

Н. Н. Сбитнева – кандидат биологических наук  
Б. С. Имангалиев – младший научный сотрудник  
Р. О. Быков – стажер-исследователь  
А. Г. Сергеев – доктор медицинских наук, профессор

#### **Information about the authors**

V. I. Chalapa – Research Assistant  
T. M. Itani – Candidate of Sciences (Biology)  
N.N. Sbitneva – Candidate of Sciences (Biology)  
B. S. Imangaliyev – Research Assistant  
R. O. Bykov – Research Assistant  
A. G. Sergeev – Doctor of Sciences (Medicine), Professor

УДК: 614.4

### **ПРОГНОЗИРОВАНИЕ ДИНАМИКИ ЗАБОЛЕВАЕМОСТИ НОРОВИРУСНОЙ ИНФЕКЦИЕЙ С ПРИМЕНЕНИЕМ МОДЕЛЕЙ ВРЕМЕННЫХ РЯДОВ НА ПРИМЕРЕ СВЕРДЛОВСКОЙ ОБЛАСТИ**

Владислав Игоревич Чалапа<sup>1</sup>, Анна Александровна Косова<sup>2</sup>, Александр  
Владимирович Семенов<sup>3</sup>

<sup>1,2</sup>ФГБОУ ВО «Уральский государственный медицинский университет»  
Минздрава России, Екатеринбург, Россия

<sup>1-3</sup>Екатеринбургский научно-исследовательский институт вирусных инфекций  
Федерального бюджетного учреждения науки «Государственный научный  
центр вирусологии и биотехнологии «Вектор» Федеральной службы по надзору  
в сфере защиты прав потребителей и благополучия человека, Екатеринбург,  
Россия

<sup>1</sup>chalapa\_vi@eniivi.ru

#### **Аннотация**

**Введение.** Норовирусная инфекция (НВИ) является широко распространенным заболеванием, особенно среди детей дошкольного возраста, специфические средства профилактики НВИ отсутствуют. **Цель исследования** – оценить возможность создания статистической модели для прогнозирования заболеваемости НВИ на примере Свердловской области. **Материалы и методы.** Материалом для исследования послужили данные государственной статистики заболеваемости НВИ в Свердловской области за 2009-2020 гг. (помесечно). Для прогнозирования заболеваемости были применены интегрированные модели авторегрессии – скользящего среднего с сезонностью (SARIMA) и без таковой (ARIMA). **Результаты.** Порядок авторегрессии был определен по результатам построения частичной автокорреляционной функции. Значения остальных параметров модели подбирались вручную путем сравнения результатов прогнозирования. Оптимальный результат продемонстрировала модель SARIMA (2, 0, 0) (0, 0, 1), ее средняя точность составила 93%. Признаков переобучения модели не обнаружено. **Обсуждение.** Некоторая неточность результатов прогнозирования может быть объяснена тем, что ранние наблюдения не были включены в модель из-за большой вариабельности

и отсутствия сезонности. В перспективе полученная модель может быть улучшена по мере накопления актуальных статистических данных и дополнена дополнительными предикторами (климатические параметры). **Выводы.** Прогнозирование заболеваемости НВИ с применением модели временных рядов (SARIMA) является возможным при условии хорошо отлаженной лабораторной диагностики и регистрации случаев. Перспективным направлением развития прогнозирования может стать использование климатических факторов в качестве дополнительных предикторов.

