$Accuracy = \frac{PSAI + NSAI}{Total\ samples} \times 100$$

AUC compares overall performance of models. The overall performance of a model was evaluated as false, poor, moderate, high, and very high for (a)$0.5< AUC \leq 0.6$, (b) $0.6< AUC \leq 0.7$, (c)$0.7< AUC \leq 0.8$, (d) $0.8< AUC \leq 0.9$, (e) $0.9<AUC\leq1$, respectively(Kannan & Menaga, 2022; Thamrin et al., 2021).

Results

Table 1 presents the characteristics of the population of study. Of the 149 children selected for the study, 77 (51.7%) were female, while 72 (48.3%) were male. Children from the Bariba socio-cultural group represented 83.9% (n = 125) of the sample, followed by Gando (8.7%, n = 13), Gourmantche (4.7%, n = 7), and Zerma (2.7%, n = 4). According to living district, the largest proportions of children were enrolled in Somperekou (34.9%, n = 52) and Kokiborou (32.2%, n = 48). In contrast, Soroko contributed only 6.7% (n = 10). Half of the households (50.3%, n = 75) were monogamous. In addition, polygamous households came in second place with 35.6% (n = 53). Finally, the remaining households, defined by no marital status, represented 14.1% (n = 21).

The Table 2 presents the longitudinal determinants of EFF according to children gender, socio cultural group, living district and marital status of head of household. EFF consumption did not differ between male and female children. However, compared to children from Bariba socio cultural group, EFF consumption was lower among Gourmantche children (OR = 0.30; 95 % CI: 0.15 - 0.59; $p = 0.001$) and Zerma children (OR = 0.35; 95 % CI: 0.20 - 0.62; $p = 0.0003$). Meanwhile, no difference was observed for Gando children. Furthermore, using Banikoara as living district reference, EFF consumption was lower and higher in Founougo (OR = 0.39; 95 % CI: 0.26 - 0.60; $p < 0.001$) and Soroko (OR = 2.07; 95 % CI: 1.16 - 3.68; $p = 0.01$), respectively. In contrast, EFF consumption in Kokiborou and Somperekou remained similar to Banikoara. Moreover, children from polygamous households were less likely to consume EFF than those from monogamous households (OR = 0.73; 95 % CI 0.54–0.99; $p = 0.04$). Overall, the consumption of EFF increased over months (OR = 1.32; 95 % CI 1.18–1.47; $p < 0.001$).

The time-dependent distribution of EFF consumption across gender, socio-cultural group, living district and marital status is illustrated in Figure 2. Regardless of the variables (i.e. Gender, socio cultural group, living district and marital status of the head of household), the EFF consumption could be sequenced in three parts. First, EFF consumption was (1) low in November and January, then (2) higher in March and May, and finally (3) declining in July.

The Table 3 presents the longitudinal determinants of FV according to children gender, socio cultural group, living district and marital status of head of household. FV consumption did not differ between male and female children. Furthermore, using Banikoara as living district reference, FV consumption was markedly elevated in Somperekou (OR = 12.97; 95% CI 3.13–53.70; $p = 0.0004$) and very strongly elevated in Soroko (OR = 4.7×10^{10}; $p < 0.001$). Meanwhile, FV consumption remained similar in Kokiborou and Banikoara. Likewise, FV consumption remained similar (a) between the socio cultural groups and (b) in all the households independently of the marital status of the household head. Overall, FV consumption increased over time (OR = 3.24; 95% CI 2.16 - 4.85; $p < 0.001$).

