DOI: 10.7860/JCDR/2026/81409.23737

Research Protocol

Microbiology Section

Genotypic Diversity of Norovirus Variants in Acute Gastroenteritis Cases and its Correlation with the Severity of Diarrhoea: A Research Protocol

NANDKISHOR J BANKAR[1], GARGI MUDEY[2], ANKIT K BADGE[3], MEENA MISHRA[4]

ABSTRACT

Introduction: Norovirus (NoV) is a leading cause of acute gastroenteritis globally, especially in children under five. Despite its burden, data from India on circulating genotypes and their clinical implications remain limited.

Need of the study: This study is needed for the early diagnosis and treatment of NoV-related diarrhoea also it will help epidemiologists to enable the detection of emerging and prevalent strains of NoV for better disease control, thereby reducing morbidity.

Aim: To determine the genotypic diversity of NoV variants in patients with acute gastroenteritis and to assess the correlation between specific genotypes and the clinical severity of diarrhoea.

Materials and Methods: This cross-sectional study will be conducted in December 2024 – November 2026 at the Department of Microbiology, Datta Meghe Medical College, Nagpur, Maharashtra, India enrolling 766 participants with a history of acute gastroenteritis. Stool or rectal swab samples will be collected and analysed using Reverse Transcriptase Polymerase Chain Reaction (RT-PCR) for NoV detection, followed by genotyping and bioinformatics-based sequence analysis to determine genetic diversity and variant distribution. Demographic parameters, including age, sex, and residential background (urban/rural), will be recorded for all participants. Additionally, machine learning, specifically graph-based techniques and graph neural networks, will be applied for advanced genomic feature extraction and analysis. Descriptive statistics, Chi-square or Fisher's exact test, multiple regression, and Spearman correlation will be used to analyse clinical and molecular data, with a p-value <0.05 considered statistically significant. Temporal and spatial trends will also be assessed.

Keywords: Genetic variation, Norovirus infections, Virus genotyping

INTRODUCTION

The NoV is widespread and is reported to be the main cause of sporadic and epidemic gastroenteritis in all age groups globally [1]. NOV is the most common cause of diarrhoea in all ages, resulting in 699 million cases and 219,000 deaths annually [2]. The NOV are only next to rotavirus in infecting children <5 years of age, resulting in 200 million cases and 54,214 deaths [3].

The NOV are non enveloped and contain a single-stranded positive-sense RNA genome belonging to the *Caliciviridae* family, with the name calicivirus derived from the characteristic appearance of the viral particle with 'cuplike' indentations [4]. The virus has three long Open Reading Frames (ORF), ORF-1 encodes Non-Structural Proteins (NSP1-NSP7), such as RNA polymerase, helicase, and protease vital in replication, while capsid (VP1) is encoded by ORF 2 and minor structural components (VP2) by ORF3 [5,6]. NOV are divided into seven genotypes (GI to GVII) based on the amino acid of the capsid protein. GI and GII primarily infect humans and are further divided into nine capsid genotypes in GI and 22 in GII. Other genotypes were found in porcine, canine, feline, bovine, murine, etc., [6,7]. The dominant strain is replaced every 2-3 years because of antigenic mutations, which requires tracking the prevalent strains in the respective geographical regions. NOVs usually present with vomiting, diarrhoea, abdominal cramps, and low-grade fever. NOV is often diagnosed using highly sensitive and specific molecular assays, as immunological assays, such as Enzyme-Linked Immunosorbent Assay (ELISA), while immunochromatography have a poor sensitivity of around 50% [7]. Symptomatic treatments are usually administered because there is no antiviral drug to treat NOV. Several NOV vaccines are under trial; however, no licensed vaccines are currently available [8]. In developing countries such as India, continuous assessment of NOVs is necessary as NOV is the most prominent viral agent causing diarrhoea disease contributing to the global burden of diseases [9]. Limited data is available on the incidence and circulating genotypes. Also, during their post-rotavirus vaccination era, NOV potentially emerges as the most frequent cause of diarrhoeal disease [10].

Therefore, the problem addressed by this research proposal is the limited knowledge of the molecular epidemiology of NOV variants. The absence of robust surveillance and genotyping systems impedes our ability to determine the prevalent NOV genotypes, understand their genetic characteristics, and evaluate their impact on clinical outcomes.