**Ключевые слова:** норовирусная инфекция, прогнозирование, временные ряды, Свердловская область.

## FORECASTING THE INCIDENCE OF NOROVIRUS INFECTION IN SVERDLOVSK REGION BASED ON A TIME-SERIES MODELS

Vladislav I. Chalapa<sup>1</sup>, Anna A. Kosova<sup>2</sup>, Alexandr V. Semenov<sup>3</sup>

<sup>1,2</sup>Ural state medical university, Yekaterinburg, Russia

<sup>1-3</sup>Yekaterinburg Research Institute of Viral Infections, Federal Budgetary Research Institution «State Research Center of Virology and Biotechnology «Vector», Federal Service for Surveillance on Consumer Rights Protection and Human Well-being, Yekaterinburg, Russia

<sup>1</sup>chalapa\_vi@eniivi.ru

### Abstract

**Introduction.** Norovirus infections are a leading cause of gastro-enteritis, especially in children, and currently no licensed vaccine is available. **The aim of the study** – to evaluate the applicability of a time-series forecasting to monitor norovirus infections in Sverdlovsk region. **Materials and methods.** Official statistical data on norovirus infection in the Sverdlovsk region were obtained monthly, and used in this study. ARIMA and SARIMA time-series models were used to forecast monthly incidence of norovirus infection. **Results.** Autoregression order was estimated using partial autocorrelation function. The best model was SARIMA (2, 0, 0) (0, 0, 1) with mean accuracy of 93%. Overfitting was undetermined. **Discussion.** Minor imprecision was estimated due to the exclusion of partial dataset with underreporting. In the future, obtained models may be improved via the extension of datasets and including climatic variables. **Conclusions.** Forecasting the incidence of norovirus infection using a time-series model is possible in case of less under-reporting. Including climatic variables may improve the model for forecasting norovirus infection.

**Keywords:** norovirus infection, forecasting, time-series, Sverdlovsk region.

### ВВЕДЕНИЕ

Норовирусная инфекция (НВИ) является широко распространенным заболеванием, особенно среди детей дошкольного возраста [1], меры специфической профилактики которого на данный момент не разработаны [2]. В связи с этим прогнозирование заболеваемости НВИ может способствовать эффективному планированию противоэпидемических мероприятий и рациональному распределению медицинских ресурсов.

**Цель исследования** – оценить возможность создания статистической модели для прогнозирования заболеваемости НВИ на примере Свердловской области.

## МАТЕРИАЛЫ И МЕТОДЫ

Материалом для исследования послужили данные государственной статистики заболеваемости НВИ в Свердловской области за 2009-2020 гг. (помесечно). Для прогнозирования заболеваемости были применены интегрированные модели авторегрессии – скользящего среднего с сезонностью (SARIMA) и без таковой (ARIMA). Сравнительная оценка полученных моделей проводилась с вычислением средней абсолютной ошибки в процентах (MAPE) и информационного критерия Акаике (AIC).

Статистическая обработка результатов проводилась в среде R (R Core Team, 2020). Для тестирования статистических гипотез применялись двусторонние критерии. Нулевые гипотезы отвергались при достигнутом уровне статистической значимости соответствующего критерия  $p < 0,05$ .

## РЕЗУЛЬТАТЫ

Для выбора оптимальной модели был проведен анализ исходного временного ряда (число случаев НВИ в Свердловской области по месяцам). Было обнаружено, что на этапе становления диагностики НВИ заболеваемость была крайне низкой и характеризовалась практическим отсутствием сезонности, которая приобрела выраженный зимний характер с 2015 г. Кроме того, в 2020 г. произошло резкое снижение заболеваемости НВИ, по-видимому, обусловленное введением ограничительных мероприятий в отношении COVID-19, и многолетняя тенденция динамики заболеваемости радикально изменилась (рисунок 1). По этой причине для построения модели были использованы данные за 2015-2018 гг., тестирование модели производилось на данных 2019 г.