The Table 4 presents the longitudinal determinants of MDD according to children gender, socio cultural group, living district and marital status of head of household. MDD prevalence was similar between male and female children (OR = 1.17; 95 % CI 0.85–1.62; p = 0.33). However, compared to children from Bariba socio cultural group, MDD prevalence was lower in Gourmantche children (OR = 0.50; 95 % CI 0.28–0.91; p = 0.02). Meanwhile, no significant differences were observed for Gando and Zerma children. Furthermore, using Banikoara as living district reference, MDD prevalence was similar in Founougo and Kokiborou. In contrast, MDD prevalence was higher in Somperekou (OR = 2.52; 95 % CI 1.40–4.56; p = 0.002) and Soroko (OR = 6.0; 95 % CI 2.03–17.70; p = 0.001). Moreover, MDD prevalence remained similar in all the households independently of the marital status of the household head. Conversely, to EFF and FV, MDD prevalence declined significantly over time (OR = 0.83; 95 % CI 0.74–0.94; p = 0.004). Accordingly, the MDD prediction was performed on the three determinants (i.e. socio cultural group, living district and the month).

The Table 5 presents performance metrics of models predicting one on the main diet quality index (i.e. MDD). Models were compared based on their sensitivity and their overall performance (AUC) since the focus was on MDD achievement. Except SVM, the overall performance of all models was moderate ($0.7 < AUC \leq 0.8$) with sensitivity ranging between 0.87 and 0.89. Tree based model (decision tree, extreme gradient model) and probabilistic classifier (naïve bayes) were the most suitable categories of models to predict MDD achievement over time based only on socio cultural group and living district.

Discussion

This study provides evidence that diet quality among children aged 6–23 months in northern Benin is varied according toliving district, sociocultural groups, and time (months). In addition, it supports that MDD can be predicted with moderate accuracy using a limited set of contextual variables.

Diet quality (especially FV, EFF and MDD) was constantly associated in this study to living districts over time and to a lesser extent to sociocultural groups. Spatial inequality in the local food environment might be a major driver of complementary feeding quality in the study area. Food availability and dietary diversity could vary according to geographic patterns across

West Africa. In addition to the disparity of MDD prevalence between west African countries (e.g. Côte d'ivoire, Niger, Senegal), children in rural area are more likely to fail MDD (Janmohamed et al., 2024). However, the disparity within rural area had never been documented.This gap reflects unequal geographic access to diverse food sources and market systems within in the same study area.In general food system resilience is built on asset, diversification, education, and household support system. Household with (a) bigger assets (land, livestock, income), (b) diverse activities (farming and other small businesses), (c) access to knowledge and (d) access to social safety nets (community-based support and access to credit from established institutions) demonstrate greater resilience during lean period (Dereje et al., 2024; Myeki& Bahta, 2021).

Compared with Bariba children, Gourmantcheand Zermachildren were less likely to consume EFF and less likely to meet MDD, whereas FV consumption did not vary significantly by sociocultural group.It suggests that sociocultural groups remained a determinant factor for evaluating the access to animal-source food. Although Gourmantche and Zerma are pastoralism communities, they view livestock primarily as capital and social status rather than as immediate food resource(Potts et al., 2019).

Children from polygamous households had lower odds of consuming EFF. Household structure could affect access to animal-source protein.Children of polygamous householdsuffered from long-term nutritional outcomes compared to children of monogamous mother(Owoo, 2018).Household-level food security might not ensure equitable access EFF to all family membersof polygamous household.

Most models achieved only moderate discrimination, with AUC values around 0.74–0.75 except the support vector machine model.The models support the finding that dietary adequacy is socially and spatially patterned.

High sensitivity in identifying MDD-positive children ensures that nearly all true positives are correctly flagged, minimizing the loss of potential positive-deviance informants. This property directly supports the Positive Deviance/Hearth methodology and aligns with WHO and SDG monitoring targets, preserving representativeness while reducing recall workload. Although false positives require extra confirmation visits, the operational trade-off is acceptable since missing true positives would undermine both intervention design and progress tracking.

The present findings offer a practical route to substantially leaner monitoring in Positive Deviance (PD)/Hearth programmes. Conventional PD evaluation relies on resource-intensive dietary recalls and anthropometric surveys, repeated across cycles to track rehabilitation gains(Chek et al., 2022a; Levinson et al., 2007). The classifier derived here requires depend only on information that PD facilitators already record during community visits. This suggests that PD/Hearth progress monitoring could shift from full-coverage dietary assessment to a two-stage design in which the model select potential positive deviance before fuller investigation(Chek et al., 2022b).