REVIEW OF THE LITERATURE

The NOV is one of the major causes of non bacterial gastroenteritis. According to Ahmed SM et al., the global prevalence of NOV in cases of gastroenteritis is 18%, and there is a clear gap remaining in developing countries (14%) [11]. The results are, therefore, incompatible with previous reports, which indicate that the prevalence of NOV in gastroenteritis in developing countries should be higher than in developed countries because of the limited availability of safe water, sanitation and hygiene. The lack of surveillance of NOV and the tendency not to opt for medical services for mild diseases may underestimate the prevalence in developing countries. Authors found a significantly higher estimate of the prevalence of NoV in Africa (11%) and Latin America (15%), suggesting that the prevalence of other areas of the study (such as Asia and the Middle East) may be higher.

According to Kobayashi S et al., the serological prevalence of IgG antibodies in seven noV-types (GI.1, GI.4, GII.3, GII.4, GII.10, GII.12, and GII.15) between the ages of 1 to 62 in Japan was investigated [12]. The seroprevalence related to age was measured by ELISA using recombinant VLP antibodies. The seropositive rate of all seven VLP antigens increased gradually with age. According to Menon VK et al., age-stratified sera from populations in India and the United Kingdom were analysed for the presence of specific IgGs from NOV and Group II by time-resolved immunofluorescence experiments and compared to the relative levels of antibodies in the two populations [13]. Antibody levels were higher among all ages in India than in the United Kingdom and increased with age in India, while in the United Kingdom, antibody levels decreased with age. The results show different patterns of exposure to NOVs in both countries.

According to Kulkarni R et al., the seroprevalence of GII.4 NoV antibodies were reported among children aged ≤5 years in Pune, India [14]. Of the 191 serum samples, 98 (51.3%) were positive. The antibody blocking Histoblood Group Antigen (HBGA) was detected in 33 of the 54 positive samples tested. The prevalence and number of IgG and blocking antibodies vary depending on age, and the lowest prevalence was in children aged six to 23 months. Children with antibodies showed a significantly lower faecal NOV RNA detection rate than children without antibodies, suggesting a previous NOV exposure.

According to Nagamani K et al., variants of GII.4 have been reported as the main cause of the global gastroenteritis pandemic [15]. GII is the most common genotype, accounting for more than 70% of NOV infections. By characterising the predominant genotypes and analysing their correlation with disease severity, the study will contribute to finding the specific NoV genotypes to the clinical spectrum of gastroenteritis and to identifying potential molecular determinants linked to enhanced virulence or pathogenicity.

This study aims to find out the circulation patterns, genetic diversity, and clinical significance of NoV strains among patients with acute gastroenteritis.

Primary Objective

1. To identify the prevalent NoV genotypes in the study population.
2. To identify demographic patterns in the occurrence of NoV infections among the study population.

Secondary Objective

1. To correlate which genotype is related to the severity of the disease.
2. To assess the clinical implications of different NoV variants to inform evidence-based treatment strategies.

MATERIALS AND METHODS

This cross-sectional study will be conducted in the Department of Microbiology, Datta Meghe Medical College, Wanadongri, Nagpur, over a period of two years (December 2024 – November 2026). All participants will be enrolled after obtaining informed consent. Ethical approval has been obtained from the Institutional Ethics Committee with clearance number IEC/DMMC/2024/12-04.

Sample size calculation:

$$n = \frac{z^2 p\,(1-p)}{e^2}$$

$$n = \frac{(1.96)^2 \times 0.50 \times (1-0.50)}{(0.05)^2}$$

Where,

'n' is sample size

'z' is selected critical value of desired confidence level (CL 95%)

'p' is estimated proportion of an attribute that is present in population. (50%) as no study was carried out in the region.

'e' level of precision (5%)

A 385, when considering two variables, for examining relationship (e.g., correlation or interaction effects), the sample size needs to be increased to maintain the same level of confidence and precision. This means 766 or more samples are needed to have a confidence level of 95% that the real value is within ±5% of the samples.

Inclusion criteria: Participants with a history of acute gastroenteritis, cases with laboratory-confirmed NoV infections (both symptomatic and asymptomatic), and individuals from whom appropriate biological samples, such as stool or vomitus have been collected for diagnostic purposes will be included in the study.

Exclusion criteria: Individuals who are unwilling to participate, cases of gastrointestinal infections due to pathogens other than NoV, instances with inadequate sample collection or insufficient data for genotyping and variant analysis, and situations where the required samples for molecular testing are inaccessible will be excluded from the study.

Participants will be selected who attend the tertiary care hospital with a history of acute gastroenteritis. Stool or rectal swab samples will be collected from the selected participants using standard procedures. Sample collection will be performed by trained healthcare professionals, ensuring proper handling and storage to maintain sample integrity.

