IDSA
Infectious Diseases Society of America

hivma
hiv medicine association

OXFORD

Derivation and External Validation of a Clinical Prediction Model for Viral Diarrhea Etiology in Bangladesh

Stephanie Chow Garbern,[1,] Md. Taufiqul Islam,[2] Kamrul Islam,[2] Sharia M. Ahmed,[3] Ben J. Brintz,[4] Ashraful Islam Khan,[2] Mami Taniuchi,[5] James A. Platts-Mills,[5,] Firdausi Qadri,[2] and Daniel T. Leung[3]

[1]Department of Emergency Medicine, Alpert Medical School, Brown University, Providence, Rhode Island, USA, [2]Infectious Diseases Division, icddr,b, Dhaka, Bangladesh, [3]Division of Infectious Diseases, University of Utah School of Medicine, Salt Lake City, Utah, USA, [4]Division of Epidemiology, University of Utah School of Medicine, Salt Lake City, Utah, USA, and [5]Division of Infectious Diseases and International Health, University of Virginia, Charlottesville, Virginia, USA

Background. Antibiotics are commonly overused for diarrheal illness in many low- and middle-income countries, partly due to a lack of diagnostics to identify viral cases, in which antibiotics are not beneficial. This study aimed to develop clinical prediction models to predict risk of viral-only diarrhea across all ages, using routinely collected demographic and clinical variables.

Methods. We used a derivation dataset from 10 hospitals across Bangladesh and a separate validation dataset from the icddr,b Dhaka Hospital. The primary outcome was viral-only etiology determined by stool quantitative polymerase chain reaction. Multivariable logistic regression models were fit and externally validated; discrimination was quantified using area under the receiver operating characteristic curve (AUC) and calibration assessed using calibration plots.

Results. Viral-only diarrhea was common in all age groups (<1 year, 41.4%; 18–55 years, 17.7%). A forward stepwise model had AUC of 0.82 (95% confidence interval [CI], .80–.84) while a simplified model with age, abdominal pain, and bloody stool had AUC of 0.81 (95% CI, .78–.82). In external validation, the models performed adequately although less robustly (AUC, 0.72 [95% CI, .70–.74]).

Conclusions. Prediction models consisting of 3 routinely collected variables can accurately predict viral-only diarrhea in patients of all ages in Bangladesh and may help support efforts to reduce inappropriate antibiotic use.

Keywords. antimicrobial resistance; Bangladesh; clinical prediction model; diarrhea; viruses.

Despite significant reductions in mortality over the past several decades, diarrheal disease remains the second most common acute condition globally, causing >6 billion episodes and 1.3 million deaths annually [1, 2]. While a major cause of morbidity worldwide, diarrheal diseases disproportionately affect patients in low- and middle-income countries (LMICs) in communities with poor access to healthcare, safe water, and sanitation [2]. Although the majority of diarrhea episodes are self-limiting and the mainstay of treatment is rehydration, clinicians must also make decisions regarding the appropriate use of antibiotics [3]. For most cases of diarrhea, antibiotics are not beneficial,

particularly for viral etiologies of diarrhea in which antibiotics have no role. Guidelines from the World Health Organization (WHO) recommend avoiding antibiotics for treating most cases of diarrhea, and only recommend antibiotics when there is a suspicion of *Vibrio cholerae* infection with severe dehydration, suspicion of *Shigella* infection as indicated by bloody stool, or concurrent illness such as severe malnutrition [4]. However, despite guidelines, overuse of antibiotics for viral diarrhea remains widespread, particularly in LMICs, partly due to a lack of diagnostic testing to guide management, shortages of trained healthcare providers, nonprescription antibiotic availability, and patient expectations for antibiotics [5, 6].

Inappropriate antibiotic use can lead to increased costs, adverse effects (eg, hemolytic uremic syndrome in Shiga toxin–producing *Escherichia coli*), and antimicrobial resistance (AMR), which has been identified as a serious global public health concern [7–9]. Recent studies have demonstrated widespread resistance among numerous enteric pathogens with multidrug-resistant *V cholerae*, *Shigella* spp, and *Campylobacter* identified in Asia, Africa, and Latin America [10–15]. AMR impedes the treatment of patients and outbreak management, and drug-resistant pathogens quickly spread worldwide [16–18]. Ideally, antibiotic use decisions would be guided by molecular diagnostics or stool cultures; however, these tests are unavailable in the vast majority of LMICs

[19, 20]. As a result, clinicians often use syndromic guidelines or clinical suspicion to decide on antimicrobial use [4, 20]. Unfortunately, physician judgment and syndromic guidelines poorly predict need for antibiotics; a study of patients with diarrhea in Kenya showed that syndrome-based guidelines for *Shigella* led to the failure to diagnose shigellosis in nearly 90% of cases [21]. Clinician assumptions regarding etiologies of diarrhea that are not evidence-based may lead to poor management decisions. Better tools to assist clinician decision-making, that do not rely on costly laboratory tests, are urgently needed.

Much of the existing research on diarrhea epidemiology in LMICs has focused primarily on children <5 years old, while there remains a paucity of data on diarrheal etiology in older individuals [22, 23]. Prior research has shown that viral etiologies of diarrhea (eg, rotavirus, norovirus) are most common in younger children, whereas bacterial causes predominate in older children, adults, and the elderly globally [24–27]. However, recent research using multiplex molecular platforms suggests that viral cases of diarrhea in older individuals may be more common than previously described, with adults remaining highly susceptible to viral enteric infections [28, 29]. Additionally, there is scant evidence on the clinical predictors most associated with viral etiologies in older children and adults that may help clinicians identify patients with viral diarrhea who do not warrant antibiotic use. Clinical tools to better determine diarrhea etiology at the point-of-care without relying on laboratory tests are greatly needed to reduce antibiotic overuse while conserving scarce healthcare resources. The aim of this study was to develop and validate new clinical prediction models to predict the risk of viral-only diarrhea etiology among patients of all ages to assist clinician decision-making for appropriate use of antibiotics.

