



# OPEN Retrospective multidisciplinary analysis of human alveolar echinococcosis in Hungary using spatial epidemiology approaches

Eszter Csulak<sup>1</sup>, Ágnes Csivincsik<sup>2</sup>, Tamás Sréter<sup>3</sup>, Norbert Solymosi<sup>4</sup>, József Danka<sup>3</sup>, Zsolt Káposztás<sup>1</sup>, Gábor Nagy<sup>2</sup>✉ & Balázs Dezsényi<sup>5</sup>

Human alveolar echinococcosis (HAE), which is caused by the larval stage of the *Echinococcus multilocularis* tapeworm, is an increasing healthcare issue in Hungary. Among the 40 known cases in the country, 25 were detected in the last five years. Our study aimed to reveal the geographically underlying risk factors associated potentially with these cases. We investigated the spatial pattern and the impact of potential risk factors of HAE by cluster analysis, and local and global regression models. Also, a questionnaire survey on the patients' lifestyle was implemented. We found two HAE hyperendemic foci in the country with very dissimilar biotic and climatic features, and controversial impact of different environmental factors. Four factors, viz. forest cover ( $\beta = 0.291$ ,  $p < 0.0001$ ), surface soil wetness ( $\beta = -0.157$ ,  $p = 0.033$ ), fox infection rate ( $\beta = 0.369$ ,  $p < 0.0001$ ) and socio-economic development ( $\beta = -0.216$ ,  $p = 0.009$ ), proved important countrywide. The most forested and the least developed districts showed the highest HAE risk. Among the patients, kitchen gardening (67.86%) and dog ownership (67.86%) seemed the riskiest activities. Our models detected an anomaly in one of the poorest regions of Hungary where all risk factors behaved contrary to that of the neighboring areas. This phenomenon was supposed to be the result of under-detection of the disease, and it called attention to the urgent priority of knowledge dissemination to the public and the healthcare professionals.

**Keywords** Human alveolar echinococcosis, *Echinococcus multilocularis*, Spatial epidemiology, Culture of poverty, Hungary

Human alveolar echinococcosis (HAE) is a zoonotic disease caused by the larval stage of the *Echinococcus multilocularis*, an indirectly developing cestode. Its natural definitive host is the red fox (*Vulpes vulpes*) while its intermediate hosts are rodents<sup>1</sup>. Humans can acquire infection by accidental ingestion of parasite eggs with contaminated food or water, or during contact with infected animals<sup>2,3</sup>. Although the distribution area of HAE extensively grows in the Holarctic realm<sup>1</sup>, it does not get enough attention from either the researchers or the health care systems. Therefore, HAE is considered as a neglected disease, which affects the most vulnerable communities<sup>4</sup>.

In Europe, outside its Alpine historic distribution area, HAE is hypothesized to be under-detected because of low disease awareness of both the public and the healthcare system<sup>3,4</sup>. Therefore, determination of risk factors of HAE and its potentially high-risk areas contribute to better disease awareness, which can enhance diagnostic and treatment success, and also preventive measures<sup>5</sup>. In Hungary, HAE was first described in 2008<sup>6</sup>. A later retrospective study, which analyzed country-wide HAE cases between 2003 and 2018, ascertained that the disease has low incidence but high case-fatality rate due to diagnostic delay and inappropriate treatment<sup>7</sup>.

Epidemiological analyses based on conventional, non-spatial analysis often do not provide satisfactory answers during the investigation of disease occurrence, spread and root causes. This is especially true for rare

<sup>1</sup>Department of Surgery, Somogy County Teaching Hospital, Kaposvár 7400, Hungary. <sup>2</sup>Department of Physiology and Animal Health, Hungarian University of Agriculture and Life Sciences, Kaposvár Campus, Kaposvár 7400, Hungary. <sup>3</sup>Laboratory Department of Bacteriology, Mycology and Parasitology, National Center for Public Health and Pharmacy, Budapest 1097, Hungary. <sup>4</sup>Centre for Bioinformatics, University of Veterinary Medicine, Budapest 1078, Hungary. <sup>5</sup>Department of Surgery, Transplantation and Gastroenterology, Semmelweis University, Budapest 1082, Hungary. ✉email: gabor.nagy.oh@gmail.com

diseases such as infections caused by *E. multilocularis*. In this case, the occurrence and transmission of the parasite in the final hosts or human populations could vary conspicuously on a wide spatial scale. The variations in trends of the disease may occur even globally and regionally, but it could also be characteristic locally<sup>8–10</sup>. The advancement of geographic information technology, sophisticated statistics and geospatial analytical software can improve the knowledge and understanding of diseases. This complex approach could contribute to the foundation of competent integrated preventive control programs against epidemics<sup>11</sup>.

For investigation of spatial heterogeneity, geographically weighted regression (GWR) and multiscale geographically weighted regression (MGWR) models were introduced<sup>12</sup>. In contrast to traditional linear regression models (e.g. ordinary least squares - OLS), these methods can borrow data from neighboring locations over spatial weights. Thus, they can manipulate spatially heterogeneous associations between the dependent and explanatory variables<sup>11</sup>.

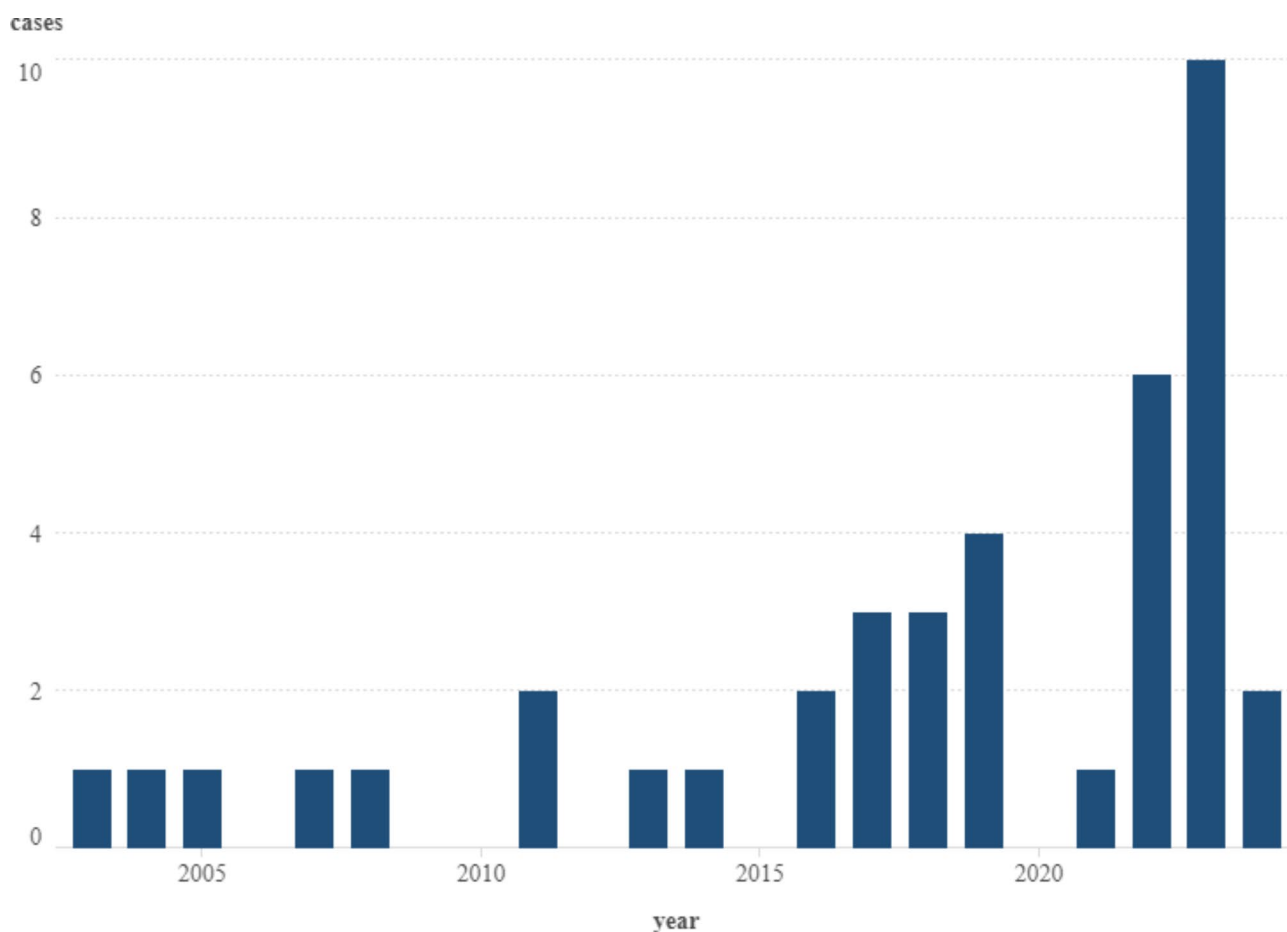
Our study aimed to analyze the spatial epidemiological background of 40 HAE cases that were detected between April of 2003 and April of 2024, in Hungary. Therefore, we compared the cases by their geographical location in a district-based geospatial analysis, and by the patients' life-style features. Besides natural biotic and climatic factors, a socio-economic variable was also involved to estimate the potential additional risk of vulnerable regions.

## Results

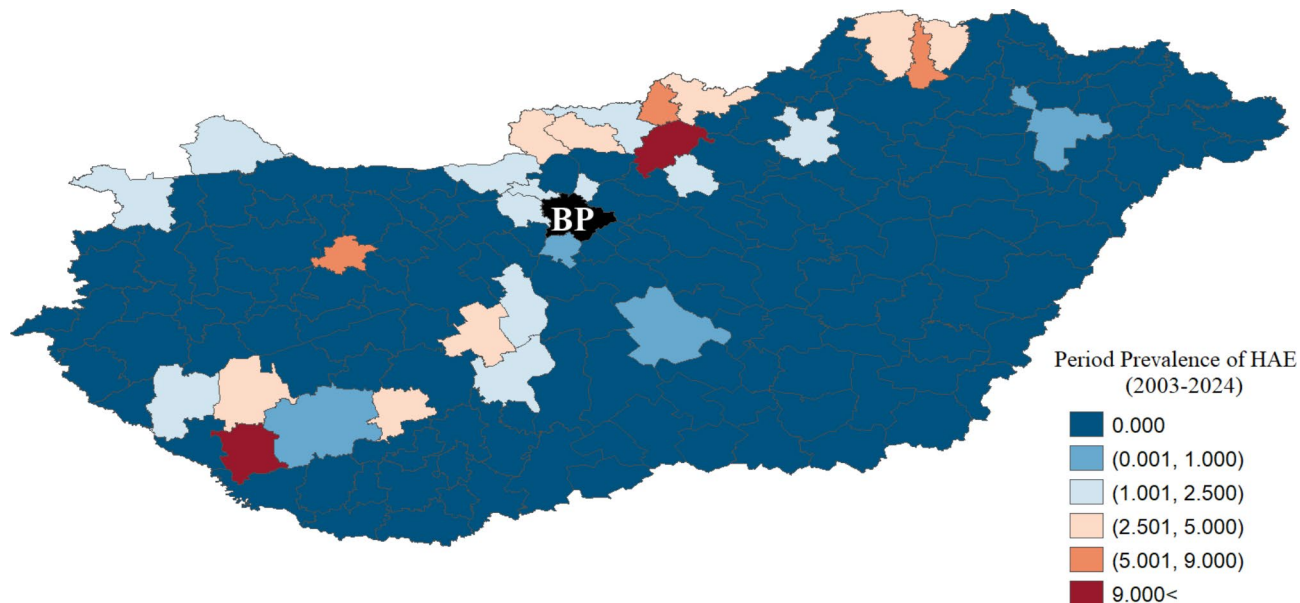
### Descriptive findings

At the beginning of the investigation period, the number of cases was low but increased sharply after 2020 (Fig. 1).

