UDC 57.087.1: 616-036.22: 578.82/.83
DOI: 10.21668/health.risk/2022.3.11.eng

Read online

Research article

SUBSTANTIATION OF STATISTICAL MODEL TO DESCRIBE AND PREDICT RISKS OF TICK BITES FOR POPULATION

V.A. Mishchenko[1,2], I.A. Kshnyasev[2], Yu.A. Davydova[2], I.V. Vyalykh[1]

[1]Yekaterinburg Research Institute of Viral Infections, State Research Center of Virology and Biotechnology "Vector", Federal Service for Surveillance on Consumer Rights Protection and Human Well-being, 23 Letnyaya Str., Ekaterinburg, 620030, Russian Federation
[2]Institute of Plant and Animal Ecology, Ural Branch of Russian Academy of Sciences, 202 8 Marta Str., Ekaterinburg, 620144, Russian Federation

Incidence of tick-borne encephalitis and other tick-borne infections correlates with a number of people applying for medical aid due to tick bites. Obviously, the number of registered tick bites is proportionate to people's economic and recreational activities on an endemic territory and the quantity of hungry ticks. In its turn, the quantity of ticks depends on abundance of main hosts for blood-feeding stages but with a certain time lag caused by their life cycle parameters such as molting to the next stage, diapauses, and apparent seasonality in a continental boreal climate zone.

Our research goal was to analyze and synthesize an adequate formalized/parameterized statistical model to describe and predict risks of tick bites for population.

To describe dynamics and to predict a number of people bitten by ticks exemplified by the Sverdlovsk region, we used several linear (by parameters) logistic regression models. We applied a multimodel inference framework to assess whether the observed dynamics was described adequately. Long-tern dynamics of the number of people bitten by ticks in the Sverdlovsk region is characterized with an occurring high-amplitude slow long-wave oscillation (circadecadal one, with a quasi-period being approximately 10 years) and a short-wave 2–3-year cyclicity. The former may be associated with climatic rhythm and socioeconomic trends; the latter may be caused by biotic factors.

By using the logit-regression model, we showed that the number of small mammals, both in the previous year and at the beginning of the current tick activity season can be a valuable predictor of a risk for population to be bitten by ticks.

Predictive values of the created statistical model adequately describe an initial time series of chances / probabilities of tick bites.

Keywords: *ticks, small mammals, affected by tick bites, pathogen transmission, population dynamics, population cycles, odds ratio, time series.*

Tick-borne encephalitis (TBE) is an endemic disease typical for the Central and Eastern Europe, some parts of the Northern Europe as well as for the Northern and Central Asia. Annually, 10 to 12 thousand TBE cases are registered on these territories [1].

The Urals Federal District (UFD) is highly endemic in terms of tick-borne encephalitis; within its boundaries, the greatest number of bitten people and the highest incidence rates are usually registered in the Sverdlovsk and Kurgan regions [2, 3]. Registered TBE incidence correlates with a number of people who apply for medical aid due to tick bites. Typically, there are periodical changes in incidence of tick-borne encephalitis. They have been detected over many decades and are associated with such factors as demography, changes in land use and relevant density of organisms (population) in wilderness, and

Vladimir A. Mishchenko – Researcher at the Laboratory of Vector-Borne Viral Infections and Tick-Borne Encephalitis, Research Engineer at the Laboratory of Evolutionary Ecology (e-mail: mischenko_va@eniivi.ru; innamoramento23@yandex.ru; tel.: +7 (343) 261-99-47, +7 (965) 515-34-89; ORCID: https://orcid.org/0000-0003-4280-283X).

Ivan A. Kshnyasev – Candidate of Biological Sciences, Senior Researcher at the Laboratory of Population Ecology and Modeling (e-mail: kia@ipae.uran.ru; tel.: +7 (343) 210-38-58; ORCID: https://orcid.org/0000-0002-6281-7644).

Yulia A. Davydova – Candidate of Biological Sciences, Senior Researcher at the Laboratory of Population and Community Ecotoxicology (e-mail: davydova@ipae.uran.ru; tel.: +7 (343) 210-38-58; ORCID: https://orcid.org/0000-0002-3721-629X).