METHODS

Study Design and Setting
This was a retrospective study using datasets from 2 established diarrhea epidemiology surveillance systems in Bangladesh: (1) Institute of Epidemiology Disease Control and Research nationwide diarrhea surveillance network, which included data from 10 sentinel hospital sites across Bangladesh; and (2) the routine diarrheal surveillance system at the icddr,b in Dhaka, Bangladesh (Supplementary Figure 1). The icddr,b provides free clinical services to >100 000 patients with diarrhea each year in the capital city of Dhaka and surrounding rural districts and is not part of the nationwide surveillance network. The 10 hospital sites are public, mostly district-level hospitals, which provide free medical services to their surrounding catchment areas.

Study Populations: Derivation and Validation Datasets
The nationwide surveillance network (n = 2516) was used as a derivation dataset and the icddr,b diarrhea surveillance (n =

3000) as the validation dataset. In the derivation dataset collected from February 2014 through June 2018, the first patient <5 years of age and the first patient >5 years of age was enrolled each week at each of the 10 sites. In the validation dataset collected during August 2014–June 2017, every 50th patient presenting with acute diarrhea was enrolled. Diarrhea was defined as ≥3 loose or liquid stools within 24 hours or ≤3 liquid stools causing dehydration in the last 24 hours; for children <2 months of age, diarrhea was defined as a change in stool habits from the usual frequency or nature of stool [30]. Patients who met the case definition, had a stool sample with diagnostic testing performed, and had a valid test result, with no other severe comorbidity (eg, respiratory illness, acute cardiovascular symptoms, or severe neurological disorder), were included.

Predictor Variables
All demographic and clinical variables collected during routine clinical care and available in both derivation and validation datasets were considered for inclusion in the models, including age (predetermined age categories <12 months, 12–23 months, 24–59 months, 5–17 years, 18–55 years, and >55 years), sex (male or female), duration of diarrhea (days), severe diarrhea (yes/no >10 episodes of diarrhea in the past 24 hours), dehydration status (none, some, severe), history of vomiting (yes/no), bloody stool (yes/no), and abdominal pain (yes/no). Dehydration status was determined as none, some, or severe dehydration by the treating physician following WHO criteria [31]. Given the known association between season and diarrhea epidemiology in Bangladesh, season was included as a predictor using the patient encounter date. Prior research on diarrhea seasonality has considered Bangladesh to have 3 primary seasons: winter (November–February, cool/dry), summer (March–May, hot/dry), and monsoon (June–October, rainy/wet) [32, 33].

Microbiological Testing and Attribution of Diarrhea Etiology
A stool sample was collected from each enrolled patient. Custom multiplex TaqMan Array Cards (TACs) containing compartmentalized probe-based quantitative real-time polymerase chain reaction (qPCR) assays for 32 pathogens were used to perform qPCR assays for a broad range of pathogens at the icddr,b laboratory (pathogen targets are listed in Supplementary Table 1); a full description of TAC laboratory methods have been previously described [34].

The primary outcome (dependent) variable was defined as identification of only viral pathogen(s) ("viral-only") (ie, no bacterial or parasitic pathogens) on TAC PCR. This clinically relevant outcome was selected because patients with viral-only diarrhea do not warrant antibiotics. Stool samples without detection of any viral targets or evidence of coinfection of viral and bacterial or protozoal pathogen based on TAC results were considered to have nonviral-only diarrhea. A TAC cycle

threshold (Ct) of <30 was used to define a positive result for pathogen detection based on expert consensus.

Data Analysis, Model Development, and Validation

Descriptive analyses using frequencies with percentages, medians with interquartile ranges (IQRs), or means with standard deviations were performed as appropriate. Comparisons between the derivation and validation datasets were conducted with Pearson χ^2 test or Mann-Whitney U test. Univariable logistic regression was performed to assess the unadjusted associations between viral diarrhea etiology and predictors, with magnitudes of effect given as odds ratios (ORs) and their respective 95% confidence intervals (CIs).

Three multivariable logistic regression models were fit using clinically relevant candidate predictors available from the derivation dataset, with each model's included predictors shown in Supplementary Table 2. Model 1 (full model) contained all clinically relevant predictors available in both the derivation and validation datasets and the season variable; model 2 (forward stepwise model) was fit using forward stepwise selection using $P < .1$ for entry and $P > .2$ for removal; model 3 (simplified model) contained only 3 predictors (age, bloody stool, abdominal pain) selected a priori based on diarrhea syndromic guidelines regarding viral versus bacterial causes of diarrhea [4, 35, 36]. The predictors' adjusted associations with viral diarrhea etiology were expressed as adjusted ORs (aORs) and their respective 95% CIs. The models were then externally validated using the icddr,b surveillance dataset.

Model performance was calculated using the area under the receiver operating characteristic curve (AUC) to evaluate discrimination. The Delong method was used to calculate 95% CIs around the AUC. Calibration was assessed by comparing the predicted probability to the observed probability of viral-only diarrhea and reported using calibration-in-the-large and calibration slope [37]. Flexible calibration curves were obtained using locally estimated scatterplot smoothing using val.-prob.ci.2 in R software as recommended in the literature [38]. Nagelkerke pseudo R^2 was calculated to provide a global measure of the estimated explained variance of the models on a new dataset. From the AUC curves, probability cutoffs were used to calculate sensitivity and specificity and positive and negative predictive values [39, 40]. R software (R Foundation for Statistical Computing, Vienna, Austria) was used for all analyses. The Transparent Reporting of a Multivariable Prediction Model for Individual Diagnosis (TRIPOD) Checklist for Prediction Model Validation was used.

Sensitivity Analysis

Two sensitivity analyses were conducted. First, an alternative categorization for age was explored using classification and regression trees (CART) from step_discretize_cart in the R package embed (version 1.0.0) to optimally discretize the age variable using supervised binning. Second, different Ct values were explored to evaluate the effect of using TAC PCR Ct cutoff thresholds of <25 and <35 for attribution of diarrhea etiology.