Outcomes: All the samples suspected of NoV will be analysed by RT-PCR (DiAGSure™ NoV TaqMan PCR kit). This involves the extraction of RNA, reverse transcription and PCR amplification of genomic regions. The genome will be detected using multiplex RT-PCR and further characterised by Nanopore sequencing for genomic analysis. The genotyping results identify the predominant genotypes and variants of NOV.

Sequence analysis: For all positive samples, the genomic region sequence will be performed by the Sanger sequencing method. The obtained sequences will be analysed in bioinformatics, including sequence alignment, phylogenetic analysis and comparison with reference sequences from the global NOV database. Specifically, the NCBI GenBank database will be used for sequence comparison, and the NoV reference sequence NC_001959.2 will serve as the standard for alignment and phylogenetic inference [16]. This analysis provides information on the genetic diversity and relation of NOV species.

Analysis using Machine Learning

To achieve feature extraction from genomic data, a graph-based technique is used. This technique is mainly suitable for complex tasks such as finding relationships between genomic elements. The following steps are involved in this method:

i. **Graph construction:** This is the first step, which includes identifying nodes that will represent individual nucleotide sequences and representing them in an appropriate form. Edges connect these nodes based on characteristics like proximity, sequence similarity or functional interactions.

ii. **Feature extraction:** Features are extracted from nodes and edges. These features include various attributes related to nodes or edges and are created using various machine learning techniques. Edge features are extracted by calculating sequence similarity, genomic distance, or functional associations.

iii. **Graph topology:** The overall structure of the graph is known as graph topology. It helps in identifying important features, degree centrality, and betweenness centrality.

iv. **Graph neural network:** It is a special type of neural network designed for graphs. It can be directly applied to genomic data represented in terms of a graph and can help in doing meaningful

analysis, such as clustering, classification, etc. Using the rich information stored in graph-based representations of genomic data, researchers may extract features that reflect intricate interactions and dependencies within the genome, allowing for more accurate and thorough study of genomic sequences [17].

[Table/Fig-1] shows Gantt chart of the study.

Activity	Dec 2024 - May 2025	June 2025 - Nov 2025	Dec 2025 - Mar 2026	Apr 2026 - July 2026	Aug 2026 - Nov 2026
Sample collection	▨	▨	▨		
Laboratory Analysis (RT-PCR and Genotyping)		▨	▨	▨	
Sequence analysis and bioinformatics		▨	▨	▨	
Data compilation, Interpretation				▨	▨
Report writing				▨	▨

[Table/Fig-1]: Gantt chart.

STATISTICAL ANALYSIS

All statistical analyses will be conducted using IBM Statistical Package for the Social Sciences (SPSS) Statistics software, version 26.0 (IBM Corp., Armonk, NY, USA). Descriptive statistics will be used to summarise the population data, clinical characteristics, NOV genotypes and variants in study participants. Statistical tests, such as the Chi-square test or Fisher's exact test, will be used to examine the relationship between the genotype/variation of NOV and the clinical results. Multiple regression and Spearman correlation for non parametric observations. Temporal and spatial analysis will be done using time sequence analysis and spatial clustering techniques, to explore patterns and trends in the infection of NOVs.