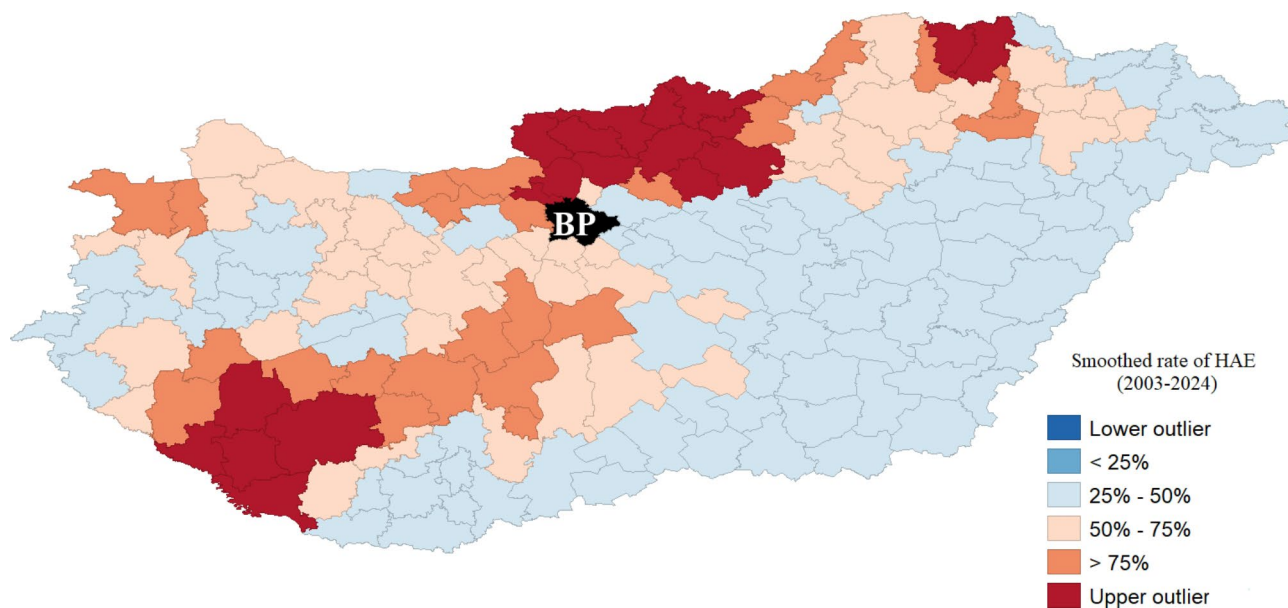
The countrywide overall period prevalence (PP) in humans was 0.4025/period/10<sup>5</sup> with a broad range of 0.0/period/10<sup>5</sup> and 23.07/period/10<sup>5</sup> (Fig. 2). During the study period, 4 out of 40 patients died directly from alveolar echinococcosis. Thus, the case fatality rate proved 10%. The two districts (indicated red in Fig. 2) with the highest PP had six and three human patients each. Of these nine patients, six were detected during the last five years. Due to Budapest's exclusion, 38 out of 40 confirmed cases were included in the spatial analysis which



**Fig. 1.** Bar chart demonstrating the annual number of human alveolar cases during the study period: from 2003 April to 2024 April (Note: The year 2024 represents only four months, using the Software for creation: Scimago Graphica version 1.0.42).



**Fig. 2.** Map of the raw period prevalence of human alveolar echinococcosis (2003–2024) over the 174 Hungarian districts. (Note: The black color with BP indicates the excluded districts of Budapest.) (Software used for creation: GeoDA version 1.22. and PhotoScape version 3.7).



**Fig. 3.** Map of the smoothed rate of HAE risk (note: The black color with BP indicates the excluded districts of Budapest.) (Software used for creation: GeoDA version 1.22. and PhotoScape version 3.7).

were aggregated in 29 districts. Based on the raw data, the most infected districts were mainly observed in the country's northern, northeastern, and southwest regions.

After smoothing, we obtained an extended risk area, which covered 87 districts. On the other hand, owing to smoothing, the areas between the affected local endemics (northeast and southwest) were also estimated to be at higher risk. After smoothing the raw data, it was found that the areas most affected by HAE were the North Hungarian Mountains (north-northeast), a part of South Transdanubia (southwest), and the Little Hungarian Plain (northwest) region. The least endangered area is the southcentral and southeastern parts of the country, thus the Great Hungarian Plain (Fig. 3, Supplementary Information 1).

### Spatial clustering

In the case of HAE, the high-rated cluster (log likelihood ratio = 12.42;  $p = 0.001$ ) involved 16 districts, of which 11 had human cases. The relative risk was 6.03, representing that the probability of HAE occurring within this window was more than six times higher than in areas outside the cluster. The low-rated cluster (log likelihood ratio = 9.40;  $p = 0.012$ ) consisted of 44 districts, in which no human cases were confirmed (relative risk = 0). Three significant clusters were confirmed for foxes, of which two highly overlapped with human clusters (Fig. 4). The low-rated cluster characterized by 17.68 log likelihood ratio ( $p < 0.001$ ), while the high-rated clusters had 12.28 ( $p = 0.001$ ) and 9.60 ( $p = 0.026$ ) log likelihood ratio, respectively.

The most conspicuous difference in the clustering of human and fox cases was observed in the Transdanubian region, which was proved to be a high-risk area for fox infection contrary to the lower number of human cases (Table 1).

### Global (OLS) and local (GWR and MGWR) regression models

During the multicollinearity checking, we excluded one variable (proportion of agricultural areas - AGRO) because its variance inflation factor (VIF) exceeded the threshold level ( $5 < \text{VIF}$ ). The final result of our OLS model was relatively poor ( $R^2 = 0.237$ ). Based on the standard t-value threshold ( $\pm 1.96$ ), four independent variables (the surface soil wetness - GWTOP, the mean earth skin temperature - TSMEAN, the proportion of forests - FOR, the district development index - DDI and the smoothed yield of fox infection - FOXRISK) had a stationary statistically verifiable effect on the smoothed yield of human infection (HUMRISK). When the t-value was greater than 1.96, we accepted it as a positive significant relationship. At the same time, the t-test generated a lesser number than  $-1.96$ , and the association was evaluated as negatively significant. The examination of the residuals suggested a strong autocorrelation (Moran's  $I = 0.496$ ,  $p = 0.0002$ ).

The GWR and MGWR models showed considerable improvement in residual sum of squares (RSS), the Akaike Information Criterion (AIC), and the adjusted coefficient of the model (adj.  $R^2$ ), and revealed the spatial relationships between the variables (Table 2). The results made it possible to conclude that the best model was MGWR. This statement was reinforced by the fact that spatial autocorrelation could no longer be detected in the case of the residuals of the model (Moran's  $I = 0.042$ ,  $p = 0.134$ ).

When the OLS and MGWR models were compared, it was found that the effect of certain variables varied significantly from model to model. The OLS model indicated that the forest cover (FOR) and the smoothed fox infection rate (FOXRISK) had a positive correlation with the smoothed HAE risk. The district development (DDI) and, interestingly, the wetness of the ground surface (GWTOP) were negatively associated with HUMRISK. The MGWR results confirmed that two of the four above-mentioned variables were not considered stationary (GWTOP and FOXRISK), as their significant local effects were also confirmed. The soil surface temperature (TSMEAN) estimation also showed a non-stationary effect. The p-value of DDI and the FOR in the MGWR proved that these variables had a universal relationship with HAE risk across Hungary (Table 3). The bandwidth in the MGWR model denoted the side effect of the influencing determinants. The smaller the bandwidth, the larger the spatial heterogeneity over the study area and vice versa. When bandwidth is higher, the impact of a particular variable may not be local but regional or global<sup>13</sup>.

#### *Spatial non-stationarity in HAE patterns*

The local  $R^2$  map indicated that the total variance of the MGWR model ranged between 0.459 and 0.866. The best-fitting values were observed in Hungary's central, northern and southwestern parts (Fig. 5). The local model did not fit as firmly with data from the southernmost part of the Great Plain, so the local  $R^2$  value in these areas was only moderate.

Figure 6 demonstrates the spatial fluctuation between the significant non-stationary explanatory variables and HUMRISK. The regression coefficient of TSMEAN varied widely (from  $-0.611$  to  $0.511$ ). Significant relationships were detected in the north-south direction in the central zone of the country and the western region. The local standard t-value thresholds indicated that both lower and higher mean ground surface temperatures influenced human infection rates. The TSMEAN estimation in the OLS was very weak and non-significant (see Table 4).

GWTOP showed a homogeneous impact on the districts. The ground surface wetness had a negative non-stationary association with HUMRISK on a large portion of the country (mainly the total Transdanubian region). This variable's effect was insignificant in the dryer regions (e.g. Great Hungarian Plain).