RESULTS

Participant Characteristics

In the derivation dataset, 2532 patients had samples available for testing and 2516 (99.4%) had valid results. In the validation dataset, 3154 patients had samples available for testing, of whom 3000 (95.1%) had valid results for inclusion in the analysis (Figure 1). The derivation cohort had slightly more females, patients with severe dehydration, and older patients (median age, 22 [IQR, 1–40] years in the derivation dataset compared to 7 [IQR, 0–32] years in the validation cohort). Further characteristics of the study populations are shown in Table 1.

Diarrhea Etiology

The prevalence of viral-only etiology of diarrhea was similar between the 2 study cohorts with 802 (31.9%) of patients in the derivation dataset and 1016 (33.9%) of patients in the validation dataset having viral-only pathogens detected (Supplementary Table 3). Coinfection of viral with nonviral pathogens was found in 104 (4.1%) of the derivation cohort and 309 (10.3%) of the validation cohort. Pathogens detected using TAC PCR are listed in Supplementary Table 3.

In unadjusted analysis, longer duration of diarrhea (OR, 1.41 [95% CI, 1.32–1.49]) and having severe dehydration (OR, 1.76 [95% CI, 1.26–2.46] for severe vs no dehydration) predicted higher odds of viral-only diarrhea, whereas older age (OR, 0.07 [95% CI, .06–.08] for adults 18–55 years of age vs <12 months), abdominal pain (OR, 0.25 [95% CI, .21–.30]), presenting during monsoon season (OR, 0.4 [95% CI, .37–.55] for monsoon vs winter season), and bloody stool (OR, 0.24 [95% CI, .06–.69]) predicted lower odds of viral-only diarrhea (Supplementary Table 4).

The adjusted association of predictors of viral-only diarrhea etiology in the 3 prediction models are shown in Table 2. In the full model (model 1), longer duration of diarrhea (aOR, 1.11 [95% CI, 1.04–1.20]) and having severe dehydration (aOR, 2.32 [95% CI, 1.53–3.54] for severe vs no dehydration) predicted higher odds of viral-only etiology, while older age (aOR, 0.80 [95% CI, .64–.98] for age >55 years vs <12 months), abdominal pain (aOR, 0.63 [95% CI, .50–.79]), presenting during monsoon season (aOR, 0.59 [95% CI, .47–.75] for monsoon vs winter/dry season), and male sex (aOR, 0.80 [95% CI, .64–.98]) predicted lower odds of viral-only diarrhea (Table 2). In the simplified model (model 3), older age (aOR, 0.06 [95% CI, .04–.10] for age >55 years vs <12 months), abdominal pain (aOR, 0.63 [95% CI, .51–.79]), and bloody stool (aOR, 0.28 [95% CI, .07–1.02]) predicted lower odds of viral-only diarrhea.
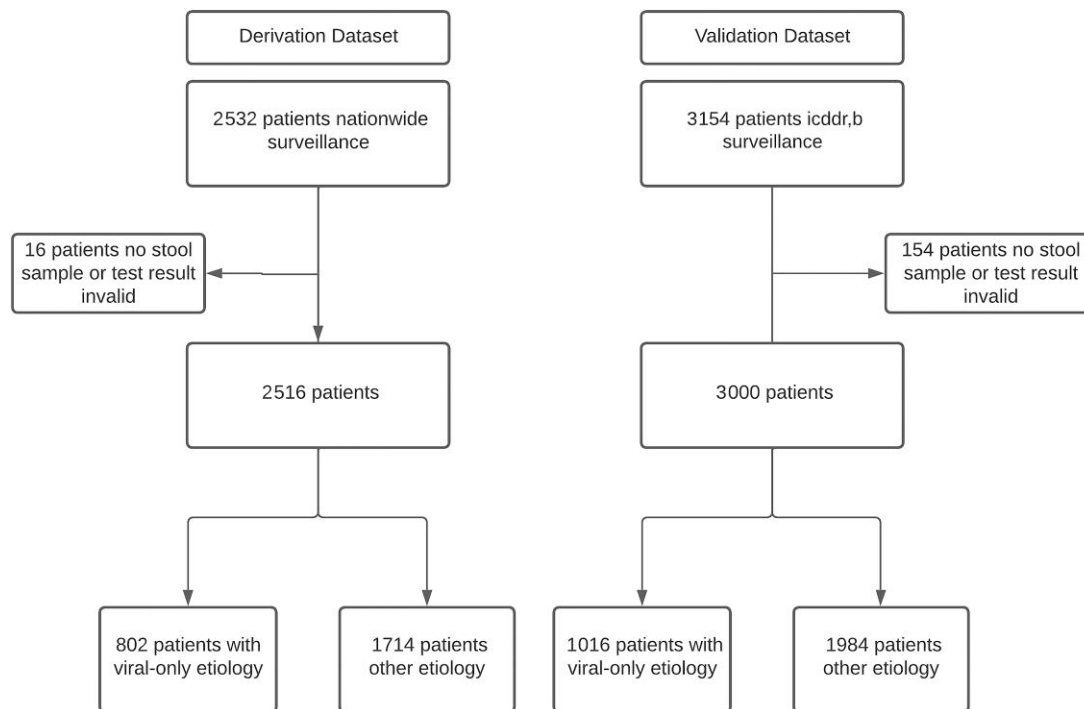
Figure 1. Study flow diagram.