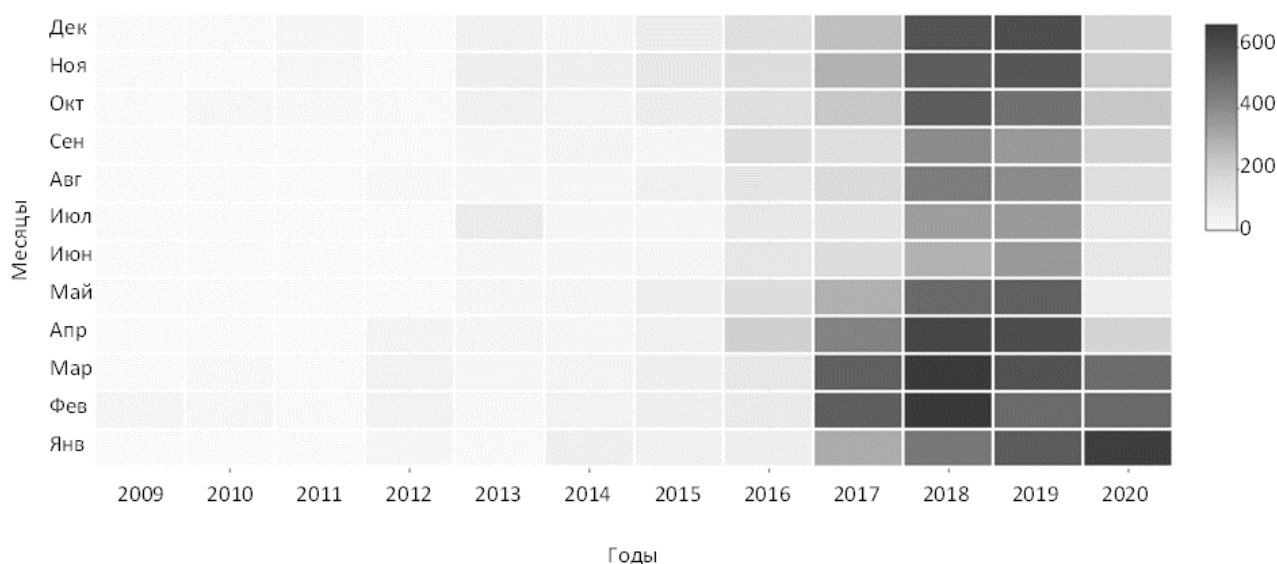


Рис. 1. Помесячная динамика числа случаев норовирусной инфекции (абс.) в Свердловской области, 2009-2020 гг.

Порядок авторегрессии был определен по результатам построения частичной автокорреляционной функции, статистически значимым оказался лаг, равный единице. Значения остальных параметров модели подбирались

вручную путем сравнения результатов прогнозирования по величине средней абсолютной ошибки в процентах и критерия Акаике. Оптимальный результат продемонстрировала модель SARIMA (2, 0, 0) (0, 0, 1). Средняя точность модели составила 93% с диапазоном варьирования от 0% (полное совпадение для октября 2019 г.) до 74% (для февраля 2019 г., когда наблюдалось необычное снижение числа случаев, рисунок 2). В сравнении с моделью ARIMA, у SARIMA наблюдалось меньшее значение средней абсолютной ошибки в процентах (MAPE 35,4 и 13,9 соответственно). Признаков переобучения (overfitting) модели SARIMA не наблюдалось, средняя абсолютная ошибка в процентах для тестового временного ряда оказалась даже ниже, чем для обучающего ряда.

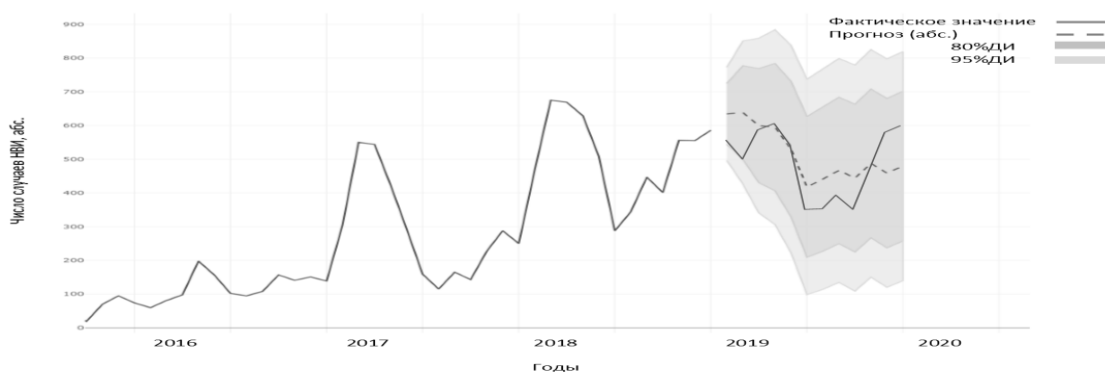


Рис. 2. Прогноз заболеваемости норовирусной инфекцией (НВИ) в Свердловской области на 2019 г. (абс.)