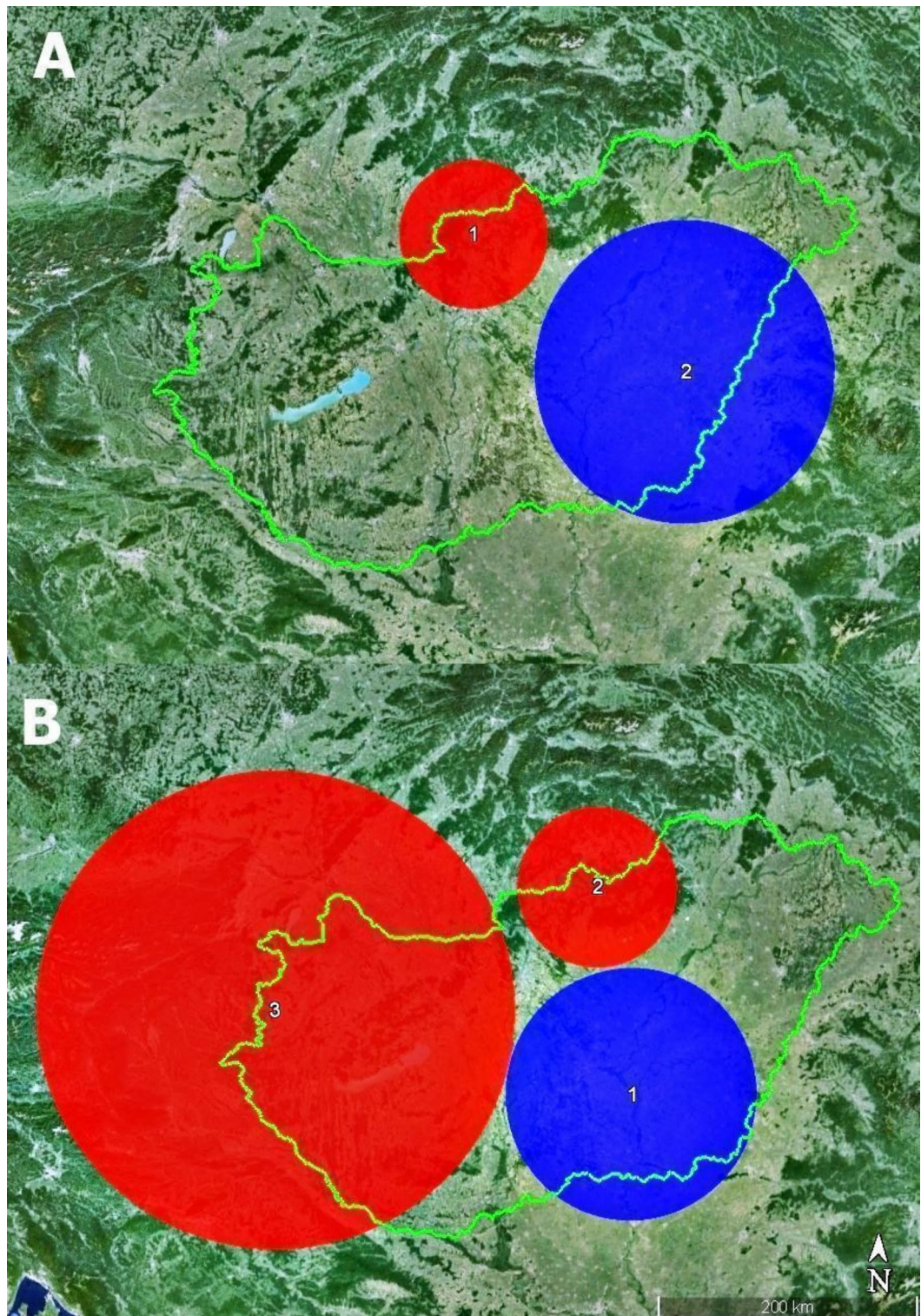
The effect of the smoothed fox infection rate (FOXRISK) showed a similar spatial aggregation pattern to GWTOP. However, in this case, the positive, significant spatial relationship was confirmed in Hungary's central, northern and eastern regions. On the other hand, two districts in the South Transdanubia showed a provable negative correlation (in both cases  $\beta = -0.225$ ).

### Questionnaire

During the study period (2003–2024), 40 patients were diagnosed with HAE in Hungary (female,  $n = 22$ ; male,  $n = 18$ ), thus the sex ratio (male/female = 0.8) differed slightly, but this difference was not statistically significant ( $p = 0.530$ ). The median age of patients was 56.6 years (CI 95% 51–67), with the youngest and the oldest patients being 12 and 80 years old, respectively. The median age of the two sexes was statistically not different,  $p = 0.341$  (female: 60.5 years, CI 95% 51–71; male: 53 years, CI 95% 32–62).

Our questionnaire consisted of 13 decidable questions, mainly related to pet keeping, outdoor activity, and some hygiene habits related to fruit and vegetable consumption (Supplementary Information 2). Education levels were sorted into four categories. Among 18 of the 28 respondents who gave education level, two thirds of them reached at least secondary level (Supplementary Information 2). Based on the responses, it was possible to conclude that the proportion of the four most frequently chosen answers among the respondents exceeded





**Fig. 4.** The confirmed human (A) and fox (B) spatial clusters. (Note: Red indicates the significant high-risk areas with higher relative risk, while blue marks the significant low-risk areas where the relative risk is low. The order of numbering was given by SaTScan Software.) (Software used for creation: SaTScan version 10.1, Google Earth Pro version 7.3.6.9345 and PhotoScape version 3.7).

50%. Owning and playing with dogs occurred in 67.86% (CI 95% 48.2–82.5) and 57.14% (CI 95% 38.06–74.24) of respondents, respectively. The other two factors were related to outdoor activities (kitchen garden and regular forest walk), accounting for more than 60%. One of the factors next in the ranking (agricultural activity) was also related to outdoor activities. At the same time, the consumption of unwashed fruits, vegetables, and forest

Cluster type	Central coordinates	Number of districts	Radius (km)	Relative risk	p-value
Human clusters					
High-rated 1*	47.929150 N 19.137118 E	16	53.28	6.03	0.001
Low-rated 2**	47.024704 N 21.172044 E	44	109.72	0	0.012
Fox clusters					
High-rated 2***	48.102382 N 19.805404 E	15	59.72	2.41	0.001
High-rated 3****	47.235514 N 16.622270 E	63	178.04	1.99	0.024
Low-rated 1*****	46.709441 N 20.143731 E	39	94.16	0.13	<0.001

**Table 1.** Spatial clusters of *E. multilocularis* infection were detected in the human and fox populations in Hungary. \*Indicated in Fig. 4A this cluster signed with 1. \*\*Indicated in Fig. 4A this cluster signed with 2. \*\*\*Indicated in Fig. 4B this cluster signed with 2. \*\*\*\*Indicated in Fig. 4B this cluster signed with 3. \*\*\*\*\*Indicated in Fig. 4B this cluster signed with 1.

	OLS	GWR	MGWR
Adj. R <sup>2</sup>	0.237	0.519	0.751
AIC	455.546	397.242	289.085
RSS	127.396	68.056	32.966

**Table 2.** The goodness of fit statistics of ordinary least regression (OLS), geographically weighted regression (GWR) and multiscale geographically weighted regression (MGWR) models. *adj. R<sup>2</sup>* adjusted coefficient of the model, *AIC* akaike information criterion, *RSS* residual sum of squares.

Variable	(OLS)	Mean (MGWR)	SD of (MGWR)	Min (MGWR)	Median (MGWR)	Max (MGWR)	Monte Carlo p-value (MGWR)	Bandwidth (CI 95%) <sup>#</sup> MGWR
TSMEAN	0.035	-0.121	±0.314	-0.545	0.175	0.619	0.040	43 (43–49)
GWTOP	-0.157*	-0.248	±0.138	-0.452	-0.198	-0.089	0.033	114 (92–143)
URB	0.099	0.086	±0.005	0.075	0.088	0.093	0.855	173 (124–173)
FOR	0.291***	0.142	±0.221	-0.204	0.070	0.939	0.108	43 (43, 47)
WETHABIT	-0.068	-0.056	±0.015	-0.081	-0.052	-0.036	0.544	173 (124–173)
DDI	-0.216**	-0.172	±0.130	-0.491	-0.184	0.078	0.318	43 (43–49)
FOXRISK	0.369***	0.294	±0.322	-0.225	0.261	1.049	0.000	43 (43–47)

**Table 3.** Comparison of the standardized regression coefficients in OLS and MGWR models. \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.0001$ , <sup>#</sup>confidence interval 95%.

products was higher than 46% of respondents. Direct contact with foxes had the lowest proportion of responses (Fig. 7).

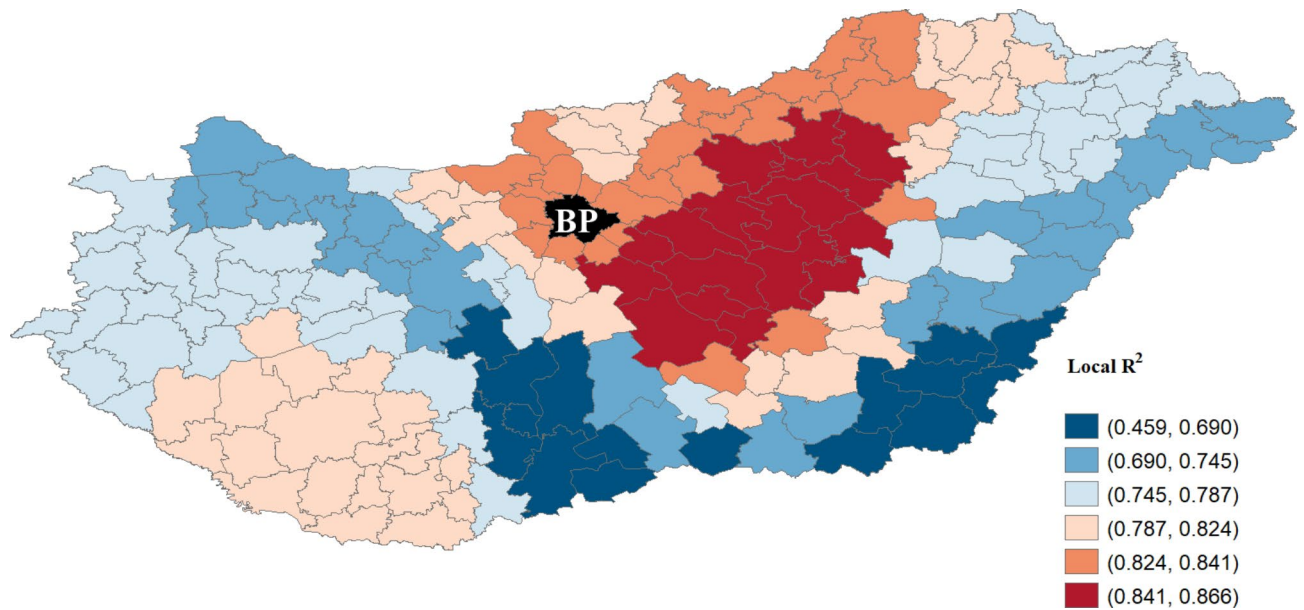
## Discussion

Human alveolar echinococcosis is an increasing health problem in Hungary. Between 2003 April and 2024 April, 40 cases were detected, of which 25 entered the health care system during the last 5 years. Two hyperendemic foci are supposed to form in the country, one in the Northern Hungarian Mountains, and another in the southwestern part of the country, the South Transdanubia. The period prevalences for these focus areas,  $9.46/10^5$  and  $23.07/10^5$ , are similar to those, which were detected in the previously identified highly endemic regions close to the historical Alpine focus area<sup>2</sup>.

Cluster analysis by SaTScan software could detect a high-risk and a low-risk HAE cluster in the eastern part of the country while the South Transdanubian case accumulation remained non-detected by this method. The Bernoulli-based approach uses the presence of cases in a district as a quality regardless of the prevalence in the district concerned. The South Transdanubian case accumulation involved only four districts as a contiguous area, and three districts further with plenty of negative districts around these spots, therefore the software identified this phenomenon as a coincidence.