Table 1. Clinical Characteristics of the Study Population

Characteristic	Derivation (Nationwide) (n = 2516)	Validation (icddr,b) (n = 3000)	P Value
Age, y, median (IQR)	22 (1–40)	7 (0–32)	<.01[a]
Age category			<.01[b]
<12 mo	488 (19.4)	950 (31.7)	
12–23 mo	355 (14.1)	401 (13.4)	
24–59 mo	148 (5.9)	122 (4.1)	
6–17 y	136 (5.4)	135 (4.5)	
18–55 y	1159 (46.1)	1241 (41.4)	
>55 y	230 (9.1)	151 (5.0)	
Male sex	1387 (55.1)	1752 (58.4)	.02[b]
Duration of diarrhea, d, median (IQR)	3 (2–3)	2 (1–2)	<.01[a]
Abdominal pain	1563 (62.1)	1762 (58.7)	.01[b]
>10 episodes of diarrhea in 24 h	1821 (72.4)	2170 (72.3)	.97[b]
Blood in stool	29 (1.2)	46 (1.5)	.22[b]
Presence of vomiting	1786 (71.0)	2303 (76.8)	<.01[b]
Dehydration severity			<.01[b]
None	922 (36.6)	983 (32.8)	
Some	1426 (56.7)	1179 (39.3)	
Severe	168 (36.6)	838 (27.9)	
Season			<.01[b]
Winter	733 (29.1)	1106 (36.9)	
Summer	517 (20.6)	957 (31.9)	
Monsoon	1266 (50.3)	937 (31.2)	

Data are presented as No. (%) unless otherwise indicated.

Abbreviation: IQR, interquartile range.

[a]Mann-Whitney *U* test.

[b]χ^2 test.

Model characteristics including AUC and pseudo R^2 were similar between the 3 models with AUC of 0.82 (95% CI, .80–.84) for the full model (model 1) and forward stepwise model (model 2) and AUC of 0.81 (95% CI, .78–.82) for the simplified model (model 3) as shown in Table 3. Pseudo R^2 was similar for all models at 0.26 for model 1 and 2 and 0.25 for model 3, indicating that the models explained a moderate amount of the total variability.

In external validation, all models performed similarly although were less robust (AUC, 0.72 [95% CI, .70–.74]) (Figure 2). Calibration was similar for all models with slope of 0.74 (95% CI, .66–.81) and intercept of −0.26 (95% CI, −.35 to −.18) for model 1 and slope of 0.69 (95% CI, .63–.76) and intercept of −0.21 (95% CI, −.30 to −.12) for model 3 (Table 4). Calibration plots for each model are shown in Supplementary Figure 2. At a specificity level of 70%, model sensitivity was 82%, 83%, and 82% for model 1, 2, and 3, respectively. Sensitivity, specificity, positive predictive value, and negative predictive values for the 3 prediction models for viral-only diarrhea are shown in Table 5.

Sensitivity Analyses

In sensitivity analysis, different Ct values were used to attribute viral-only etiology (Supplementary Table 5). Model performance at varying thresholds was similar to the original cutoff of <30. Using Ct <25, model discrimination was similar with AUC of 0.82 (95% CI, .80–.84), 0.81 (95% CI, .80–.84), and

Table 2. Adjusted Association of Predictors of Viral-Only Diarrhea Etiology Among 3 Candidate Prediction Models in the National Surveillance (Derivation) Dataset, Bangladesh, 2014–2018

Characteristic	Model 1 (Full Model) OR (95% CI)	Model 2 (Forward Stepwise) OR (95% CI)	Model 3 (Simplified) OR (95% CI)
Age category			
<12 mo	Ref	Ref	Ref
12–23 mo	0.81 (.60–1.10)	0.81 (.60–1.09)	0.84 (.63–1.13)
24–59 mo	0.44 (.29–.65)	0.43 (.29–.64)	0.39 (.26–.57)
6–17 y	0.10 (.06–.18)	0.10 (.06–.18)	0.09 (.05–.15)
18–55 y	0.09 (.07–.12)	0.09 (.07–.12)	0.08 (.06–.11)
>55 y	0.07 (.04–.11)	0.07 (.04–.11)	0.06 (.04–.10)
Male sex	0.80 (.64–.98)	0.80 (.65–.98)	…
Duration of diarrhea, d	1.11 (1.04–1.20)	1.11 (1.04–1.19)	…
Abdominal pain	0.63 (.50–.79)	0.63 (.50–.79)	0.63 (.51–.79)
>10 episodes of diarrhea in 24 h	1.09 (.87–1.38)	…	…
Blood in stool	0.28 (.08–1.03)	0.28 (.08–1.02)	0.28 (.07–1.02)
Presence of vomiting	0.98 (.78–1.23)	…	…
Dehydration severity			
None	Ref	Ref	…
Some	1.18 (.95–1.48)	1.18 (.95–1.47)	…
Severe	2.32 (1.53–3.54)	2.33 (1.54–3.55)	…
Season			
Winter	Ref	Ref	…
Summer	1.00 (.75–1.33)	1.00 (.75–1.33)	…
Monsoon	0.59 (.47–.75)	0.59 (.47–.75)	…

Abbreviations: CI, confidence interval; OR, odds ratio; Ref, reference group.

Table 3. Model Characteristics Using the Area Under Receiver Operating Characteristic Curve and Pseudo R^2 for Each Model in the National Surveillance (Derivation) Dataset, Bangladesh, 2014–2018

Model	Included Predictors	AUC (95% CI)	Pseudo R^2
Model 1 (full model)	Age, sex, duration, >10 diarrhea episodes/24 h, abdominal pain, vomiting, bloody stool, dehydration severity, season	0.82 (.80–.84)	0.26
Model 2 (forward stepwise)	Age, sex, duration, abdominal pain, bloody stool, dehydration severity, season	0.82 (.80–.84)	0.26
Model 3 (simplified)	Age, abdominal pain, bloody stool	0.81 (.78–.82)	0.25

Abbreviations: AUC, area under the receiver operating characteristic curve; CI, confidence interval.

0.80 (95% CI, .79–.82) among model 1, 2, and 3, respectively (Supplementary Table 6). Using Ct <35, AUC was 0.80 (95% CI, .78–.82), 0.80 (95% CI, .78–.82), and 0.78 (95% CI, .76–.80) among model 1, 2, and 3, respectively (Supplementary Table 7). Using binary categorization for age (age <4 vs ≥4 years), the models performed nearly identically to the original models using the a priori clinically determined age categories with AUC 0.82 (95% CI, .80–.84) for model 1 and 2 and AUC 0.82 (95% CI, .78–.82) in model 3 (Supplementary Table 8).