Study limitations

Several limitations should be considered when interpreting these findings. First, the study was conducted in a single commune in northern Benin, which limits the generalizability of the results. Second, dietary data were collected using caregiver-reported 24-hour recall, which may be affected by recall bias and day-to-day variation in food intake. Third, some key determinants of child diet quality, including maternal education, household income, and food security, were not included in the analysis or the predictive models. Finally, the follow-up period covered only nine months and did not capture repeated annual cycles, which limits the ability to fully characterize seasonal variation in child feeding practices. These limitations suggest that the present findings should be interpreted as context-specific, while also pointing to important directions for future research.

Conclusion

This study provides the first longitudinal characterization of growth-supporting food consumption among children aged 6–23 months in Banikoara, northern Benin, and demonstrates that dietary adequacy in this setting is strongly patterned by geography and sociocultural affiliation.Although consumption of EFF and FV rose over the nine-month period, the overall prevalence of MDD declined. In addition, the disadvantage of (a) Gourmantche and Zerma children and (b) children in polygamous households indicates that household-level food security does not guarantee equitable intra-household allocation of animal-source foods to young children. Interventions must therefore be targeted at the sociocultural and household-structure level. Futhermore, supervised ML analysis shows that MDD can be predicted with moderate discrimination (AUC 0.74–0.75) and high sensitivity (0.89) from just three routine contextual variables, with Decision Tree, Extreme Gradient Boosting and Naive Bayes performing best.Community monitoring programmes such as

Positive Deviance/Hearth could adopt a two-stage design where models first flag children likely to meet MDD, and only those flagged undergo full dietary assessment.

Acknowledgments

The authors are grateful to the participantsfor their collaboration.

CRediT authorship contribution statement

FUGA: Conceptualization, Methodology, Validation, Visualization, Writing – original draft, **RBT**: conceptualization, methodology, investigation, formal analysis, Writing – reviewing & editing, **IBC**: conceptualization, methodology, Writing – reviewing & editing, **JMK**: conceptualization, methodology, Writing – reviewing & editing, **CNAS:** conceptualization, methodology, Writing – reviewing & editing, **FH:** Supervision,Conceptualization, Methodology, Writing – reviewing & editing, **APPK**:conceptualization, methodology, Writing – reviewing & editing.

Funding

This study did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Ethical statement

This study involved interviewing participants in Banikoara.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be available on request.

References

Bitew, F. H., Sparks, C. S., & Nyarko, S. H. (2022). Machine learning algorithms for predicting undernutrition among under-five children in Ethiopia. *Public Health Nutrition*, *25*(2), 269–280.

Chek, L. P., Gan, W. Y., Chin, Y. S., & Sulaiman, N. (2022a). A nutrition programme using positive deviance approach to reduce undernutrition among urban poor children under-five in Malaysia: A cluster randomised controlled trial protocol. *Plos One*, *17*(10), e0275357.

Chek, L. P., Gan, W. Y., Chin, Y. S., & Sulaiman, N. (2022b). A nutrition programme using positive deviance approach to reduce undernutrition among urban poor children under-five in Malaysia: A cluster randomised controlled trial protocol. *Plos One*, *17*(10), e0275357.

del Valle M, M., Shields, K., & Boza, S. (2024). Using Household Dietary Diversity Score and Spatial Analysis to Inform Food Governance in Chile. *Nutrients*, *16*(17), 2937.

Dereje, D., Tilahun, T., &Yirgu, T. (2024). Determinants of Household Resilience to Food Insecurity for Comparison: the Case of Chencha District, Southern Ethiopia. *F1000Research*, *13*, 693.

Diatta, A. D., Olney, D. K., Honeycutt, S., Mitchodigni, I. M., &Bliznashka, L. (2025). *Nutrition and diet profile: Benin*. Intl Food Policy Res Inst.