The spatial rate smoothing could highlight the potential risk of HAE in the South Transdanubia, which remained undetected by cluster analysis. In the case of foxes, the smoothing process displayed a pattern that agrees with the strong hypothesis that *E. multilocularis* spread to Hungary from the northwest direction<sup>14</sup>. Smoothed risks of both HAE and fox echinococcosis removed the least risky classification from the Great Plain districts. This procedure can provide a more appropriate estimation of risk than use of crude proportions, especially when





**Fig. 5.** Spatial distribution of the local regression coefficient (Local  $R^2$ ) for MGWR. (note: the black color with BP indicates the excluded districts of Budapest) (software used for creation: GeoDA version 1.22. and PhotoScape version 3.7).

the sample is not optimal to represent the whole population at risk<sup>15</sup>. Therefore, increased smoothed risk within the apparently free area reflected the possibility of disease occurrence.

For analyzing those variables that could influence the smoothed risk of HAE in a district, we applied the MGWR model, which could handle the heterogeneity of the candidate explanatory variables. The model provided a high  $R^2$  values (0.459–0.866) for almost the whole territory of the country, except for the southernmost areas of the Great Plain where the fitness of the model was moderate. The model detected two factors which affect HUMRISK globally. The socio-economic factor (DDI) showed negative influence on HUMRISK, thus less developed districts represented higher risk for HAE. Forest cover proportion of the district (FOR) also prevailed globally but with a positive influence, whereas the more the forest cover caused the higher the risk of HAE.

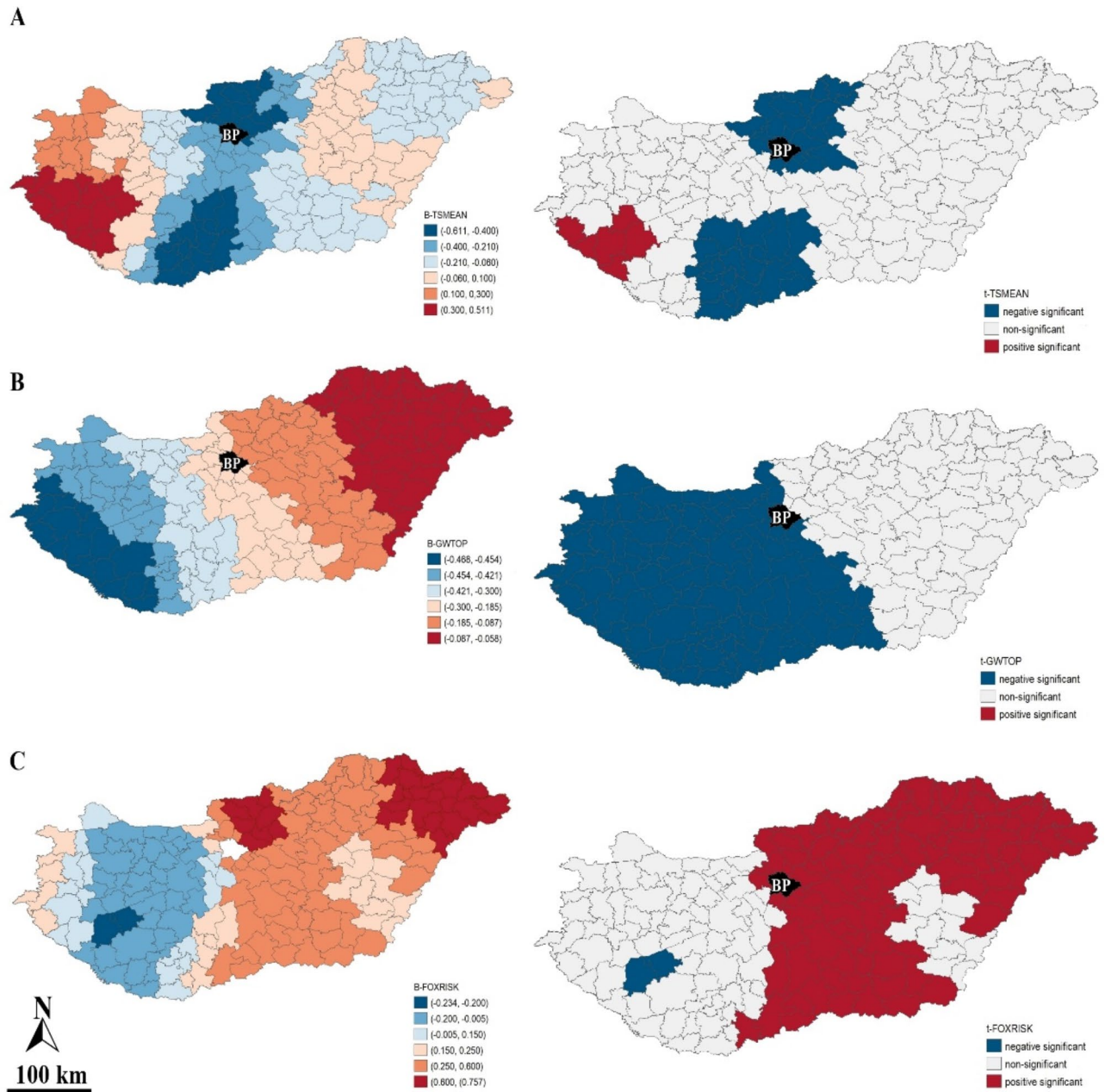
Those variables, which were involved as previously proven drivers of *E. multilocularis* endemics, such as earth skin temperature, ground surface wetness, and infection prevalence in red foxes<sup>2,16</sup>, showed various impacts on HAE by regions. In the northern part of the country, which was reached by the parasite for the first time<sup>14</sup>, both the soil temperature (TSMEAN) and the smoothed prevalence in foxes (FOXRIK) behave very similar as it was detected within the historical focus of *E. multilocularis*<sup>16</sup>. On the other hand, the moisture content of the ground did not show a significant impact on the smoothed period prevalence of HAE (HUMRISK) here.

The Northern Hungarian Mountains are under a strong influence of continental climate, which is modified by the microclimatic effect of the relief. This complex climatic impact causes harsh winters but less hot and dry summers in the mountains than in the surrounding plain areas.

Although the North Hungarian Mountains have a more humid climate than of the Great Plain, both showed the same pattern in correlation between ground surface wetness and HAE risk. The bandwidth for the humidity of the soil surface (GWTOP) proved large, thus this variable prevails in large areas. Besides the direct linkage between the two regions, a remarkable west-to-east change to a more continental climate can explain the similar GWTOP effect of the North Hungarian Mountains and the Great Plain. The controversial relationship between soil surface humidity and HAE prevalence apparently questioned the previously proven impact of humidity on HAE risk. Nevertheless, it is worth noting that our study investigated district level values of different variables, therefore it was not suitable to detect microfoci within the concerned district. However, these small habitat spots can be the keys in the transmission of *E. multilocularis*<sup>17</sup>.

The South Transdanubian Region of Hungary is affected by sub-Mediterranean climatic impact<sup>18</sup>. Hot summers and mild winters are very different from the climatic features of the historic focus of *E. multilocularis*<sup>19</sup>. In spite of this difference, a remarkable case accumulation could be observed in this area, especially during the last five years. Moreover, neither the comprehensive surveillance that was carried out in the red fox population of the region could provide an explanation for the phenomenon<sup>20</sup>. In the western part of this area both the skin surface temperature and soil humidity showed a negative effect on smoothed period prevalence of HAE, thus the warmer the temperature and the lower the humidity resulted in the higher the risk of HAE. Moreover, on a spot inside this area, a HAE case accumulation in humans accompanied by a low prevalence in red foxes.

It is less probable that the *E. multilocularis* population in the southern part of Hungary differs this rate in its environmental demands from the northern populations. Though, genetic studies proved that the Hungarian *E. multilocularis* population has a higher genetic diversity than of Slovakian or Czech populations, the origin of all Central Eastern European populations are proved to be the historical focus in the Alps<sup>21,22</sup>.



**Fig. 6.** Local effects of TSMEAN, GWTOP and FOXRISK on HUMRISK, (A)  $\beta$ -TSMEAN = regression coefficient of TSMEAN, (B) t-TSMEAN = t-value of TSMEAN, (C)  $\beta$ -GWTOP = regression coefficient of GWTOP, (D) t-GWTOP = t-value of GWTOP, (E)  $\beta$ -FOXRISK = regression coefficient of FOXRISK, (F) t-FOXRISK = t-value of FOXRISK (note: the black color with BP indicates the excluded districts of Budapest) (software used for creation: GeoDA version 1.22. and PhotoScape version 3.7).

A possible explanation for the parasite's apparent preference for unsuitable district level environmental conditions could be a compensatory mechanism by other factors. These factors can also prevail in microfoci, in which the optimal conditions of survival and transmission exist in spite of the generally existing suboptimal climatic conditions. These factors can be detected by micro-epidemiological investigations, which can disclose local risk factors<sup>10,17</sup>. Another reason for the parasite's success in less appropriate conditions is the presence of regionally prevailing factors, which compensate for the risk mitigation effect of the suboptimal climate.

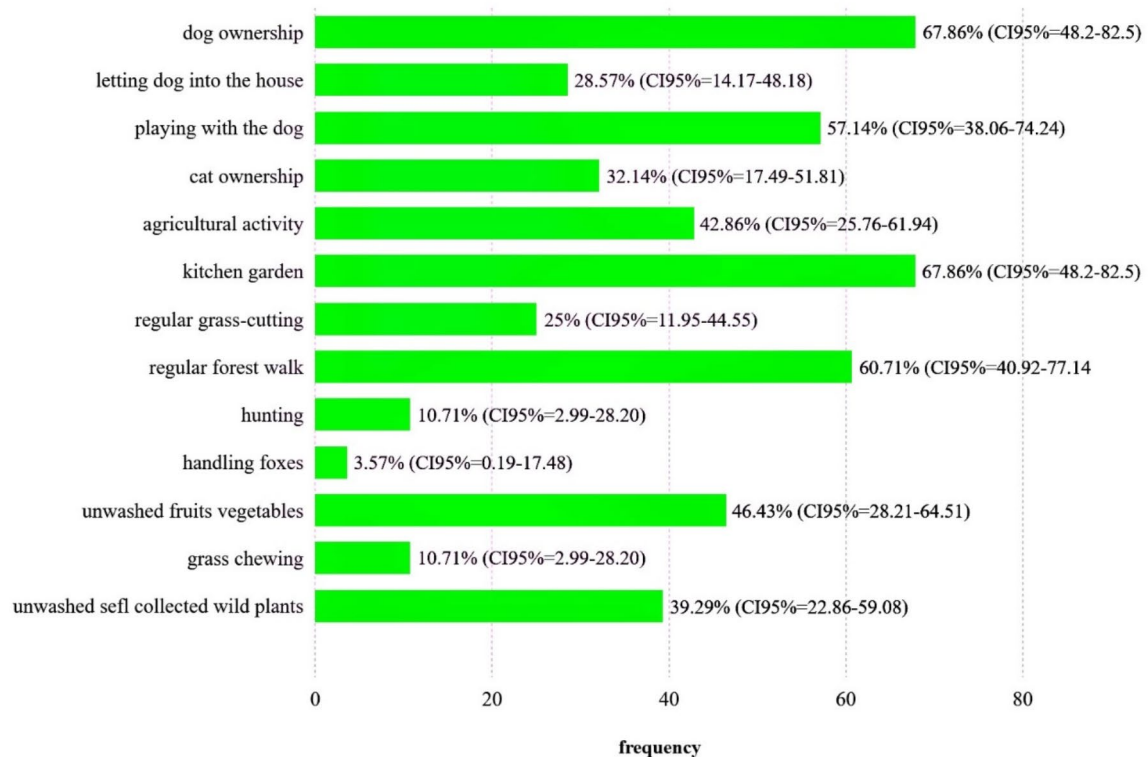
In South Transdanubia, this risk raising factor can be the composition of the predator community in this region, which is characterized by a large density of golden jackal (*Canis aureus*) populations. This wild canine can carry *E. multilocularis* and can significantly increase the risk of HAE in smaller geographical habitats. In South Transdanubia, the stable presence of golden jackals modifies the risk caused by foxes. These two mesocarnivores strongly compete for sources, thus resulting in increased occurrence of foxes around human settlements because