DISCUSSION

In this study, we have derived and externally validated several clinical prediction models that can adequately discriminate viral-only diarrhea from other etiologies using relatively few, easily collected predictor variables that are commonly obtained during routine clinical care, using data from these 2 large surveillance systems in Bangladesh. Our findings provide an improved understanding of the key predictors of viral-only diarrhea etiology among patients of all ages and add insights regarding the prevalence of viral-only diarrhea etiologies in older age groups, which may allow for more judicious use of antibiotics.

Prior research has shown that up to 50%–80% of children <5 years of age inappropriately receive antibiotics for diarrhea, the majority of whom have viral etiologies [19, 41–44]. Additionally, a reanalysis of results from the Global Enteric Multicenter (GEMS) study estimated that there were approximately 12.6 inappropriate-treated diarrhea cases for each appropriately treated case, with viruses among the leading etiologies inappropriately treated with antibiotics [45]. While multiple factors contribute to high levels of antibiotic use, 1 major contributing factor is a lack of evidence-based tools to assist clinicians with assessing the risk of viral diarrhea etiology, which may subsequently influence antibiotic prescribing behaviors. Clinical prediction models such as those developed in this study may provide clinicians with evidence-based tools to better identify patients with viral-only diarrhea and increase their confidence in following guidelines when antibiotic use is not indicated. Recently, a randomized crossover trial of a clinical decision support tool for viral diarrhea prediction among pediatric patients in Bangladesh and Mali found that a 10% increase in predicted probability of viral-only diarrhea was associated with a 14% decrease in the odds of antibiotic prescribing [46].

Consistent with prior studies, this study shows that age is a major predictor of viral-only etiology diarrhea, with the predicted odds of viral-only etiology decreased with increasing patient age, particularly dropping off after early childhood [47]. In our sensitivity analysis using CART to discretize age, age was optimally split into a binary variable, differentiating young children (<4 years) versus older children and adults (≥4 years), and performed nearly as well as the original age categories, indicating that the age remains one of the most important epidemiological predictors of viral-only diarrhea. However, viral diarrhea still constitutes a substantial burden of diarrhea in adults, who are often presumed to have nonviral diarrhea by clinicians based on age alone. In our study, nearly 1 in 5 adults aged 18–55 years had viral-only pathogens detected on TAC PCR, indicating that viruses remain an important cause of diarrhea in older populations. This finding may be driven by the high incidence of viral pathogens in young children, as this age group often consists
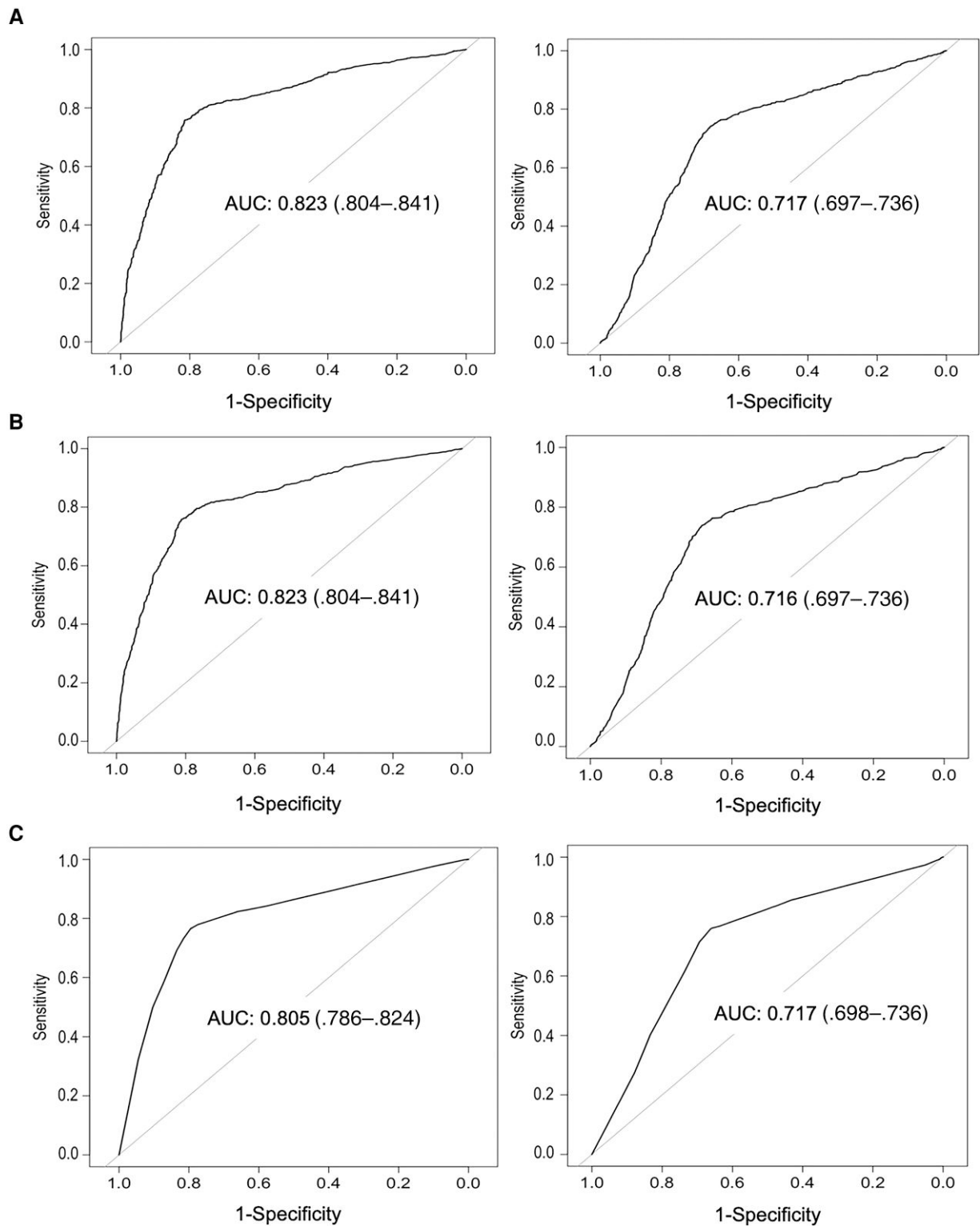
Figure 2. Area under the receiver operating characteristic curve (AUC) for model 1 (*A*), model 2 (*B*), and model 3 (*C*) in the derivation (left) and validation (right) datasets.