Ivan V. Vyalykh – Candidate of Veterinary Sciences, Leading Researcher, Head of the Laboratory of Vector-Borne Viral Infections and Tick-Borne Encephalitis (e-mail: vialykh_iv@eniivi.ru; tel.: +7 (343) 261-99-47; ORCID: https://orcid.org/0000-0002-3123-8359).

people's recreational behavior. We should also remember that the climate change makes a significant contribution to the process as a possible driving force for cyclicity of TBE-related processes [4].

In Western Europe, virus of tick-borne encephalitis (TBE) is mostly transmitted by *Ixodes ricinus* whereas Siberian and Far Eastern strains are carried by *I. persulcatus* [5]. Ticks usually have several stages in their life cycle. Each stage in tick life cycle lasts from several months to a year, therefore, the whole cycle usually takes from two to three years. However, it may last longer, from two to six years, depending on a geographical location and sustenance. A transfer between stages in tick development is directly linked to changes related to populations of small mammals (SM) since they are basic feeders for larvae and nymphs and (though, to a lesser extent) for imago [6]. Apparent drastic changes in density of small mammal populations with the 2–3 order amplitude and a quasi-period from 3 to 5 years are well known as an example of so called population cycles [7, 8]. Abundance of small animals is assumed to have a positive effect (with certain lags) on quantities in which subsequent stages of ticks (basically, larvae and nymphs) are generated. There is a direct relation between a number of free and hungry ticks that failed to find a usual feeder and risks for people to get bitten by them [9].

Our research goal was to analyze and synthesize an adequate formalized/parameterized statistical model to describe and predict risks of tick bites for population.

Materials and methods. We took initial data (a number of bitten people in the Sverdlovsk region) from the federal statistical report form No. 2 "Data on communicable and parasitic diseases" (Section 1) (issued in 2017–2021) and from the materials provided by the Sverdlovsk Regional Center for Hygiene and Epidemiology (data collected in 1992–2006).

SM quantities were calculated at stationary spots by using break-back traps and wooden live traps [10] in spring and fall over the period from 1992 to 2021.

We calculated odds of tick bites in order to quantify influence exerted by SM density on number of people who applied for medical aid due to tick bites. *Odds* are a ratio of a number of bitten people (N_1) to the whole population of Sverdlovsk region, bitten people excluded (N_0).

There is a possible way from odds to a self-explanatory scale of likelihood or risks (1):

$$P(X) = \frac{\exp(LogOdds)}{1+\exp(LogOdds)} = \frac{Odds}{1+Odds}. \quad (1)$$

We applied logit regression, a conventional apparatus of the generalized linear models theory[1] (2):

$$Ln\left(\frac{N_1}{N_0}\right) = b_0 + \sum b_i X_i. \quad (2)$$

We estimated effects produced by the following (exogenous) predictors of (X_i): SM density in spring and fall in the previous year (SMspr(t–1) and SMfall(t–1)) and in spring in the current year (SMspr(t)). A time series (a trend) of the dependable variable Log odds ticks attacks, or LOTA, was apparently non-stationary. To consider it, we used several approaches (some strict, some heuristic to a certain extent):

1) an autoregression member of the first order being included into the model: AR(1)X where the LOTA(t–1) values observed in the previous year served as an additional predictor;

2) preliminary LOTA time series smoothing (and subsequent inclusion into the model as an additional predictor) by using local regression (*loess*); the optimal value of the smoothing parameter (*span* = 0.427) was identified with 10-fold cross-validation;

3) preliminary STL (Seasonal and Trend decomposition) of LOTA time series by using Loess and identification of the trend by smoothing it with local regression.

The initial time series and its smoothed variants are shown in Figure 1.

[1] McCullagh P., Nelder J.A. Generalized linear models. London, Chapman and Hall, 1989, 511 p.