## ОБСУЖДЕНИЕ

Поскольку заболеваемость НВИ в средних широтах имеет выраженную зимнюю сезонность, преимущество модели SARIMA в целом было ожидаемо. Вместе с тем, риск переобучения модели SARIMA, обусловленный ее большей сложностью, оправдывал первоначальный выбор модели ARIMA в качестве альтернативы. Некоторая неточность результатов прогнозирования может быть объяснена тем, что ранние наблюдения не были включены в модель из-за большой вариабельности и отсутствия сезонности. В перспективе полученная модель может быть улучшена по мере накопления актуальных статистических данных. Кроме того, учитывая, что сезонность НВИ по крайней мере частично обусловлена влиянием климатических факторов [3], последние могут быть использованы в качестве предикторов в более сложных моделях.

## **ВЫВОДЫ**

Прогнозирование заболеваемости НВИ с применением моделей временных рядов является возможным при условии хорошо отлаженной лабораторной диагностики и регистрации случаев. В связи с выраженной сезонностью заболевания, модели, учитывающие этот фактор (в частности, SARIMA), демонстрируют лучшие результаты. Перспективным направлением развития прогнозирования заболеваемости НВИ может стать применение моделей, учитывающих в качестве предикторов климатические факторы.

## **СПИСОК ИСТОЧНИКОВ**

1. Global prevalence and genotype distribution of norovirus infection in children with gastroenteritis: A meta-analysis on 6 years of research from 2015 to 2020 / Farahmand M., Moghoofei M., Dorost A. et al. // *Reviews in Medical Virology*. – 2021; 32(1): e2237.
2. Insights and Enhanced Human Norovirus Cultivation in Human Intestinal Enteroids / Ettayebi K., Tenge V., Cortes-Penfield N. et al. // *mSphere*. – 2021;6(1): 1136-20.
3. Greer A.L., Drews S.J., Fisman D.N. Why “winter” vomiting disease? Seasonality, hydrology, and Norovirus epidemiology in Toronto, Canada. *EcoHealth*. – 2009; 6(2): 192-199.

## **Сведения об авторах**

В. И. Чалапа – аспирант

А. А. Косова – кандидат медицинских наук, доцент

А. В. Семенов – доктор биологических наук

## **Information about the authors**

V. I. Chalapa – Postgraduate Student

A. A. Kosova – Candidate of Sciences (Medicine), associate professor

A. V. Semenov – Doctor of Sciences (Biology)

УДК 616.6

## **ОЦЕНКА БИОЦЕНОЗА УРОГЕНИТАЛЬНОГО ТРАКТА У МУЖЧИН РАЗНОГО ВОЗРАСТА**

Милана Арсеновна Эржибова<sup>1</sup>, Леонид Иосифович Савельев<sup>2</sup>, Полина Геннадьевна Аминева<sup>3</sup>

<sup>1,2</sup>ФГБОУ ВО «Уральский государственный медицинский университет» Минздрава России, г. Екатеринбург, Россия

<sup>3</sup>Медицинский центр «Кволити Мед», г. Екатеринбург, Россия

<sup>1</sup>miss.erjibowa@yandex.ru

## **Аннотация**

**Введение.** Урогенитальные инфекции имеют важную медицинскую и социальную значимость, так как представляют угрозу репродуктивному здоровью лицам мужского пола. Этиологическим фактором могут являться как специфические патогены, так и условно-патогенные микроорганизмы. **Цель**