of parents and caregivers. Our findings suggest that clinicians should simultaneously maintain a high index of suspicion for viral diarrhea in infants and young children, while avoiding broad generalizations of presumed nonviral etiology among older children and adults. It is also important for clinicians to assess for additional clinical characteristics such as bloody stool and

Table 4. Model Performance Using the Area Under Receiver Operating Characteristic Curve, Calibration-in-the-Large (α), and Calibration Slope (β) for Each Model in the Validation Dataset, Bangladesh, 2014–2017

Model	AUC (95% CI)	α: Calibration-in-the-Large (95% CI)	β: Calibration Slope (95% CI)
Model 1 (full model)	0.72 (.70–.74)	−.26 (−.35 to −.18)	.73 (.66–.81)
Model 2 (forward stepwise selection)	0.72 (.70–.73)	−.27 (−.35 to −.18)	.73 (.66–.81)
Model 3 (simplified)	0.72 (.70–.74)	−.21 (−.30 to −.12)	.69 (.63–.76)

Abbreviations: AUC, area under the receiver operating characteristic curve; CI, confidence interval.

Table 5. Sensitivity, Specificity, Positive Predictive Value, and Negative Predictive Value for the 3 Prediction Models for Viral-Only Diarrhea

Specificity	Sensitivity	PPV	NPV	Probability Cutoff
Model 1				
0.6	0.84	0.50	0.89	0.15
0.7	0.82	0.56	0.89	0.17
0.8	0.77	0.64	0.88	0.33
0.9	0.53	0.71	0.80	0.61
Model 2				
0.6	0.85	0.50	0.89	0.15
0.7	0.83	0.56	0.89	0.17
0.8	0.77	0.64	0.88	0.33
0.9	0.53	0.71	0.81	0.61
Model 3				
0.6	0.85	0.50	0.89	0.15
0.7	0.82	0.53	0.89	0.17
0.8	0.74	0.65	0.87	0.41
0.9	0.50	0.71	0.79	0.67

Abbreviations: NPV, negative predictive value; PPV, positive predictive value.

abdominal pain in all age groups, as currently recommended in clinical guidelines, including from the WHO [4, 36]. Models such as those developed here may assist clinicians with making more accurate assessments of risk of viral versus nonviral diarrhea, based on more than a single clinical symptom.

Bloody stool and abdominal pain predicted lower odds of viral diarrhea, consistent with prior guidelines and studies indicating their association with invasive bacterial etiologies [4, 36, 48]. Our simplified prediction model (model 3) consisting of only age, abdominal pain, and bloody stool performed nearly as well as the alternative models. Model 2, which eliminated the variables of diarrhea severity and vomiting, performed nearly identically to model 1, which contained all candidate variables. We found that diarrhea severity and vomiting were not predictive of viral-only diarrhea, indicating that assessing for these symptoms is of limited benefit for risk prediction of viral etiology diarrhea. Given the models' similar performance, the simplified model may be most feasible for use in high-

volume and highly resource-constrained settings where a rapid assessment using only a few variables is desirable, while model 2 is optimal by adding a clinical assessment of dehydration and accounting for seasonal variations and could be used when resources allow.

Season was found to be predictive of viral-only etiology, with presentation during the monsoon versus winter season having a lower predicted odds of viral-only diarrhea. This finding is consistent with the known seasonal cycles of diarrhea in Bangladesh, with rotavirus peaks typically occurring in the winter and cholera peaks in the beginning and end of monsoon season [49–51]. Interestingly, severe dehydration (vs no dehydration) was found to be a positive predictor of viral-only etiology of diarrhea. This was unexpected, as some bacterial causes of diarrhea (eg, cholera) are characterized by high-volume stool output leading to severe dehydration very rapidly [4]. This may be related to those with viral diarrhea having a longer duration of symptoms and perhaps be reflective of greater delays in seeking care with viral diarrhea, which may be more insidious, in contrast to bacterial causes, which may cause more rapid-onset and alarming symptoms prompting care-seeking.

This study has several limitations. The data were collected exclusively at hospitals and therefore may be skewed toward patients with more severe illness; these study findings may therefore not be generalizable to patients presenting to outpatient clinics. Additionally, while the data were collected from diverse settings throughout Bangladesh, the prediction model may not be generalizable to other countries with different epidemiology of diarrhea, particularly locations that have initiated new rotavirus vaccination campaigns.

As evidenced by the modest pseudo R^2, other predictor variables associated with viral diarrhea etiology may have improved the performance of the model; however, this was not possible given the retrospective nature of this study. For example, while data on history of fever were collected in the derivation dataset, this variable was not collected in the validation dataset, which precluded its inclusion in the models. However, given the variable association of fever with both viral and bacterial pathogens (eg, Shigella and Salmonella often cause fever, whereas other bacteria such as V cholerae rarely do), the inclusion of this variable is of questionable utility [36]. Further research using a wider array of easily collected clinical variables may improve model performance across diverse settings.

There is currently no existing gold standard to attribute etiology using TAC, therefore a Ct value of <30 was set a priori based on expert consultation, although this may not be the ideal cutoff. However, our sensitivity analysis using varying Ct values found similar results to the primary analysis, showing that this threshold is likely reasonable when using similar methodology. Additionally, differences in the age, sex, and symptomatology of the populations that visit the hospitals in the 2 datasets indicate that the epidemiology of diarrhea may vary depending on

healthcare facility type and may have contributed to decreased performance of the models in external validation.