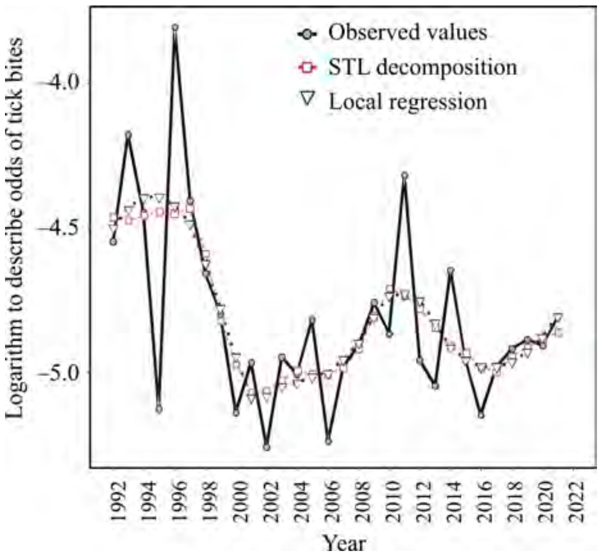

Figure 1. Dynamics of the logarithm describing odds of tick bites and smoothed variants of the rime series used to identify the trend. Sverdlovsk region, 1992–2021

Odds ratios (*OR*) and their confidence intervals (95 % CI) are given after the transformation: $OR = \exp(b_i)$ or $OR^{-1} = 1/\exp(b_i)$, where b_i are the logit regression parameters (odds ratio logarithms). Odds ratios for rare events (frequency < 10^{-1}) approach relative risk levels and it simplifies the interpretation.

We applied *Akaike information criterion or AIC* to compare and rank our logit regression models. This criterion determines optimality as a compromise between a model accuracy and complexity. A model with greater statistical adequacy corresponds to the smallest *AIC* value. The models were compared based on the modification of the initial *AIC*, *consistent Akaike information criterion (CAIC)*, calculated as per the formula (3):

$$CAIC = -2LL + k\left[1 + \ln(m)\right], \qquad (3)$$

where *LL* is the logarithm of the likelihood function maximum, *k* is the number of parameters, and *m* is the number of observations. This modification, in comparison with *AIC*, imposes stricter "penalties" for additional parameters[2].

"Weight" (relative plausibility) of each model calculated as per the formula (4) was applied to rank and compare competing models:

$$w_i = \frac{\exp(-0.5\Delta CAIC_i)}{\sum \exp(-0.5\Delta CAIC_i)}. \qquad (4)$$

This "weight" can be interpreted as a posterior probability that the *i*-th model is the best given the examined multiplicity of other candidate models. If the "weight" differed from w_{max} by less than 10 %, we considered these models to be identical as per the quality of the best one [11].

The results were statistically analyzed and visualized with Statistica v. 10.0 applied software package (StatSoft, Inc) and the R (v. 3.4.4) system for statistical computation and graphics [12].

Results and discussion. The observed long-term dynamics of odds for people to be bitten by ticks is obviously a non-stationary (not as per the mean value, or amplitude, or a period) time series (Figure 1). We can spot out the following in it:

– a trend as a long-term drift of mean values, relatively slow changes over time;

– a quasi-periodical component that is repeated relatively fast;

– remains as an irregular component in the series, relatively high-frequency noise.

The trend is probably an "echo" of both 1) a long circadecadal weather and climate cycle associated with large-scale phenomena (formation and duration/stability of atmospheric circulations and such phenomena as anticyclones that block the transfers from west) responsible for typical long-term temperature fluctuations and precipitations in a given region [13, 14]; and 2) a peculiar socioeconomic situation that existed in the Russian Federation in 90ties last century and its following improvement [15].

Quasi three-year cycles (T = 2–3 years) might be associated both with parameters of tick lifecycle and dynamics of SM numbers

[2] Akaike H. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 1974, vol. 19, pp. 716–723.

[16], such periods being quite typical for the latter [17]. SM and tick populations are interrelated as years when ticks are highly active (their activity is estimated as per number of bitten people) follow years with maximum density of small rodents [18]. Larvae and nymphs mostly feed on small mammals and this leads to natural TBE transmission.

