Received: 23 October 2023 | Accepted: 24 November 2023

DOI: 10.1111/hiv.13601

SHORT COMMUNICATION

Areas with high HIV prevalence: A spatial analysis of nationwide claims data in Germany

M. K. Akmatov[1] | E. Hu[1] | R. Rüsenberg[2] | C. Kollan[3] | D. Schmidt[3] | C. Kohring[1] | J. Holstiege[1] | M. Bickel[4] | J. Bätzing[1]

[1]Department of Epidemiology and Health Care Atlas, Central Research Institute of Ambulatory Health Care (Zi), Berlin, Germany

[2]German Association of Physicians in HIV-Care (dagnä e.V.), Berlin, Germany

[3]Department of Infectious Disease Epidemiology, Robert Koch Institute (RKI), Berlin, Germany

[4]Infektiologikum, Frankfurt am Main, Germany

Correspondence
M. K. Akmatov, Department of Epidemiology and Health Care Atlas, Central Research Institute of Ambulatory Health Care, Salzufer 8, 10587, Berlin, Germany.
Email: makmatov@zi.de

Abstract

Objective: We aimed to identify spatial clusters of high HIV prevalence in Germany.

Methods: Using nationwide outpatient claims data comprising information of about 88% of the total German population ($N = 72\,041\,683$), we examined spatial variations and spatial clusters of high HIV prevalence at the district level ($N = 401$). People with HIV were identified using the *International Statistical Classification of Diseases and Related Health Problems, Tenth Revision* (ICD-10 codes) B20, B22, and B24 (HIV disease) documented as 'confirmed'.

Results: Among 72 041 683 people with statutory health insurance in Germany in 2021, 72 636 had diagnosed HIV, which corresponds to a prevalence of 101 per 100 000 individuals (0.10%). Of these, 56 895 were males (78%). At a district level, the HIV prevalence varied by a factor of 32 between 13 in a rural district in Bavaria and 417 per 100 000 individuals in the German capital, Berlin. The spatial autocorrelation coefficient was 0.24 ($p < 0.0001$, Global Moran's I). Several high-prevalence spatial clusters of different sizes were identified, mostly located in western Germany. The largest cluster comprised eight districts in the southern part of Hesse, including the city of Frankfurt and the city of Mainz in Rhineland-Palatinate. The second cluster consisted of four districts in North Rhine-Westphalia, including the cities of Cologne and Düsseldorf. Two districts in southern Germany (Mannheim and Ludwigshafen) formed the third cluster. Only urban districts were observed in spatial clusters of high HIV prevalence.

Conclusions: The current study identified for the first time spatial clusters with high HIV prevalence in Germany. This understanding is of particular importance when planning the general and specialized medical care of patients with HIV and to support preventive measures.

KEYWORDS

Germany, HIV, prevalence, small-area variations, spatial clusters

INTRODUCTION

According to estimations from a mathematical modelling within the yearly national HIV reporting, an estimated 90 800 people (95% confidence interval [CI] 85 300–96 000) were living with HIV in Germany at the end of 2021 [1]. Of these, about 82 100 individuals (95% CI 76 900–87 200) were aware of their diagnosis, which corresponds to a prevalence of about 99 per 100 000 individuals. For the remaining 8600 individuals, HIV infection had not yet been diagnosed. Furthermore, the number of people living with HIV was estimated by sex, age, region of origin (e.g., foreign origin) and groups of increased risk of HIV by mode of HIV transmission or transmission risk groups (e.g., men who have sex with men [MSM] or intravenous drug users). For example, about 80% of people living with HIV were male, about 16% were of foreign origin, and 10% were intravenous drug users [1]. Both mathematical modelling studies and secondary analyses of routine data are available. The latter report the so-called administrative prevalence of HIV among individuals with statutory health insurance (SHI) [2] or private health insurance [3]. None of these studies examined variations in the regional distribution of people living with HIV. The above-mentioned mathematical modelling study estimated the number of people living with HIV in 16 German federal states, ranging between 320 people in Brandenburg (a federal state in eastern Germany) and 19 400 in North Rhine-Westphalia (a federal state in western Germany) [1]. Data on people living with HIV with a finer geographical resolution (e.g., at the district level) are not available in Germany. These data are of particular importance as they can support the planning of medical care of people with HIV. In addition, it is also important to know the local areas where the chance of being exposed to HIV could be increased in order to offer targeted preventive measures in these regions. The aim of the current study was therefore to examine the geographic variations in HIV prevalence at the district level and to identify spatial clusters with a high prevalence of HIV in Germany.