Last, while the aim of this study was to develop a prediction model for risk of viral-only etiology, conversely a prediction model for nonviral/bacterial causes was considered. However, we opted to predict viral-only etiologies since this information would be most clinically actionable (ie, clinicians could avoid prescribing antibiotics given consensus on the nonuse of antibiotics for viral diarrhea) whereas there is no uniform standard for treatment of bacterial causes, although most cases still do not warrant antibiotic use per WHO recommendations. However, further research examining the preferences of clinicians on the type of prediction model output for use in clinical decision-making is needed.

CONCLUSIONS

We have derived and retrospectively validated several clinical prediction models for viral-only diarrhea in Bangladesh that can be applied to patients of all ages. Further research to develop this model into a mobile application may help clinicians identify patients with viral diarrhea who do not warrant antibiotics, without the need for laboratory stool diagnostics, thereby supporting the reduction of inappropriate antibiotic use.

Supplementary Data

Supplementary materials are available at *Open Forum Infectious Diseases* online. Consisting of data provided by the authors to benefit the reader, the posted materials are not copyedited and are the sole responsibility of the authors, so questions or comments should be addressed to the corresponding author.

References

1. James SL, Abate D, Abate KH, et al. Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet **2018**; 392:1789–858.
2. Troeger C, Blacker BF, Khalil IA, et al. Estimates of the global, regional, and national morbidity, mortality, and aetiologies of diarrhoea in 195 countries: a systematic analysis for the Global Burden of Disease Study 2016. Lancet Infect Dis **2018**; 18:1211–28.
3. Shane AL, Mody RK, Crump JA, et al. Infectious Diseases Society of America clinical practice guidelines for the diagnosis and management of infectious diarrhea. Clin Infect Dis **2017**; 2017:e45–80.
4. World Health Organization (WHO). The treatment of diarrhoea: a manual for physicians and other senior health workers. Geneva, Switzerland: WHO; **2005**.
5. Okeke IN, Laxminarayan R, Bhutta ZA, et al. Antimicrobial resistance in developing countries. Part I: recent trends and current status. Lancet Infect Dis **2005**; 5: 481–93.
6. Laxminarayan R, Duse A, Wattal C, et al. Antibiotic resistance—the need for global solutions. Lancet Infect Dis **2013**; 13:1057–98.
7. Guarino A, Bruzzese E, Giannattasio A. Antibiotic treatment of acute gastroenteritis in children. F1000Res **2018**; 7:193.
8. World Health Organization (WHO). Global action plan on antimicrobial resistance. Geneva, Switzerland: WHO; **2017**.
9. Wong CS, Jelacic S, Habeeb RL, Watkins SL, Tarr PI. The risk of the hemolytic–uremic syndrome after antibiotic treatment of *Escherichia coli* O157: H7 infections. N Engl J Med **2000**; 342:1930–6.
10. Platts-Mills JA, Kosek M. Update on the burden of *Campylobacter* in developing countries. Curr Opin Infect Dis **2014**; 27:444–50.
11. Mandomando I, Espasa M, Vallès X, et al. Antimicrobial resistance of *Vibrio cholerae* O1 serotype Ogawa isolated in Manhica District Hospital, southern Mozambique. J Antimicrob Chemother **2007**; 60:662–4.
12. Mahbubur R, Shoma S, Rashid H, et al. Increasing spectrum in antimicrobial resistance of *Shigella* isolates in Bangladesh: resistance to azithromycin and ceftriaxone and decreased susceptibility to ciprofloxacin. J Health Popul Nutr **2007**; 25:158.
13. Tjaniadi P, Lesmana M, Subekti D, et al. Antimicrobial resistance of bacterial pathogens associated with diarrheal patients in Indonesia. Am J Trop Med Hyg **2003**; 68:666–70.
14. Kahsay AG, Muthupandian S. A review on sero diversity and antimicrobial resistance patterns of *Shigella* species in Africa, Asia and South America, 2001–2014. BMC Res Notes **2016**; 9:422.
15. Garbern SC, Chu T-C, Yang P, et al. Clinical and socio-environmental determinants of multidrug resistant *Vibrio cholerae* 01 in older children and adults in Bangladesh. Int J Infect Dis **2021**; 105:436–41.
16. Kumarasamy KK, Toleman MA, Walsh TR, et al. Emergence of a new antibiotic resistance mechanism in India, Pakistan, and the UK: a molecular, biological, and epidemiological study. Lancet Infect Dis **2010**; 10:597–602.
17. Okeke IN, Edelman R. Dissemination of antibiotic-resistant bacteria across geographic borders. Clin Infect Dis **2001**; 33:364–9.
18. World Health Organization (WHO). Antimicrobial resistance: global report on surveillance. Geneva, Switzerland: WHO; **2014**.
19. Bebell LM, Muiru AN. Antibiotic use and emerging resistance: how can resource-limited countries turn the tide? Glob Heart **2014**; 9:347–58.
20. Hatchette TF, Farina D. Infectious diarrhea: when to test and when to treat. CMAJ **2011**; 183:339–44.
21. Pavlinac PB, Denno DM, John-Stewart GC, et al. Failure of syndrome-based diarrhea management guidelines to detect *Shigella* infections in Kenyan children. J Pediatric Infect Dis Soc **2016**; 5:366–74.
22. Bray AE, Ahmed S, Das SK, et al. Viral pathogen-specific clinical and demographic characteristics of children with moderate-to-severe diarrhea in rural Bangladesh. Am J Trop Med Hyg **2019**; 101:304–9.
23. Dey SK, Sharif N, Sarkar OS, et al. Molecular epidemiology and surveillance of circulating rotavirus among children with gastroenteritis in Bangladesh during 2014–2019. PLoS One **2020**; 15:e0242813.
24. Kotloff KL. The burden and etiology of diarrheal illness in developing countries. Pediatr Clin North Am **2017**; 64:799–814.
25. Fischer Walker CL, Sack D, Black RE. Etiology of diarrhea in older children, adolescents and adults: a systematic review. PLoS Negl Trop Dis **2010**; 4:e768.
26. Kotloff KL, Nataro JP, Blackwelder WC, et al. Burden and aetiology of diarrhoeal disease in infants and young children in developing countries (the Global Enteric Multicenter Study, GEMS): a prospective, case-control study. Lancet **2013**; 382: 209–22.
27. Platts-Mills JA, Babji S, Bodhidatta L, et al. Pathogen-specific burdens of community diarrhoea in developing countries: a multisite birth cohort study (MAL-ED). Lancet Glob Health **2015**; 3:e564–75.
28. Ballard S-B, Reaves EJ, Luna CG, et al. Epidemiology and genetic characterization of noroviruses among adults in an endemic setting, Peruvian Amazon Basin, 2004–2011. PLoS One **2015**; 10:e0131646.
29. Nataraju SM, Pativada M, Chatterjee D, et al. Molecular epidemiology of norovirus infections in children and adults: sequence analysis of region C indicates genetic diversity of NVGII strains in Kolkata, India. Epidemiol Infect **2011**; 139: 910–8.