Dynamics of the logarithm for odds of tick bites is described optimally ($w \approx 1$, $CAIC$ = 13005588) with a model that contains a trend obtained by STL-decomposition. Autoregression and local smoothing models are at a "significant disadvantage" in comparison to it: $w \sim 0$, $CAIC$ = 13009459 and $w \sim 0$, $CAIC$ = 13072412, accordingly.

The optimal model includes five parameters: the trend that was isolated from the LOTA time series by STL-decomposition (trend-STL); SM density: in spring in the current year, in spring and fall of the previous year; a free term b_0 (Table 1).

We included a trend estimate as an additional component into the list of predictors in the logit regression model. This made it possible to consider, though only heuristically, non-stationarity of the time series and estimate certain effects which we thought were interesting (Figure 2A).

A growth in SM density in spring last year will lead to ($p < 0.01$) odds of tick bites (to the Sverdlovsk region population) growing by 1.04 times (by 3.53 % higher, Figure 2B). A growth in the number of ticks at different stages (nymphs or imago) that are out from the diapause is very likely in the current year due to abundance of feeders that occurred last spring.

Greater SM density this spring will lead to a statistically significant decrease by 1.06 times in odds of tick bites (by 5.87 % lower, Figure 2C). Therefore, SM population in its growing phase "intercepts" tick nymphs and imago since sufficient quantities of feeders make tick bites less likely for large mammals and people.

Growing SM density last fall will lead to a statistically significant growth by 1.02 times in odds of tick bites (by 1.98 % higher, Figure 1D). SM population in fall is a projection of its density (quantity) in spring and summer. Therefore, we can expect a growing number of ticks at different stages in this tick season and it makes for growing likelihood of people being bitten during it.

Figure 3 shows the current risks of tick bites and risks (likelihood) expected (predicted) by the model for the Sverdlovsk region population.

Obviously, levels predicted by the model are quite consistent with the initial time series: both the trend and the periodical component are similar both for actual levels and those predicted by the statistical model (the correlation is strong and direct: Spearman's correlation coefficient r_S = 0.70; $p < 0.0001$). The determination coefficient is quite high: R^2 = 0.66, therefore, the model is able to explain 66 % of dispersion regarding dynamics of the logarithm for a risk (or likelihood) of tick bites. If we include additional predictors into our model, for example, those associated with dynamics of tick populations and their feeders' ones, climatic changes and socioeconomic factors, this is expected to improve its predictive capabilities.

Table 1

Parameters of the optimal model to describe long-term dynamics of the logarithm for odds of tick bites. Sverdlovsk region, 1992–2021 ("the best" logit regression model: LR (4) = 99827; $p < 0.0001$; R^2 = 0.66)

Parameters	b	SE(b)	Wald test (Z)	p	Odds ratio		
					OR	*95 % CI*	
b_0	1.10	0.02	47.72	< 0.0001	–	–	–
trend-STL	1.24	0.004	272.10	< 0.0001	3.45	3.42	3.48
SMspr(t–1)	0.03	0.001	41.04	< 0.0001	1.04	1.03	1.04
SMspr(t)	-0.06	0.001	-80.81	< 0.0001	1.06^{-1}	1.057^{-1}	1.061^{-1}
SMfall(t–1)	0.02	0.001	20.83	< 0.0001	1.02	1.017	1.021

Note: b_0 is a free term; *OR* is odds ratio; *LR(df)* is the test of ratios between likelihood and a number of degrees of freedom; *trend-STL* is the trend isolated from LOTA time series by STL-decomposition; SM density is: *spr(t–1)* last spring, *spr(t)* this spring, *fall(t–1)* last fall; -1 – OR^{-1} = 1/exp (b_i).