Source	Data (applied factor)	Period	References
Humanitarian data exchange	Shapefile of Hungarian districts (N/A*)	2015–2022	51
Hungarian central statistical office	Census (human population size)	2011	52
Copernicus land monitoring service	Land cover categories URB = urban, industrial areas and transport infrastructures AGRO = arable lands, permanent crops, grasslands, and heterogenous agricultural areas FOR = forests WETHABIT = wetlands and water bodies	2018	53
NASA prediction of worldwide energy resources	Meteorological data TSMEAN = mean earth skin temperature GWTOP = surface soil wetness	1992–2022	54
Institute for economic and enterprise research Hungarian chamber of commerce and industry	DDI = district development index (calculated from 22 socio-economic indicators)	2017	50
National food chain safety office, Budapest, Hungary	<i>E. multilocularis</i> infection in fox (sample size, number of infected foxes)	Hunting seasons of 2008 and 2012 years	20

**Table 4.** The involved data sources utilized in the retrospective analysis of Hungarian HAE (2003–2024). \*Not applicable.

#### risk factor



**Fig. 7.** Bar chart on risk factors of HAE in Hungary in the period 2003–2024 (note CI95% indicates 95% confidence interval of the counted frequency) using the software Scimago Graphica version 1.0.42.

they can tolerate humans better than golden jackals do<sup>23,24</sup>. Besides competition and spatial distribution of these two canids, dissimilar host suitability might enhance each other's role in disease transmission<sup>1,25</sup>.

The other, district level, factors that could affect the prevalence of HAE are DDI and FOR, which have the same impact in both hyperendemic focuses of HAE. It is curious that both the socio-economic factor (DDI) and forest coverage (FOR) level had homologous effects in each investigated district, though with opposite signs. Between DDI and FOR variables, the statistical analysis could not identify multicollinearity, therefore both factors have their own impact on HAE risk in humans.

In the case of forest coverage, other research resulted in various conclusions. Most studies concluded that mature forest stands with high biodiversity provide less suitable habitat for *E. multilocularis* spread than

meadows or shrublands, which are generally agreed as the most hazardous habitats for maintenance of *E. multilocularis*<sup>2,5,26</sup>. In our research, the applied data source did not distinguish between different types of forests, e.g. shrublands or woody pastures, thus the exact risk of different sylvatic environments could not be identified.

In Hungary, extended grasslands are characteristic especially for the Great Plain<sup>27</sup>, which proved quite the less risky area of the country. On the other hand, smaller patches of grasslands are very important natural habitats in both the Northern Mountains and in Transdanubia<sup>28</sup>. In Hungary, the proportion of true grassland area decreased for two reasons. On the one hand, human demands for area, such as urbanization and intensification of agriculture, reduced the area of grasslands. On the other hand, the end of extensive animal farming resulted in reduced pasturing and mowing activity, which led to afforestation of pastures and meadows<sup>27</sup>. For this reason, we presumed that a high proportion of forests at a district level meant a higher proportion of natural habitats compared to urban areas and large-scale arable lands. Our finding that forest coverage was identified as a general risk of HAE agreed with the early findings within the historical focus, thus mainly rural environments carried high infection risk<sup>2</sup>.

For infection with HAE, a higher risk of low-income regions and low education level was investigated in China<sup>3,29</sup> and Kyrgyzstan<sup>4</sup>. All studies emphasize the importance of individual knowledge on prevention. In our study, the education level groups of the patients were exactly equal, whereas one third of the patients had low, intermediate, and higher education levels, respectively. Therefore, the individual's educational level apparently did not affect the chance of infection. However, the socio-economic environment seemed to have a strong impact on the patients' risk for HAE acquisition. A district's development index depends on a list of factors, which determine access to information, healthcare, and services. The central role of an appropriately functioning and achievable healthcare system in prevention and early diagnosis could be well illustrated in Kyrgyzstan studies. After the end of the Soviet Union, the collapse of the local healthcare system led to the dramatic increase of both cystic echinococcosis, which is caused by *E. granulosus*, and HAE cases<sup>4</sup>.

The questionnaire survey was based solely on the interrogation of diseased people. Therefore, in the lack of a control group, it was not suitable to determine relative risk of different potential risk factors but frequency of occurrence of these factors in cases. The questions involved in the survey were adapted from a previously performed study<sup>30</sup>. Our survey revealed that 66.67% of HAE patients engaged in kitchen gardening. Thus, this lifestyle element was the most frequently chosen one among the respondents. This finding agrees with the experiences of a case-control study completed in Germany<sup>31</sup>.

A European large-scale study on food self-provisioning revealed that the central motivation for gardening is financial as the poorest social groups produced a part of their food demand<sup>32</sup>. The questionnaire did not ask directly about the financial background of the respondents, therefore gardening as a socio-economic indicator of limited income could not be proven by this study. Moreover, less developed districts are characterized by countryside abandonment all over Europe<sup>33</sup>. This phenomenon creates plenty of appropriate, undisturbed shelters and feeding sites for mesocarnivores, such as red foxes and golden jackals, in the surroundings and inside the human settlements. Less sufficient waste management, which is also common in poor regions, provides a huge number of anthropogenic sources for survival and successful breeding of wildlife<sup>34</sup>. In densely populated fox habitats, even lower prevalence of *E. multilocularis* can generate a higher risk as it could be observed in Switzerland between urbanised foxes<sup>2</sup>.

The presence of golden jackals in the less developed districts in South Transdanubia highlights an anomaly, which could be experienced in the relationship between the soil surface temperature and HAE prevalence of a district in the eastern part of the region. In spite of very similar characteristics to the western part of the region, the eastern half seemed free from HAE. Therefore, only spatial rate smoothing suggested suspicion of risk eastward in South Transdanubia. However, the complete lack of cases in the southernmost districts of the east reversed the positive association between soil temperature and HAE risk, which was identified in the west.

The anomalous districts are part of Ormánság (Supplementary Information 1), one of the poorest regions in Hungary. The average inhabitants are characterised by low education and consequential low income, without vision and ability of self-care<sup>35</sup>. Considering these facts and the biotic and climatic similarity of this region with the neighboring HAE focus, it cannot be excluded that the apparent freedom of this area is due to the locals' inattention. During the disease course, initially, the clinical signs of HAE are lacking or less severe<sup>36,37</sup>, therefore most patients do not suspect a life-threatening disease. In addition, disease prevention and health screening tests do not fit the culture of poverty, which characterize the small villages of Ormánság<sup>35</sup>. Our lack of confidence in the disease-free status of Ormánság was supported by the study of Balen Topic et al.<sup>38</sup>, which detected a HAE case accumulation in Croatia, not far from the southern border of Ormánság. The contradictory finding about the southeastern part of Transdanubia highlighted the paramount importance of knowledge dissemination not only in known hyperendemic focuses but areas with uncertain status.

In this study, we investigated the interdependence of potential influencing factors and *E. multilocularis* infection in humans and foxes. For statistical analysis, we applied geospatial methods to determine the risk of HAE and its background at a district level. As a result, we revealed that two, very different HAE foci have been localized in the country, yet. However, the driving forces of HAE in the two foci proved remarkably different. The Northern Hungarian Mountain focus was found to be similar to the historic focus of *E. multilocularis* in the Alps<sup>19,26</sup>. This observation agrees with the studies implemented in Czechia and Slovakia where *E. multilocularis* prevalences in foxes<sup>39,40</sup> and in humans<sup>16,41</sup> are proved to be high. In the northern focus, cool soil temperature and higher parasite prevalence in red foxes seemed to increase the risk of human infection at district level. Curiously, humidity as a risk factor could not be justified, which contradicted the previous study results that confirmed the humidity dependence of *E. multilocularis*<sup>40,42</sup>.

The South Transdanubian focus showed a very different risk pattern. Warm and dry soil surfaces appeared to have higher HAE risk than a humid and cool environment. Based on this apparent contradiction to known environmental demand of *E. multilocularis*<sup>19,26</sup>, we concluded that other, not investigated factors enhanced

the disease transmission within this area, in spite of the suboptimal climatic features. This observation draws attention to the significance of micro-epidemiological investigations, which can identify microfoci and their risk factors contributing to HAE transmission even in seemingly less appropriate conditions<sup>10,17,42</sup>.

Our study recognized an area in the country, which behaved in the model as an anomaly. This area is one of the poorest zones of Hungary. The climatic and biotic characteristics of this area is very similar to the neighboring South Transdanubian focus and a Croatian area with HAE accumulation<sup>10,38,42</sup>. In spite of that, HAE has never been detected here. We concluded that without comprehensive knowledge dissemination and health screening of the inhabitants, this anomaly cannot be evaluated correctly, and the status of this area remains uncertain.

Our study possesses limitations, which mostly originated from the spatially aggregated distribution of the few confirmed HAE cases. However, the geospatial analysis of the known cases denoted two hyperendemic foci of the disease with apparently free areas around them. Based on this finding and the data of the surrounding countries, we deem that the recently seen data are incomplete and cannot describe the true situation. In Slovakia<sup>16</sup> and in Croatia<sup>38</sup>, 137 and 6 HAE cases were detected, resulting in an annual incidence of 0.187/10<sup>5</sup> and a prevalence of 4.91/10<sup>5</sup>, respectively. Both recently determined hyperendemic foci are located adjacent to similar HAE hot spots of the neighboring countries. This geographical vicinity suggests a cross-border connectivity of the endemic areas. Therefore, the diagnostic sensitivity of the healthcare service of Hungary needs improvement. The apparently increasing annual incidence, especially in the last five years, may also indicate that the early inefficiency of the diagnostic system is in a slow transition towards enhancement. The rapidly changing diagnostic efficiency cannot make sense of calculating annual incidence of HAE, which could suggest an intensively worsening epidemiological situation. Though we cannot exclude even this possibility, in this stage of the research, we hypothesize that raising disease awareness of healthcare professionals caused the apparent increase of annual incidence of HAE.

Another shortcoming of the study was that only 28 patients could take part in the questionnaire survey. Moreover, a control group was not interrogated parallelly with the case group. For this reason, the results of the survey could highlight potential risk factors without determination of relative risk. However, the factors that were chosen most frequently by the respondents of our study, such as gardening and dog ownership, are those which were determined by a previous case control study as the most important risk factors<sup>31</sup>.