REFERENCES

[1] Pires SM, Fischer-Walker CL, Lanata CF, Devleesschauwer B, Hall AJ, Kirk MD, et al. Aetiology-Specific estimates of the global and regional incidence and mortality of diarrhoeal diseases commonly transmitted through food. PLoS ONE. 2015;10(12):e0142927.
[2] Bartsch SM, Lopman BA, Ozawa S, Hall AJ, Lee BY. Global economic burden of norovirus gastroenteritis. PLoS ONE. 2016;11(4):e0151219.
[3] Yang L, Shi S, Na C, Li B, Zhao Z, Yang T, et al. Rotavirus and norovirus infections in children under 5 years old with acute gastroenteritis in Southwestern China, 2018–2020. J Epidemiol Glob Health. 2022;12(3):292-303.
[4] Infectious Enteritis and Proctocolitis - ClinicalKey [Internet]. [cited 2023 Jan 24]. Available from: https://www.clinicalkey.com/#!/content/book/3-s2.0-B9780323609623001107.
[5] Yu Y, Yan S, Li B, Pan Y, Wang Y. Genetic diversity and distribution of human norovirus in China (1999-2011). BioMed Res Int. 2014;2014:e196169.
[6] Vinjé J. Advances in laboratory methods for detection and typing of norovirus. J Clin Microbiol. 2015;53(2):373-81.
[7] Barclay L, Cannon JL, Wikswo ME, Phillips AR, Browne H, Montmayeur AM, et al. Emerging Novel GII.P16 noroviruses associated with multiple capsid genotypes. Viruses. 2019;11(6):535.
[8] Netzler NE, Enosi Tuipulotu D, White PA. Norovirus antivirals: Where are we now? Med Res Rev. 2019;39(3):860-86.
[9] Ugboko HU, Nwinyi OC, Oranusi SU, Oyewale JO. Childhood diarrhoeal diseases in developing countries. Heliyon. 2020;6(4):e03690.
[10] Amin AB, Cates JE, Liu Z, Wu J, Ali I, Rodriguez A, et al. Rotavirus genotypes in the postvaccine era: A systematic review and meta-analysis of global, regional, and temporal trends by rotavirus vaccine introduction. J Infect Dis. 2024;229(5):1460-69.
[11] Ahmed SM, Hall AJ, Robinson AE, Verhoef L, Premkumar P, Parashar UD, et al. Global prevalence of norovirus in cases of gastroenteritis: A systematic review and meta-analysis. Lancet Infect Dis. 2014;14(8):725-30.
[12] Kobayashi S, Fujiwara N, Takeda N, Minagawa H. Seroepidemiological study of norovirus infection in Aichi Prefecture, Japan. Microbiol Immunol. 2009;53(6):356-59.
[13] Menon VK, George S, Aladin F, Nawaz S, Sarkar R, Lopman B, et al. Comparison of age-stratified seroprevalence of antibodies against norovirus GII in India and the United Kingdom. PLoS One. 2013;8(2):e56239.
[14] Kulkarni R, Lole K, Chitambar SD. Seroprevalence of antibodies against GII.4 norovirus among children in Pune, India. Journal of Medical Virology. 2016;88(9):1636-40.
[15] Nagamani K, Rani M, Reddy V, Rao P, Rajyalakshmi S, Pakalapaty S. Molecular epidemiology of norovirus variants detected in children under five years of age in Hyderabad, India. Indian J Med Microbiol. 2022;40(1):12-17.
[16] ORF1 nonstructural polyprotein [Norovirus GI] - Gene - NCBI [Internet]. [cited 2025 Dec 16]. Available from: https://www.ncbi.nlm.nih.gov/gene?Cmd=DetailsSearch&Db=gene&Term=1491970.
[17] Zhang XM, Liang L, Liu L, Tang MJ. Graph neural networks and their current applications in bioinformatics. Front Genet. 2021;12:690049.

PARTICULARS OF CONTRIBUTORS:
1. Professor and Head, Department of Microbiology, Datta Meghe Medical College, Datta Meghe Institute of Higher Education and Research (DU), Nagpur, Maharashtra, India.
2. Professor and Head, Department of Microbiology, Jawaharlal Nehru Medical College, Datta Meghe Institute of Higher Education and Research (DU), Sawangi (M), Wardha, Maharashtra, India.
3. Research Assistant, NKP Salve Institute of Medical Sciences, Nagpur, Maharashtra, India.
4. Professor and Head, Department of Microbiology, All India Institute of Medical Sciences, Nagpur, Maharashtra, India.

NAME, ADDRESS, E-MAIL ID OF THE CORRESPONDING AUTHOR:
Dr. Nandkishor J Bankar,
Professor and Head Department of Microbiology, Datta Meghe Medical College, Datta Meghe Institute of Higher Education and Research (DU), Nagpur, Maharashtra, India.
E-mail: drbankarnj28@gmail.com

PLAGIARISM CHECKING METHODS: [Jain H et al.]
- Plagiarism X-checker: Jul 10, 2025
- Manual Googling: Feb 16, 2026
- iThenticate Software: Feb 18, 2026 (1%)

ETYMOLOGY: Author Origin

EMENDATIONS: 7

AUTHOR DECLARATION:
- Financial or Other Competing Interests: None
- Was Ethics Committee Approval obtained for this study? Yes
- Was informed consent obtained from the subjects involved in the study? Yes
- For any images presented appropriate consent has been obtained from the subjects. NA

Date of Submission: **Jun 19, 2025**
Date of Peer Review: **Aug 28, 2025**
Date of Acceptance: **Feb 20, 2026**
Date of Publishing: **Jul 01, 2026**