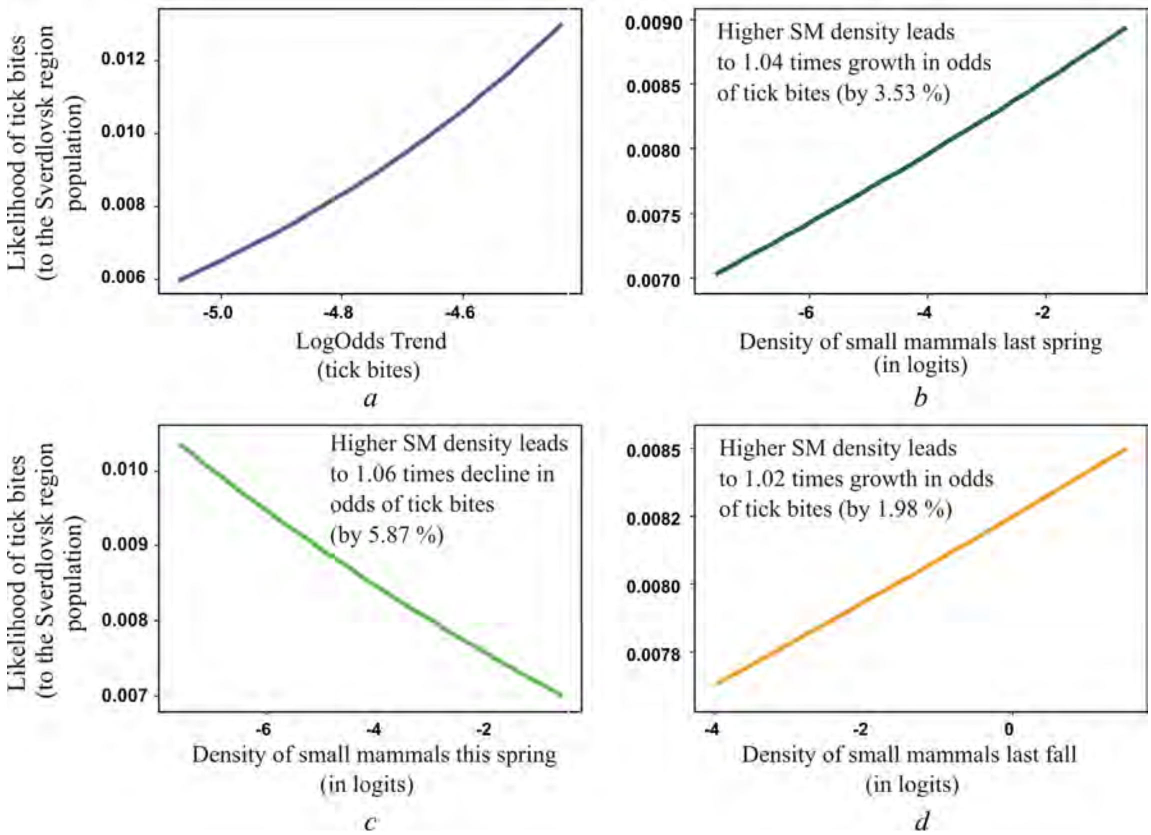

Figure 2. Dependence between likelihood of tick bites and density of small mammals: **A** is dependence on the logarithm for odds of tick bites; dependence on density of small mammals: **B** is last spring, **C** is this spring; **D** is last fall

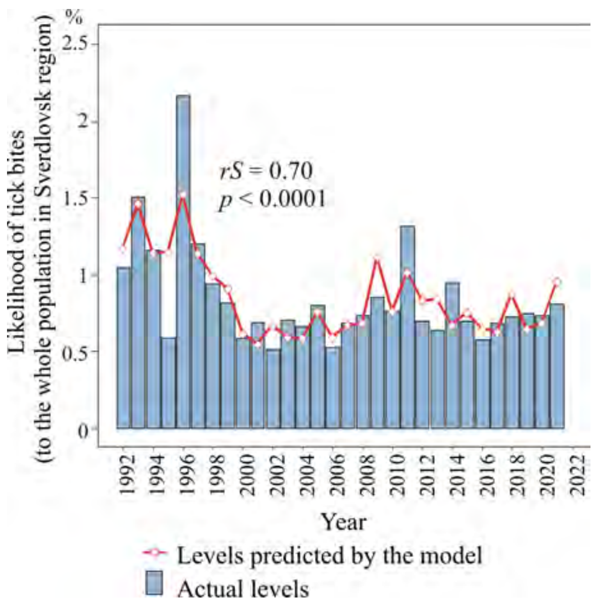

Figure 3. Actual and predicted likelihood of tick bites for the Sverdlovsk region population, 1992–2021

Conclusions:

1. Long-term dynamics of the number of people bitten by ticks can be adequately described with a model of a non-stationary time series that contains the trend and the quasi non-periodical component.