## Conclusion

This study highlighted the weight of multidisciplinary approach to *E. multilocularis*. Our investigation revealed some biological, climatic and socio-economic factors of HAE. On the other hand, the analysis resulted in contradictory findings, which need micro-epidemiological research to clarify. This study aimed to determine the most important district level risk factors for human infection with *E. multilocularis*, therefore it concentrated on districts and could not analyze microfoci. These small natural spots are very important in disease transmission<sup>10,17</sup> but are too small for large-scale analysis. However, the results of large-scale risk analysis could point at those areas, which need more detailed investigation to determine true risk of HAE.

Though the number of confirmed HAE cases in Hungary is very low, district level socio-economic development and the proportion of forested areas could be determined as risk factors by spatial analysis. Due to the rarity of the disease, the questionnaire survey of the cases was not suitable for statistical analysis. However, the survey results could call attention to the potential individual risk factors, like kitchen gardening, dog ownership, and outdoor activity in the forest areas. Based on the findings of this study, we concluded that determination of district level risk factors could be a very important element for decision makers in mapping out an efficient disease control strategy, which should rest on extensive knowledge dissemination for both the public and the health care professionals of the high-risk areas.

## Methods

### Data sources

This retrospective study enrolled 40 patients who were diagnosed with alveolar echinococcosis between 9 April 2003 and 12 April 2024. The Hungarian patients included in this study were those who had positive serology for *Echinococcus multilocularis* infection with one highly sensitive (ELISA or IHA) and one highly specific (Western blot) test and fulfilled the clinical and laboratory diagnostic criteria for probable or confirmed AE patient as proposed by the WHO-IWGE. The patients were diagnosed with AE in the framework of the healthcare system. The researchers evaluated all suspect cases based on the available diagnostic findings. A probable case was defined as any patient with typical organ lesions detected by imaging methods (CT/MRI/ultrasound), and positive serology for AE with two positive tests. A confirmed case was defined as the above, plus (1) histopathology compatible with AE and/ or (2) detection of *E. multilocularis* genetic material in a clinical specimen. Patients who were serologically negative or non-specific positive but *E. multilocularis* infection was unequivocally confirmed by histopathological methods were also included<sup>7,36,37</sup>.

As an indicator of the probability of acquiring *E. multilocularis* infection during a given period<sup>43</sup>, we calculated the period prevalence (PP) of both HAE and fox echinococcosis at the country and district levels. Regarding human infections, the index was expressed to 100,000 citizens and covered the interval between 2003 and 2024, and the case fatality rate was also counted. For calculation of the human population at risk, the population size was obtained from the 2011 census data. This census was the first that provided district level data. The district ('járás' in Hungarian) system was introduced into the Hungarian administration in 2012. Since the year 2011 is in the middle of the studied period, and its result on population size is close to both the average and the median of other two censuses on the boundaries of this period, we used 2011 census data. The calculation of PP in the fox populations was based on the results of a previously shown study<sup>20</sup>, in which 1612 animals were investigated during the hunting seasons of 2008–2009 and 2012–2013. The foxes were hunted in



the framework of the approved hunting management plans. The carcasses were submitted to the Veterinary Diagnostic Directorate, National Food-chain Safety Office to survey the *E. multilocularis* spatial distribution in Hungary<sup>20</sup>.

By application of period prevalence instead of disease incidence, we followed the recommendations of Bhopal, Thrusfield, and Spronk et al.<sup>43–45</sup>. We determined the number of human cases as the sum of the cases at the beginning of the period and new cases that occurred during the period independently of the outcome of the disease. The human population at risk was determined as the average population size that can characterize the period as Bhopal<sup>44</sup> suggests. Since the population of Hungary decreased linearly between the 2001 and 2022 censuses, we deemed that the population size measured in 2011 represented well the average population size for the study period.

In foxes, similarly proper population data cannot be achieved. For this reason, we used the sample size of the 2012–2013 national *E. multilocularis* screening programme to determine the population at risk. The number of cases were established by counting all carcasses, in which the presence of *E. multilocularis* was confirmed in the laboratory. We calculated the PP index by the following equations.

In the case of human population:

$$PP = \left( \frac{NIP}{PR} \right)$$

where PP = period prevalence between 2003 and 2024, NIP = number of infected persons during the study period (2003–2024), PR = size of population at risk based on 2011 census data in particular spatial units (country or district).

In the fox population:

$$PP = \left( \frac{NIF}{NP} \right)$$

where PP = period prevalence for 2008–2009 and 2012–2013 winter months, NIF = number of infected foxes in the sample, NP = number of population (sample size in the *E. multilocularis* screening campaign during 2008–2009 and 2012–2013 winter) in particular spatial units (country or district).

Land cover categories were obtained from the CLC2018 database with a minimum mapping unit of 25 hectares (<https://land.copernicus.eu/en/products/corine-land-cover>). We classified the surface characteristics into six variables: URB (including urban and industrial areas and transport infrastructures), AGRO (including arable lands, permanent crops, grasslands, and heterogenous agricultural areas), FOR (including forests), and WETHABIT (including wetlands and water bodies) (Supplementary Information 4). Two climatic variables were chosen in our analysis. The mean earth skin temperature (TSMEAN) and the surface soil wetness (GWTOP) were adopted. These climatic components are well-known and substantial drivers in the viability of the infective *E. multilocularis* eggs in the environment<sup>19,42,46,47</sup>.

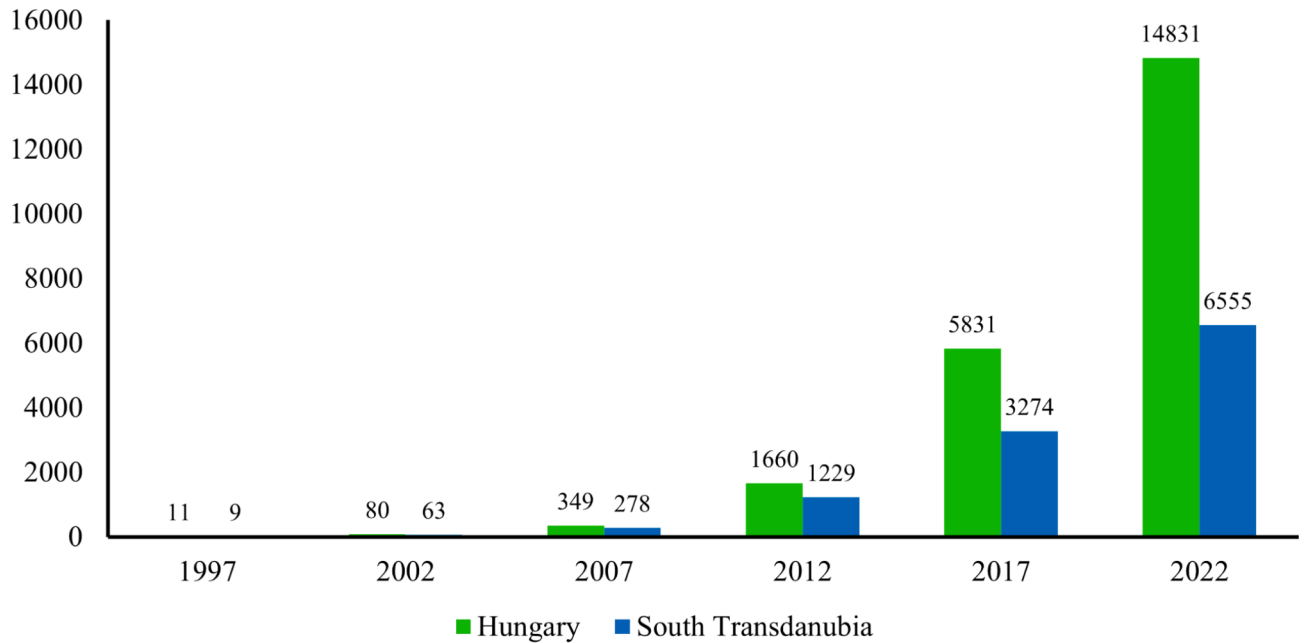
Not only abiotic and biotic environmental factors may play a role in the spread of zoonotic diseases, but also the socio-economic status of the affected human population. Strong evidence supported that some socio-economic drivers (e.g. economic disparity, higher unemployment rate, poor living conditions) could also potentially provide a higher risk of infectious diseases, including HAE<sup>3,48,49</sup>. For this reason, we implicated a socio-economic level variable (district development index = DDI). The data obtained from the Institute for Economic and Enterprise Research Hungarian Chamber of Commerce and Industry<sup>50</sup> and calculated from 21 different indicators (Supplementary Information 5). All the datasets that were applied in this study are shown in Table 4.

In the lack of district level data, we could not count with the population size of the concurrent golden jackal. However, the effect of golden jackal on the maintenance of *E. multilocularis* in the Hungarian ecosystems has already arisen<sup>55</sup>. Nowadays, the presence of golden jackals can be detected all over Hungary but the southwestern part of the country was the place where this mesopredator entered the country during the 1990s. Based on the hunting bag data, South Transdanubia still possesses the densest golden jackal population of Hungary (Fig. 8).

### Data preparation and selection

The Hungarian districts (Local Administrative Units level 1) were chosen as the units of analysis (<https://ec.europa.eu/eurostat/web/nuts/local-administrative-units>).

There are currently 197 district-level administrative units in Hungary, of which 23 represent Budapest districts. Although two cases of HAE were reported in the Budapest area, the 23 city districts were excluded from the analysis. Fox data were available in only seven districts (14 animals in total); accordingly, the proportion of the missing data was not tolerable. The urban ecosystem, which characterizes the vast majority of Budapest has a very homogenous pattern as huge continuous areas belong to URB land cover category. Thus, the analysis could have given distorted results for the city. Six of the remaining 174 countryside districts had neither fox data. Since sufficient data were available on these districts' neighbors, we used the neighbors' mean imputation method<sup>56</sup>, applying the generally accepted 5% replacement limit<sup>57</sup>.



**Fig. 8.** Trend of the golden jackal's hunting bag for Hungary as a whole and the South Transdanubian region. (Data accepted from the National Game Management Database: <http://www.ova.info.hu/index-en.html>).

We assigned each case to the administrative district of the patient's residence if the patient had lived there for at least ten years. If the patient had been living in their current residence for less than ten years, he or she had been assigned to the district, in which he or she had lived for ten years previously<sup>31</sup>.