Diop, L., Gelli, A., Huybregts, L., Arsenault, J. E., Bliznashka, L., Boy, E., Deitchler, M., Lachat, C., Moursi, M., & Ochoa-Avilés, A. M. (2025). The minimum dietary diversity for women indicator can be extended to children and adolescents aged 4-15 years as a proxy population indicator for good micronutrient adequacy of diets in low-and middle-income countries. *Current Developments in Nutrition*, *9*(1), 104508.

Ebido, M. J., Biadgilign, S., Tareke, A. A., &Zerfu, T. A. (2025). Determinants of household food security and maternal dietary diversity in rural Gedeo zone, southern Ethiopia: results from a cross-sectional study. *Frontiers in Nutrition*, *12*, 1523344.

Fenta, H. M., Zewotir, T., & Muluneh, E. K. (2021). A machine learning classifier approach for identifying the determinants of under-five child undernutrition in Ethiopian administrative zones. *BMC Medical Informatics and Decision Making*, *21*(1), 291.

Global Diet Quality Project. (2023). *Diet Quality Questionnaire (DQQ) for Benin*. https://www.dietquality.org/countries/ben

Hailu, A. A., Thornhill, S., Tessema, M., Weldesilassie, A. B., & Lahiff, E. (2025). The influence of the food environment on diet quality: Insights from an extensive household survey in Ethiopia, focusing on women of reproductive age. *BMC Nutrition*, *11*(1), 107.

Hailu, B. A., Geremew, B. M., Liverani, S., Abera, K. S., Beyene, J., &Miheretu, B. A. (2022). Mapping and determinants of consumption of egg and/or flesh foods and zero vegetables or fruits among young children in SSA. *Scientific Reports*, *12*(1), 11924.

Herforth, A. W., Ballard, T., & Rzepa, A. (2024). Development of the diet quality questionnaire for measurement of dietary diversity and other diet quality indicators. *Current Developments in Nutrition*, *8*(8), 103798.

Honfo, F. G., Hell, K., Akissoe, N., Dossa, R., &Hounhouigan, J. D. (2010). Diversity and nutritional value of foods consumed by children in two agro-ecological zones of Benin. *Afr J Food Sci*, *4*(4), 184–191.

Janmohamed, A., Baker, M. M., Doledec, D., Ndiaye, F., Konan, A. C. L., Leonce, A., Kouadio, K. L., Beye, M., Yattara, M. L., &Dissieka, R. (2024). Dietary diversity and associated factors among infants and young children in three West African countries. *Frontiers in Public Health*, *12*, 1386664.

Kannan, K., & Menaga, A. (2022). Risk factor prediction by naive bayes classifier, logistic regression models, various classification and regression machine learning techniques. *Proceedings of the National Academy of Sciences, India Section B: Biological Sciences*, *92*(1), 63–79.

Kennedy, G. L. (2009). *Evaluation of dietary diversity scores for assessment of micronutrient intake and food security in developing countries*. Wageningen University and Research.

Khan, J. R., Chowdhury, S., Islam, H., & Raheem, E. (2019). Machine learning algorithms to predict the childhood anemia in Bangladesh. *Journal of Data Science*, *17*(1), 195–218.

Koukou, E., Amoussa-Hounkpatin, W., Savy, M., Ntandou-Bouzitou, G. D., Mitchodigni, M. I., Bodjrènou, F. S. U., Tovissode, C. F., &Termote, C. (2022). Contribution of local agrobiodiversity to complementary foods for 6 to 23 months old children in southern rural Benin. *African Journal of Food, Agriculture, Nutrition and Development*, *22*(5), 20498–20522.

Levinson, F. J., Barney, J., Bassett, L., &Schultink, W. (2007). Utilization of positive deviance analysis in evaluating community-based nutrition programs: An application to the Dular program in Bihar, India. *Food and Nutrition Bulletin*, *28*(3), 259–265.