2. The logit regression model clearly showed that numbers of small mammals in the previous year and at the beginning of the current epidemiologic season when ticks are especially active can serve as a risk predictor for population warning them tick bites are quite possible. High density of small mammals in spring in the current tick activity season leads to the established "interception" effect and quantitatively estimated decline in likelihood of people being bitten by ticks. However, if density of small mammals was high in spring and fall of the year prior to the current tick activity season, then we should expect a growth in numbers of ticks at different stages in their lifecycle that have molted and survived winter (nymphs, imago) and growing risks for people to be bitten by ticks.

3. Levels predicted by the created statistical model adequately describe the initial time series of odds / likelihood of tick bites. Addi-

tional parameters included into the model will probably improve its predictive capabilities.

Funding. The study was accomplished within the research work "Complex analysis of population immunity to tick-borne encephalitis virus in regions located in the Urals Federal District" (Registration No. is 121040500096-4).

Competing interests. The authors declare no competing interests.

References

1. Kahl O., Pogodina V.V., Poponnikova T., Süss J., Zlobin V.I. A short history of TBE. Chapter 1. *The TBE Boom*, 3rd ed. In: G. Dobler, W. Erber, M. Bröker, H.-J. Schmitt eds. Singapore, Global Health Press, 2020, pp. 10–15. DOI: 10.33442.26613980_TBE-3

2. Zlobin V.I. Epidemiologicheskii monitoring i profilaktika iksodovykh kleshchevykh infektsii v usloviyakh sochetannykh prirodnykh i antropurgicheskikh ochagov [Epidemiological monitoring and prevention of tick-borne cases in conditions of combined natural and anthropurgic foci]. *Epidemiologiya i vaktsinoprofilaktika*, 2008, vol. 39, no. 2, pp. 10–14 (in Russian).

3. Luchinina S.V., Stepanova O.N., Pogodina V.V., Sten'ko E.A., Chirkova G.G., Gerasimov S.G., Kolesnikova L.I. Sovremennaya epidemiologicheskaya situatsiya po kleshchevomu virusnomu entsefalitu v Chelyabinskoi oblasti [Modern epidemiological situation of tick-borne encephalitis in Chelyabinsk region, Russia]. *Epidemiologiya i vaktsinoprofilaktika*, 2014, vol. 75, no. 2, pp. 32–37 (in Russian).

4. Rubel F., Walter M., Vogelgesang J.R., Brugger K. Tick-borne encephalitis (TBE) cases are not random: explaining trend, low- and high-frequency oscillations based on the Austrian TBE time series. *BMC Infectious Diseases*, 2020, vol. 20, no. 448, pp. 1–12. DOI: 10.1186/s12879-020-05156-7

5. Gritsun T.S., Nuttall P.A., Gould E.A. Tick-Borne Flaviviruses. In: T.J. Chambers, T.P. Monath eds. *Advances in Virus Research, Academic Press*, 2003, vol. 61, pp. 317–371. DOI: 10.1016/S0065-3527(03)61008-0

6. Bournez L., Umhang G., Moinet M., Richomme C., Demerson J.-M., Caillot C., Devillers E., Boucher J.-M., Hansmann Y., Boué F., Moutailler S. Tick-Borne Encephalitis Virus: Seasonal and Annual Variation of Epidemiological Parameters Related to Nymph-to-Larva Transmission and Exposure of Small Mammals. *Pathogens*, 2020, vol. 9, no. 7, pp. 518. DOI: 10.3390/pathogens9070518

7. Kshnyasev I.A., Davydova Y.A. Dinamika plotnosti i struktury populyatsii lesnykh polevok v yuzhnoi taige [Dynamics of density and structure of forest vole populations in the southern taiga]. *Vestnik Nizhegorodskogo universiteta im. N.I. Lobachevskogo. Seriya: Biologiya*, 2005, no. 1, pp. 113–123 (in Russian).