As a first step, we determined the period prevalence of the infection using the total number of cases and the number of inhabitants in each district. For mapping the results, the calculated ratio was expressed to 100,000 people. In the case of foxes, the prevalence in each district was given as the proportion of infected individuals among all animals tested for *E. multilocularis*. These crude proportions, calculated from small areas (districts) and small populations (humans and foxes) of different sizes, provide unstable and imprecise estimations. The estimation accuracy is linked to the size of the population at risk, which can cause a high variance in results (i.e. districts with a larger population size have lower variance than districts with a smaller population size). For this reason, adjusting the crude rates could improve the precision of the estimates. This stabilization can be achieved by spatial rate smoothing, which was implemented by the GeoDa software (v1.22). The fundamental principle of smoothing is that the precision of the crude rate can be improved by borrowing strength from other observations. In this study, the software completed the calculated PP data of a concerned district with the data of its neighborhood by spatial rate smoothing. This method estimated the smoothed risk by locally weighing the number of cases (numerator) and the population at risk (denominator) separately, rather than applying a weighted average for the rate itself. This method could eliminate the shortcomings that originated from the high variance of both case numbers and population sizes by districts<sup>15</sup>. The estimation of locally weighted rate by partial rate smoothing proved to be appropriate to define high-risk areas for a rare disease<sup>58</sup>. The calculated rate smoothing yield of human cases (HUMRISK) was used as a dependent variable, while the spatial rate smoothing yield of fox infection (FOXRISK) applied as independent variables (Supplementary Information 5).

### Spatial cluster analysis

During the spatial cluster analysis, we used the original raw data. Districts where  $PP > 0$  was observed in either humans or foxes were considered infected units. However, we evaluated them as non-infected when the calculation gave  $PP = 0$  as a result.

The scan statistic was performed using SaTScan software (v9.6.1)<sup>59</sup>. The null hypothesis of spatial scan statistics was that observed HAE cases are randomly distributed across the districts. A log-likelihood ratio (LLR) statistics were calculated using the numbers of observed and expected HAE cases within and outside the clusters. The null hypothesis was rejected if the LLR statistics differed significantly between the clusters and the outside areas.

We applied in both cases a Bernoulli-based model with Gini coefficient to assess the true human and fox clusters' pattern. The maximum size of the clusters was determined to be a maximum 50% of the population. We did not allow the geographical overlap between the clusters and boosted the statistical power of our analysis with Monte Carlo iterations ( $n = 999$ ).

Clusters with a  $p$ -value  $< 0.05$  were considered statistically significant.

### Global and local regression models

The strong correlation between some predictors can weaken the power of coefficients and make the p-value vague. To avoid this problem, before the model building, we checked the multicollinearity between the explanatory variables, and only explanatory variables with a variance inflation factor (VIF) value of less than five were accepted<sup>60</sup>.

In the first step of the spatial regression analysis, we used the conventionally accepted ordinary least squares (OLS) approach to evaluate the global (stationary) distribution of HUMRISK. This model evaluates the effects of explanatory variables in a global sense using the following equation:

$Y_i = \beta_0 + \sum \beta_k X_{ik} + \varepsilon_i$  where  $Y_i$  = the dependent variable at the  $i$ th sample point,  $\beta_0$  = the estimated parameter (coefficient) of the linear regression equation,  $\beta_k$  = the estimated parameter (coefficient) of the  $k$ th independent variable,  $X_{ik}$  = the  $k$ th independent variable at the  $i$ th sample point,  $\varepsilon_i$  = the random error.

The global effect through space may be unrealistic<sup>61</sup>. Some explanatory variables do not fluctuate from sample point to sample point, but many are space-dependent (e.g. mean temperature, land cover categories). If stationary and non-stationary relationships can be observed, the assumptions of OLS may misestimate the coefficients. This problem could be eliminated by local regression models, such as geographically weighted regression or multiscale geographically weighted regression since these models are capable of solving the problem of spatial heterogeneity<sup>62</sup>. The GWR model is an ameliorated traditional linear regression model with the power to apply a spatial weight matrix. The model determines a fixed bandwidth (neighborhood range) because it is supposed that the independent variables alter in the same spatial scale. The regression coefficients of spatial units are analyzed separately and could give a finer explanation of spatial effects. The GWR model is based on the following equation:

$$Y_i = \beta_0(u_i, v_i) + \sum \beta_k(u_i, v_i) X_{ik} + \varepsilon_i$$

where  $Y_i$  = the dependent variable,  $\beta_0(u_i, v_i)$  = the estimated parameter (regression coefficient) of the linear regression equation at  $(u_i, v_i)$ ,  $\beta_k(u_i, v_i)$  = the estimated parameter (regression coefficient) of the  $k$ th independent variable at  $(u_i, v_i)$ ,  $X_{ik}$  = the  $k$ th independent variable at the  $i$ th sample point,  $\varepsilon_i$  = the random error.

The MGWR model is an extended GWR model. In this model, the different explanatory variables have diverse bandwidths, so each independent variable has its own bandwidth in the MGWR models compared to the GWR. This scheme could interpret the non-stationary action of different independent variables better, giving an improved model estimation<sup>62</sup>. The model is given as follows:

$$Y_i = \beta_0(u_i, v_i) + \sum \beta_{bw \cdot k}(u_i, v_i) X_{ik} + \varepsilon_i$$

where  $Y_i$  = the dependent variable,  $\beta_0(u_i, v_i)$  = the estimated parameter (regression coefficient) of the linear regression equation at  $(u_i, v_i)$ ,  $\beta_{bw \cdot k}(u_i, v_i)$  = the differential bandwidth of the  $k$ th independent variable at  $(u_i, v_i)$ ,  $X_{ik}$  = the  $k$ th independent variable at the  $i$ th sample point,  $\varepsilon_i$  = the random error.

Four indices were used to evaluate the different models. In the first instance, we compared the residual sum of squares (RSS), the Akaike Information Criterion (AIC), and the adjusted coefficient of the model (adj.  $R^2$ ). The lower value of RSS and AIC, while a higher adj.  $R^2$  points out a better model fit. The calculation was performed using MGWR (v2.2) software with the following settings: projected coordinate type, adaptive spatial kernel, golden selection bandwidth search and Monte Carlo test with 1000 iterations.

In the case of RSS, we used Moran's I Z-value to estimate the spatial autocorrelation. If the RSS has significant positive autocorrelation, the standard errors of the parameter estimates are not accurate enough<sup>61</sup>. GeoDa software (v1.22) was used to map the results and detect spatial autocorrelation.

### Questionnaire

To discover the potential personal risk factors, an interview was conducted with 28 of the 40 patients. The interrogation was performed when the case was involved in the recent study. Therefore, patients of the earliest cases were questioned years after the confirmation. We could not reach the total number of patients to complete the questionnaire because several went unnoticed by the health care system. The patients were interviewed concerning their gender, age, residence, education level, pet ownership, outdoor activity regarding the potential environmental exposure, and travel history in the historical *E. multilocularis* countries (Supplementary Information 2). The majority of the potential risk factors to analyze were chosen from a previously conducted large scale European study supported by EFSA<sup>30</sup>. The proportion of answers to the questions was determined during the evaluation. Sterne's exact method was used to calculate the frequencies and their 95% confidence interval. The computation was carried out using the online version of the Quantitative Parasitology software<sup>63</sup>. If the percentage of answers to a question exceeded 50%, it was considered to be a potential primary risk factor. Patients participating in the questionnaire analysis were asked for informed consent and obtained informed consent before completing it.



## Data availability

The personal questionnaire data are not public. The datasets and supplementary information generated and analysed during the current study are available from the following link: <https://zenodo.org/records/13763482>. Balázs Dezsényi and Gábor Nagy should be contacted if other questions arise.