Mitchodigni-Houndolo, I. M., Iruhiriye, E., Egah, J., Afe, O. H. I., Madode, Y. E., Ayenan, M. A. T., Salaou, A. M., Videgla, E., Schreinemachers, P., & Piters, B. de S. (2024). *The future of fruit and vegetables in Benin's food system: visioning alternative scenarios toward 2060*.

Myeki, V. A., & Bahta, Y. T. (2021). Determinants of smallholder livestock farmers' household resilience to food insecurity in South Africa. *Climate*, *9*(7), 117.

Ndagijimana, S., Kabano, I. H., Masabo, E., & Ntaganda, J. M. (2023). Prediction of stunting among under-5 children in Rwanda using machine learning techniques. *Journal of Preventive Medicine and Public Health*, *56*(1), 41–49.

Owoo, N. S. (2018). Food insecurity and family structure in Nigeria. *SSM-Population Health*, *4*, 117–125.

Potts, K. S., Mulugeta, A., & Bazzano, A. N. (2019). Animal source food consumption in young children from four regions of Ethiopia: association with religion, livelihood, and participation in the productive safety net program. *Nutrients*, *11*(2), 354.

Qasrawi, R., Sgahir, S., Nemer, M., Halaikah, M., Badrasawi, M., Amro, M., Vicuna Polo, S., Abu Al-Halawa, D., Mujahed, D., & Nasreddine, L. (2024). Machine learning approach for predicting the impact of food insecurity on nutrient consumption and malnutrition in children aged 6 months to 5 years. *Children*, *11*(7), 810.

Raru, T. B., Merga, B. T., Mulatu, G., Deressa, A., Birhanu, A., Negash, B., Gamachu, M., Regassa, L. D., Ayana, G. M., & Roba, K. T. (2023). Minimum dietary diversity among children aged 6–59 months in East Africa countries: a multilevel analysis. *International Journal of Public Health*, *68*, 1605807.

Setegn, G. M., & Dejene, B. E. (2025). Improving machine learning models through explainable AI for predicting the level of dietary diversity among Ethiopian preschool children. *Italian Journal of Pediatrics*, *51*(1), 91.

Shang, X., Li, Y., Xu, H., Zhang, Q., Liu, A., Du, S., Guo, H., & Ma, G. (2020). Leading dietary determinants identified using machine learning techniques and a healthy diet score for changes in cardiometabolic risk factors in children: a longitudinal analysis. *Nutrition Journal*, *19*(1), 105.

Thamrin, S. A., Arsyad, D. S., Kuswanto, H., Lawi, A., & Nasir, S. (2021). Predicting obesity in adults using machine learning techniques: an analysis of Indonesian basic health research 2018. *Frontiers in Nutrition*, *8*, 669155.

Van, V. T. S., Antonio, V. A., Siguin, C. P., Gordoncillo, N. P., Sescon, J. T., Go, C. C., & Miro, E. P. (2022). Predicting undernutrition among elementary schoolchildren in the Philippines using machine learning algorithms. *Nutrition*, *96*, 111571.

WHO, & UNICEF. (2021). *Indicators for assessing infant and young child feeding practices: definitions and measurement methods*. https://policycommons.net/artifacts/8937836/indicators-for-assessing-infant-and-young-child-feeding-practices/9754598/

Zemene, M. A., Kebede, N., Anteneh, R. M., Moges, N., Tsega, S. S., Dessie, A. M., Belete, M. A., Anley, D. T., Alemayehu, E., & Chanie, E. S. (2024). Determinants of animal source food consumption among children aged 6–23 months in sub-Saharan Africa: multilevel mixed effect model. *Scientific Reports*, *14*(1), 26294.