30. Taniuchi M, Islam K, Sayeed MA, et al. Etiology of diarrhea requiring hospitalization in Bangladesh by quantitative PCR, 2014–2018. Clin Infect Dis 2021; 73: e2493–9.

31. World Health Organization (WHO). IMAI district clinician manual: hospital care adolescents and adults: guidelines for the management of illnesses with limited-resources. Geneva, Switzerland: WHO; 2012.

32. Chao DL, Roose A, Roh M, Kotloff KL, Proctor JL. The seasonality of diarrheal pathogens: a retrospective study of seven sites over three years. PLoS Negl Trop Dis 2019; 13:e0007211.

33. Das SK, Klontz EH, Azmi IJ, et al. Characteristics of multidrug resistant *Shigella* and *Vibrio cholerae* O1 infections in patients treated at an urban and a rural hospital in Bangladesh. ISRN Microbiol 2013; 2013:213915.

34. Taniuchi M, Islam K, Sayeed MA, et al. Etiology of diarrhea requiring hospitalization in Bangladesh by quantitative polymerase chain reaction, 2014–2018. Clin Infect Dis 2021; 73:e2493–9.

35. Farthing M, Salam MA, Lindberg G, et al. Acute diarrhea in adults and children: a global perspective. J Clin Gastroenterol 2013; 47:12–20.

36. Kotloff KL. Bacterial diarrhoea. Curr Opin Pediatr 2022; 34:147–55.

37. Steyerberg EW, Vergouwe Y. Towards better clinical prediction models: seven steps for development and an ABCD for validation. Eur Heart J 2014; 35:1925–31.

38. Van Calster B, Nieboer D, Vergouwe Y, De Cock B, Pencina MJ, Steyerberg EW. A calibration hierarchy for risk models was defined: from utopia to empirical data. J Clin Epidemiol 2016; 74:167–76.

39. Youden WJ. Index for rating diagnostic tests. Cancer 1950; 3:32–5.

40. Le CT. A solution for the most basic optimization problem associated with an ROC curve. Stat Methods Med Res 2006; 15:571–84.

41. Ahmed F, Farheen A, Ali I, Thakur M, Muzaffar A, Samina M. Management of diarrhea in under-fives at home and health facilities in Kashmir. Int J Health Sci (Qassim) 2009; 3:171–5.

42. De T, Kondekar S, Rathi S. Hospital based prospective observational study to audit the prescription practices and outcomes of paediatric patients (6 months to 5 years age group) presenting with acute diarrhea. J Clin Diagn Res 2016; 10: SC01–5.

43. Fink G, D'Acremont V, Leslie HH, Cohen J. Antibiotic exposure among children younger than 5 years in low-income and middle-income countries: a cross-sectional study of nationally representative facility-based and household-based surveys. Lancet Infect Dis 2020; 20:179–87.

44. Tsige AG, Nedi T, Bacha T. Assessment of the management of diarrhoea among children under five in Addis Ababa, Ethiopia. Pediatr Health Med Ther 2020; 11: 135–43.

45. Lewnard JA, Rogawski McQuade ET, Platts-Mills JA, Kotloff KL, Laxminarayan R. Incidence and etiology of clinically-attended, antibiotic-treated diarrhea among children under five years of age in low- and middle-income countries: evidence from the Global Enteric Multicenter Study. PLoS Negl Trop Dis 2020; 14: e0008520.

46. Nelson EJ, Khan AI, Keita AM, et al. Improving antibiotic stewardship for diarrheal disease with probability-based electronic clinical decision support: a randomized crossover trial. JAMA Pediatr 2022; 176:973–9.

47. Brintz BJ, Howard JI, Haaland B, et al. Clinical predictors for etiology of acute diarrhea in children in resource-limited settings. PLoS Negl Trop Dis 2020; 14: e0008677.

48. Barr W, Smith A. Acute diarrhea in adults. Am Fam Physician 2014; 89:180–9.

49. Khan AI, Rashid MM, Islam MT, et al. Epidemiology of cholera in Bangladesh: findings from nationwide hospital-based surveillance, 2014–2018. Clin Infect Dis 2020; 71:1635–42.

50. Hashizume M, Armstrong B, Wagatsuma Y, Faruque ASG, Hayashi T, Sack DA. Rotavirus infections and climate variability in Dhaka, Bangladesh: a time-series analysis. Epidemiol Infect 2008; 136:1281–9.

51. Hashizume M, Faruque ASG, Wagatsuma Y, Hayashi T, Armstrong B. Cholera in Bangladesh: climatic components of seasonal variation. Epidemiology 2010; 21: 706–10.