8. Hanski I., Turchin P., Korpimaki E., Henttonen H. Population oscillations of boreal rodents: regulation by mustelid predators leads to chaos. *Nature*, 1993, no. 364, pp. 232–235. DOI: 10.1038/364232a0

9. Ostfeld R.S., Levi T., Keesing F., Oggenfuss K., Canham C.D. Tick-borne disease risk in a forest food web. *Ecology*, 2018, vol. 99, no. 7, pp. 1562–1573. DOI: 10.1002/ecy.2386

10. Kshnyasev I.A., Davydova Y.A. Population cycles and the Chitty syndrome. *Russian Journal of Ecology*, 2021, vol. 52, no. 1, pp. 70–75. DOI: 10.1134/S1067413621010082

11. Burnham K.P., Anderson D.R. Model selection and multimodel inference: a practical information-theoretic approach. N.Y., Springer-Verlag, 2002, 496 p.

12. R Development Core Team. R: a language and environment for statistical computing. Vienna, Austria, The R Foundation for Statistical Computing, 2020. Available at: https://www.R-project.org (01.08.2022).

13. Gray J.S., Dautel H., Estrada-Peña A., Kahl O., Lindgren E. Effects of climate change on ticks and tick-borne diseases in Europe. *Interdisciplinary Perspectives on Infectious Disease*, 2009, vol. 2009, pp. 1–12. DOI: 10.1155/2009/593232

14. Daniel M., Danielová V., Fialová A., Malý M., Kříž B., Nuttall P.A. Increased relative risk of tick-borne encephalitis in warmer weather. *Frontiers in Cellular and Infection Microbiology*, 2007, vol. 8, pp. 90. DOI: 10.3389/fcimb.2018.00090

15. Noskov A.K., Nikitin A.Y., Andaev E.I., Pakskina N.D., Yatsmenko E.V., Verigina E.V., Innokent'eva T.I., Balakhonov S.V. Kleshchevoi virusnyi entsefalit v Rossiiskoi Federatsii: osobennosti epidemicheskogo protsessa v period ustoichivogo spada zabolevaemosti, epidemiologicheskaya situatsiya v 2016 g., prognoz na 2017 g. [Tick-borne virus encephalitis in the Russian Federation: features of epidemic process in steady morbidity decrease period. Epidemiological condition in 2016 and the forecast for 2017]. *Problemy osobo opasnykh infektsii*, 2017, no. 1, pp. 37–43. DOI: 10.21055/0370-1069-2017-1-37-43 (in Russian).

16. Rubel F., Brugger K. Tick-borne encephalitis incidence forecasts for Austria, Germany, and Switzerland. *Ticks and Tick-borne Diseases*, 2020, vol. 11, no. 5, pp. 101437. DOI: 10.1016/j.ttbdis.2020.101437

17. Clement J., Maes P., van Ypersele de Strihou C., van der Groen G., Barrios J.M., Verstraeten W.W., van Ranst M. Beechnuts and outbreaks of Nephropathia epidemica (NE): of mast, mice and men. *Nephrology Dialysis Transplantation*, 2010, vol. 25, iss. 6, pp. 1740–1746. DOI: 10.1093/ndt/gfq122

18. Balashov Y.S. Demography and population models of ticks of the genus Ixodes with long-term life cycles. *Entomological Review*, 2012, vol. 92, no. 9, pp. 1006–1011. DOI: 10.1134/S0013873812090072

Mishchenko V.A., Kshnyasev I.A., Davydova Yu.A., Vyalykh I.V. Substantiation of statistical model to describe and predict risks of tick bites for population. Health Risk Analysis, 2022, no. 3, pp. 119–125. DOI: 10.21668/health.risk/2022.3.11.eng

Received: 31.08.2022
Approved: 09.09.2022
Accepted for publication: 25.09.2022