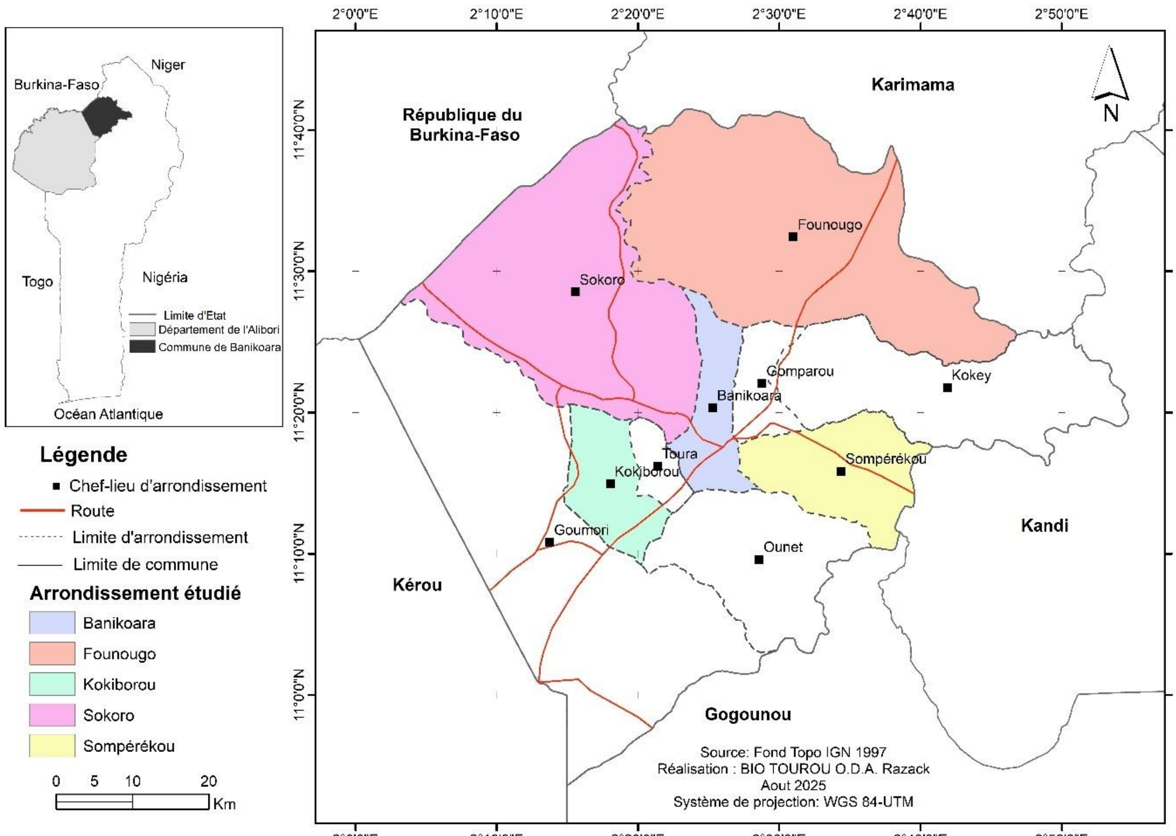

Figure 1: Study area and selected district of Banikoara.

Table 1: Sample characteristics of the population of study

Characteristics	Category	Sample size	Percentage
Gender	Female	77	51.7
	Male	72	48.3
Social cultural groups	Bariba	125	83.9
	Gando	13	8.7
	Gourmantche	7	4.7
	Zerma	4	2.7
Living district	Banikoara	20	13.4
	Founougo	19	12.8
	Kokiborou	48	32.2
	Somperekou	52	34.9
	Soroko	10	6.7
Marital status of the head of household	Monogamous	75	50.3
	Polygamous	53	35.6
	Others	21	14.1

Table 2: Multivariate GEE analysis of EFF fitted with children gender, socio cultural group, living district and marital status of head of household