Received: 14 September 2024; Accepted: 11 December 2024

Published online: 28 December 2024

## References

- Romig, T. & Wassermann, M. *Echinococcus* species in wildlife. *Int. J. Parasitol. Parasites Wildl.* **23**, 100913. <https://doi.org/10.1016/j.jppaw.2024.100913> (2024).
- Romig, T. Epidemiology of echinococcosis. *Langenbecks Arch. Surg.* **388**, 209–217 (2003).
- Conraths, F. J. et al. Potential risk factors associated with human alveolar echinococcosis: Systematic review and meta-analysis. *PLoS Negl. Trop. Dis.* **11**, e0005801. <https://doi.org/10.1371/journal.pntd.0005801> (2017).
- Casulli, A. Recognising the substantial burden of neglected pandemics cystic and alveolar echinococcosis. *Lancet Glob Health.* **8**, e470–e471 (2020).
- Giraudoux, P. et al. Long-term retrospective assessment of a transmission hotspot for human alveolar echinococcosis in mid-west China. *PLoS Negl. Trop. Dis.* **13**, e0007701. <https://doi.org/10.1371/journal.pntd.0007701> (2019).
- Horváth, A. et al. The first case of human alveolar echinococcosis in Hungary (in Hungarian). *Orv Hetil.* **149**, 795–799 (2008).
- Dezsényi, B. et al. Emerging human alveolar echinococcosis in Hungary (2003–2018): a retrospective case series analysis from a multi-centre study. *BMC Infect. Dis.* **21**, 168. <https://doi.org/10.1186/s12879-021-05859-5> (2021).
- Anselin, L. How (not) to lie with spatial statistics. *Am. J. Prev. Med.* **30**, 3–6. <https://doi.org/10.1016/j.amepre.2005.09.015> (2006).
- Guerra, D., Hegglin, D., Bacciarini, L., Schnyder, M. & Deplazes, P. Stability of the southern European border of *Echinococcus multilocularis* in the Alps: evidence that *Microtus arvalis* is a limiting factor. *Parasitology* **141**, 1593–1602 (2014).
- Moloi, S. et al. Global and local drivers of *Echinococcus multilocularis* infection in the western Balkan region. *Sci. Rep.* **13**. <https://doi.org/10.1038/s41598-023-46632-9> (2023).
- Souris, M. *Epidemiology and geography: Principles, methods and tools of spatial analysis* 9–16 (ISTE Ltd and John Wiley & Sons, Inc, 2019).
- Oshan, T. M., Li, Z., Kang, W., Wolf, L. J. & Fotheringham, A. S. mgwr: A Python implementation of Multiscale Geographically Weighted Regression for investigating process spatial heterogeneity and scale. *ISPRS Int. J. Geo-Inf.* **8**, 269. <https://doi.org/10.3390/ijgi8060269> (2019).
- Ma, Z. & Fan, H. Influential factors of tuberculosis in mainland China based on MGWR model. *PLoS One.* **18**, e0290978. <https://doi.org/10.1371/journal.pone.0290978> (2023).
- Sréter, T., Széll, Z., Egyed, Z. & Varga, I. *Echinococcus multilocularis*: an emerging pathogen in Hungary and Central Eastern Europe? *Emerg. Infect. Dis.* **9**, 384–386 (2003).
- Anselin, L., Lozano-Gracia, N. & Koschinsky, J. Rate transformations and smoothing. Technical Report. [https://www.researchgate.net/publication/249913160\\_Rate\\_Transformations\\_and\\_Smoothing](https://www.researchgate.net/publication/249913160_Rate_Transformations_and_Smoothing) (2006).
- Antolová, D. et al. Human alveolar echinococcosis in Slovakia: Epidemiology and genetic diversity of *Echinococcus multilocularis*, 2000–2023. *PLoS Negl. Trop. Dis.* **18**, e0011876. <https://doi.org/10.1371/journal.pntd.0011876> (2024).
- Eckert, J. & Deplazes, P. Alveolar echinococcosis in humans: the current situation in Central Europe and the need for countermeasures. *Parasitol. Today.* **15**, 315–319 (1999).
- Mezősi, G. Physical geography of the Transdanubian Hills in The physical geography of Hungary (ed Mezősi, G.) 269–281 (Springer, Cham, (2017).
- Cenni, L. et al. Current and future distribution of a parasite with complex life cycle under global change scenarios: *Echinococcus multilocularis* in Europe. *Glob Chang. Biol.* **29**, 2436–2449 (2023).
- Tolnai, Z., Széll, Z. & Sréter, T. Environmental determinants of the spatial distribution of *Echinococcus multilocularis* in Hungary. *Vet. Parasitol.* **198**, 292–297 (2013).
- Casulli, A., Széll, Z., Pozio, E. & Sréter, T. Spatial distribution and genetic diversity of *Echinococcus multilocularis* in Hungary. *Vet. Parasitol.* **174**, 241–246 (2010).
- Uhang, G. et al. Unravelling the genetic diversity and relatedness of *Echinococcus multilocularis* isolates in Eurasia using the EmsB microsatellite nuclear marker. *Infect. Genet. Evol.* **92**. <https://doi.org/10.1016/j.meegid.2021.104863> (2021).
- Tsunoda, H. Niche overlaps and partitioning between Eurasian golden jackal *Canis aureus* and sympatric red fox *Vulpes vulpes*. *Proc. Zool. Soc.* **75**, 143–151 (2022).
- Torretta, E. et al. Niche partitioning between sympatric wild canids: the case of the golden jackal (*Canis aureus*) and the red fox (*Vulpes vulpes*) in north-eastern Italy. *BMC Ecol. Evol.* **21**, 129. <https://doi.org/10.1186/s12862-021-01860-3> (2021).
- Kapel, C. M., Torgerson, P. R., Thompson, R. C. & Deplazes, P. Reproductive potential of *Echinococcus multilocularis* in experimentally infected foxes, dogs, raccoon dogs and cats. *Int. J. Parasitol.* **36**, 79–86 (2006).
- Fischer, I. et al. Distribution of alveolar echinococcosis according to environmental and geographical factors in Germany, 1992–2018. *Acta Trop.* **212**, 105654. <https://doi.org/10.1016/j.actatropica.2020.105654> (2020).
- Kovács, K. F., Iváncsik, V., Boromisza, Z. & Valánszki, I. Spatial trends of grassland changes based on Hungarian local studies after 1990 with a macro-regional perspective. *Eur. Countrys.* **14**, 397–419 (2022).
- Illyés, E., Bauer, N. & Botta-Dukát, Z. Classification of semi-dry grassland vegetation in Hungary. *Preslia* **81**, 239–260 (2009).
- Di, X. et al. How climate, landscape, and economic changes increase the exposure of *Echinococcus* spp. *BMC Public Health.* **22**. <https://doi.org/10.1186/s12889-022-14803-4> (2022).
- EFSA AHAW Panel (EFSA Panel on Animal Health and Welfare). Scientific opinion on *Echinococcus multilocularis* infection in animals. *EFSA J.* **13**, 4373. <https://doi.org/10.2903/j.efsa.2015.4373> (2015).
- Kern, P. et al. Risk factors for alveolar echinococcosis in humans. *Emerg. Infect. Dis.* **10**, 2088–2093 (2004).
- Vávra, J. et al. Food self-provisioning in Europe: an exploration of sociodemographic factors in five regions. *Rural Sociol.* **83**, 431–461 (2018).
- Janus, J., Božek, P., Taszakowski, J. & Doroz, A. Decaying villages in the centre of Europe with no population decline: Long-term analysis using historical aerial images and remote sensing data. *Habit Int.* **121**, 102520. <https://doi.org/10.1016/j.habitatint.2022.102520> (2022).
- Bino, G. et al. Abrupt spatial and numerical responses of overabundant foxes to a reduction in anthropogenic resources. *J. Appl. Ecol.* **47**, 1262–1271 (2010).
- Hornýák, M. & Ragadics, T. The possibilities of text mining in the examination of the local society of the Ormánság region. *Market Menedzsmet.* **51**, 61–71 (2017).
- Brunetti, E., Kern, P. & Vuitton, D. A. Writing panel for the WHO-IWGE. Expert consensus for the diagnosis and treatment of cystic and alveolar echinococcosis in humans. *Acta Trop.* **114**, 1–16. <https://doi.org/10.1016/j.actatropica.2009.11.001> (2010).
- Reinehr, M. et al. Pathology of echinococcosis: A morphologic and immunohistochemical study on 138 specimens with focus on the differential diagnosis between cystic and alveolar echinococcosis. *Am. J. Surg. Pathol.* **44**, 43–54 (2020).

38. Balen Topić, M. et al. Emergence of *Echinococcus multilocularis* in Central continental Croatia: A human case series and update on prevalence in foxes. *Life (Basel)*. **13**, 1402. <https://doi.org/10.3390/life13061402> (2023).
39. Miterpáková, M., Dubinsky, P., Reiterová, K. & Stanko, M. Climate and environmental factors influencing *Echinococcus multilocularis* occurrence in the Slovak Republic. *Ann. Agri Environ. Med.* **13**, 235–242 (2006).
40. Oksanen, A. et al. The geographical distribution and prevalence of *Echinococcus multilocularis* in animals in the European Union and adjacent countries: a systematic review and meta-analysis. *Parasit. Vectors*. **9** (519). <https://doi.org/10.1186/s13071-016-1746-4> (2016).
41. Kolářová, L. et al. Human alveolar echinococcosis, Czech Republic, 2007–2014. *Emerg. Infect. Dis.* **21**, 2263 (2015).
42. Simoncini, A. & Massolo, A. Multiscale ecological drivers of *Echinococcus multilocularis* spatial distribution in wild hosts: A systematic review. *Food Waterborne Parasitol.* **34**, e00216. <https://doi.org/10.1016/j.fawpar.2023.e00216> (2023).
43. Thrusfield, M. Describing disease occurrence. In *Veterinary epidemiology*, fourth ed Wiley,., (ed. Thrusfield, M) 58–85 (2018).
44. Bhopal, R. S. The concept of risk and fundamental measures of disease frequency: Incidence and prevalence. In *Concepts of Epidemiology: Integrating the ideas, theories, principles, and methods of epidemiology*, 3 ed., (ed. Bhopal, R. S.) 201–234 (Oxford, 2016).
45. Spronk, I. et al. Calculating incidence rates and prevalence proportions: not as simple as it seems. *BMC Public Health*. **19**, 1–9. <https://doi.org/10.1186/s12889-019-6820-3> (2019).
46. Veit, P. et al. Influence of environmental factors on the infectivity of *Echinococcus multilocularis* eggs. *Parasitology* **110**, 79–86 (1995).
47. Ma, T. et al. Geographical detector-based influence factors analysis for Echinococcosis prevalence in Tibet, China. *PLoS Negl. Trop. Dis.* **15**, e0009547. <https://doi.org/10.1371/journal.pntd.0009547> (2021).
48. Possenti, A. et al. Potential risk factors associated with human cystic echinococcosis: systematic review and meta-analysis. *PLoS Negl. Trop. Dis.* **10**, 0005114. <https://doi.org/10.1371/journal.pntd.0005114> (2016).
49. Ayorinde, A. et al. Health inequalities in infectious diseases: a systematic overview of reviews. *BMJ Open*. **13**, e067429. <https://doi.org/10.1136/bmjopen-2022-067429> (2023).
50. IEER (Institute for Economic and Enterprise Research Hungarian Chamber of Commerce and Industry). *Developing and lagging districts – 2017*. <https://gvi.hu/research-details/608/developing-and-lagging-districts-2017> (2019).
51. Humanitarian Data Exchange. *Subnational Administrative Boundaries*. <https://data.humdata.org/dataset/cod-ab-hun> (2024).
52. Hungarian Central Statistical Office. *Census database* <https://nepszamlalas2022.ksh.hu/en/database/> (2011).
53. Copernicus Land Monitoring Service. *CORINE Land Cover*. <https://land.copernicus.eu/en/products/corine-land-cover> (2018).
54. NASA (National Aeronautics and Space Administration). *Prediction Of Worldwide Energy Resources*. <https://power.larc.nasa.gov/data-access-viewer/> (2024).
55. Bijl, H., Schally, G., Márton, M., Heltai, M. & Csányi, S. From invaders to residents: The golden jackal (*Canis aureus*) expansion in Hungary since the mid-1990s. *PLoS ONE*. **19**, e0306489. <https://doi.org/10.1371/journal.pone.0306489> (2024).
56. Baker, J., White, N. & Mengersen, K. Missing in space: an evaluation of imputation methods for missing data in spatial analysis of risk factors for type II diabetes. *Int. J. Health Geogr.* **13**, 47 (2014).
57. Schafer, J. L. Multiple imputation: a primer. *Stat. Methods Med. Res.* **8**, 3–15 (1999).
58. Tan, L. et al. (ed, M.) Spatial analysis of human and livestock anthrax in Dien Bien province, Vietnam (2010–2019) and the significance of anthrax vaccination in livestock. *PLoS Negl. Trop. Dis.* **16** e0010942 <https://doi.org/10.1371/journal.pntd.0010942> (2022).
59. Kulldorff, M. A spatial scan statistic. *Commun. Stat. Theory Methods*. **26**, 1481–1496 (1997).
60. Kim, J. H. Multicollinearity and misleading statistical results. *Korean J. Anesthesiol.* **72**, 558–569 (2019).
61. Fotheringham, A. S., Yue, H. & Li, Z. Examining the influences of air quality in China's cities using multi-scale geographically weighted regression. *Trans. GIS*. **23**, 1444–1464 (2019).
62. Yang, T. C., Shoff, C., Choi, S. E. & Sun, F. Multiscale dimensions of county-level disparities in opioid use disorder rates among older Medicare beneficiaries. *Front. Public Health*. **10**, 993507. <https://doi.org/10.3389/fpubh.2022.993507> (2022).
63. Marozzi, R. J., Fabian, M., Rozsa, L. & I. & Biostatistics for parasitologists – a primer to Quantitative Parasitology. *Trends Parasitol.* **35**, 277–281 (2019).

## Acknowledgements

The study was supported by the Flagship Research Groups Programme of the Hungarian University of Agriculture and Life Sciences.

## Author contributions

All authors contributed to the study's conception and design. B.D. and E.Cs. implemented the conceptualization, supervision, manuscript reviewing and final editing. T.S., J.D. and Zs.K. performed the data collection, curation, and project administration. G.N. and N.S. designed and carried out the analytical strategy and interpretation of the findings. Á.Cs. wrote and edited the manuscript.

## Funding

The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

## Declarations

## Competing interests

The authors declare no competing interests.

## Ethics approval and consent to participate

This study was performed in line with the principles of the Declaration of Helsinki. Approval was granted by the Committee for Scientific and Research Ethics of the Medical Research Council (07-03-2024/BM/2130-3/2024). Patients participating in the questionnaire analysis were asked for informed consent before completing it. Their data was used in this study after their permission.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-024-83119-7>

[0.1038/s41598-024-83119-7](https://doi.org/10.1038/s41598-024-83119-7).

**Correspondence** and requests for materials should be addressed to G.N.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2024