Variables	β	OR	95% CI	p-value
Gender				
Female	-	-	-	-
Male	0.26	1.29	0.99 -1.69	0.06
Social cultural groups				
Bariba	-	-	-	-
Gando	−0.19	0.83	0.52 - 1.31	0.42
Gourmantche	−1.22	0.30	0.15 - 0.59	0.001
Zerma	−1.04	0.35	0.20 - 0.62	0.0003
Living district				
Banikoara	-	-	-	-
Founougo	−0.93	0.39	0.26 - 0.60	<0.001
Kokiborou	0.05	1.05	0.70 - 1.59	0.80
Somperekou	0.044	1.05	0.71 - 1.54	0.83
Soroko	0.73	2.07	1.16 - 3.68	0.01
Marital status				
Monogamous	-	-	-	-
Polygamous	−0.31	0.73	0.54 - 0.99	0.04
Others	0.02	1.02	0.73 - 1.45	0.89
Months	0.28	1.32	1.18 - 1.47	<0.001

GEE: Generalised estimated equations

EFF: Egg and flesh food

Table 3: Multivariate GEE analysis of FV fitted with children gender, socio cultural group, living district and marital status of head of household

Variables	β	OR	95% CI	p-value
Gender				
Female	-	-	-	-
Male	−0.27	0.76	0.45 - 1.30	0.32
Social cultural groups				
Bariba	-	-	-	-
Gando	−0.42	0.66	0.276 - 1.56	0.34
Gourmantche	0.86	2.36	0.42 - 13.30	0.33
Zerma	−2.29	0.10	0.009 - 1.18	0.07
Founougo	−0.79	0.45	0.18 - 1.17	0.10
Living district				
Banikoara	-	-	-	-
Kokiborou	−0.31	0.73	0.26 - 2.10	0.56
Somperekou	2.56	12.97	3.13 - 53.70	0.0004
Soroko	24.57	4.7×10^{10}	-	<0.001
Marital status				
Monogamous	-	-	-	-
Polygamous	0.035	1.04	0.55 - 1.95	0.91
Others	0.52	1.68	0.71 - 3.95	0.23
Time	1.17	3.24	2.16 - 4.85	<0.001

GEE: Generalised estimated equations

FV: Fruit and vegetable

Table 4: Multivariate GEE analysis of MDD fitted with children gender, socio cultural group, living district and marital status of head of household

Variables	β	OR	95% CI	p-value
Gender				
Female	-	-	-	-
Male	0.16	1.17	0.85 - 1.62	0.33
Social cultural groups				
Bariba	-	-	-	-
Gando	0.07	1.08	0.60 - 1.94	0.80
Gourmantche	−0.69	0.50	0.28 - 0.91	0.02
Zerma	−0.53	0.59	0.19 - 1.80	0.35
Living district				
Banikoara	-	-	-	-
Founougo	−0.43	0.65	0.362 - 1.160	0.14
Kokiborou	0.21	1.23	0.667 - 2.261	0.51
Somperekou	0.92	2.52	1.395 - 4.558	0.002
Soroko	1.79	6.0	2.03 - 17.70	0.001
Marital status				
Monogamous	-	-	-	-
Polygamous	−0.04	1.0	0.68 - 1.35	0.80
Others	0.55	1.74	1.0 - 3.04	0.05
Months	−0.18	0.83	0.74 - 0.94	0.004

GEE: Generalised estimated equations

MDD: Minimum Dietary Diversity

Table 5: Precision, sensitivity (or recall), F1- score, specificity, accuracy and Area Under ROC Curve of models predicting the achievement of MDD.

Model	**AUC***	**Accuracy**	**F1-score**	**Precision**	**Sensitivity (Recall)**	**Specificity**
SVM	0.65	0.54	0.60	0.74	0.51	0.62
K-Nearest Neighbours (kNN)	0.74	0.74	0.82	0.78	0.87	0.45
Artificial Neural Network (ANN)	0.75	0.74	0.82	0.77	0.88	0.42
Logistic Regression	0.75	0.74	0.82	0.77	0.88	0.42
Tree	0.74	0.74	0.83	0.77	0.89	0.42
Extreme Gradient Boosting (EGB)	0.75	0.74	0.83	0.77	0.89	0.42
Naive Bayes	0.75	0.74	0.83	0.77	0.89	0.42

*AUC: Area Under ROC Curve