MATERIALS AND METHODS

Data and study population

We used the nationwide outpatient claims data collected in 2021 in accordance with §295 of the Social Code Book V. The dataset contains medical information for all individuals with SHI in Germany who visited SHI-authorized physicians at least once in 2021 ($N = 72\,041\,683$). Medical information includes diagnoses coded according to the German modification of the *International Statistical Classification of Diseases and Related Health Problems, Tenth Revision* (ICD-10-GM). In addition, information on sex, age (in years), and district of residence of the insurees is available.

Case ascertainment

Individuals with HIV were identified using the ICD-10 codes B20 (HIV disease resulting in infectious and parasitic diseases), B22 (HIV disease resulting in other specified diseases), and B24 (unspecified HIV disease) coded as 'confirmed' diagnoses.

Spatial analysis

We used the smallest geographical unit available in the dataset: administrative districts according to the administrative structure as of 31 December 2016 (401 districts, corresponding to Nomenclature of Territorial Units for Statistics [NUTS]-3 level). First, for each district, the prevalence of diagnosed HIV was calculated as the proportion of individuals with HIV among all insurees with contact with SHI-authorized physicians in the respective districts per 100 000 insurees. We did not observe any single district without individuals with HIV. In districts where fewer than 30 individuals had HIV, the direct calculation of HIV prevalence was not possible because of data protection regulations. In these districts, we set the number of people with HIV at 15. Furthermore, we applied the Global [4] and Local Moran's I [5] analysis to examine the spatial autocorrelation and spatial clusters, respectively. The Global Moran's I tests whether there is a tendency to clustering in the entire region (in our case, Germany). The autocorrelation values of this test are the same as with the other correlation statistics such as Pearson's coefficient r. Negative values indicate clusters with neighbouring districts that have larger differences in the parameters examined (e.g., low and high HIV prevalence). Positive values indicate neighbouring districts with similar (low or high) prevalences. Values around zero indicate random spatial distribution. In case of significant spatial autocorrelation, the Local Moran's I is used to detect the location and size of local clusters. Four cluster types can be identified: clusters with (1) high-high, (2) low-low, (3) high-low, and (4) low-high prevalence. The Holm–Bonferroni method was used to adjust for multiple testing problem. The statistical analysis was performed with the program R (version 1.0.4).

TABLE 1 Sex, age, and regional distribution of people with statutory health insurance (SHI), those diagnosed with HIV, and corresponding prevalence, 2021.

Characteristics	Total SHI population, *N* (%)	SHI population with HIV, *n* (%)	HIV prevalence per 100 000 individuals
Total	72 041 683 (100.0)	72 636 (100.0)	101
Sex			
Female	38 579 929 (53.6)	15 741 (21.7)	41
Male	33 461 754 (46.4)	56 895 (78.3)	170
Age group (years)			
0–4	3 713 758 (5.2)	286 (0.4)	8
5–9	3 265 265 (4.5)	103 (0.1)	3
10–14	2 984 020 (4.1)	133 (0.2)	4
15–19	3 221 044 (4.5)	245 (0.3)	8
20–24	3 874 721 (5.4)	798 (1.1)	21
25–29	4 267 173 (5.9)	2519 (3.5)	59
30–34	4 881 565 (6.8)	5518 (7.6)	113
35–39	4 614 532 (6.4)	7337 (10.1)	159
40–44	4 386 156 (6.1)	9007 (12.4)	205
45–49	4 071 248 (5.7)	8918 (12.3)	219
50–54	5 230 053 (7.3)	11 854 (16.3)	227
55–59	5 834 915 (8.1)	11 549 (15.9)	198
60–64	5 110 935 (7.1)	6915 (9.5)	135
65–69	4 202 929 (5.8)	3287 (4.5)	78
70–74	3 555 659 (4.9)	1829 (2.5)	51
75–79	2 788 192 (3.9)	1224 (1.7)	44
80–84	3 262 693 (4.5)	750 (1.0)	23
85–89	1 801 581 (2.5)	263 (0.4)	15
≥90	975 244 (1.4)	101 (0.1)	10
Federal states			
Baden-Württemberg	9 278 596 (12.9)	7514 (10.3)	81
Bavaria	11 177 498 (15.5)	9358 (12.9)	84
Berlin	3 167 185 (4.4)	13 215 (18.2)	417
Brandenburg	2 235 264 (3.1)	1368 (1.9)	61
Bremen	603 182 (0.8)	995 (1.4)	165
Hamburg	1 623 170 (2.3)	4383 (6.0)	270
Hesse	5 428 971 (7.5)	6670 (9.2)	123
Lower Saxony	7 027 764 (9.8)	4416 (6.1)	63
Mecklenburg-Western Pomerania	1 449 088 (2.0)	844 (1.2)	58
North Rhine-Westphalia	15 743 034 (21.9)	15 408 (21.2)	98
Rhineland-Palatinate	3 427 102 (4.8)	2645 (3.6)	77
Saarland	848 822 (1.2)	556 (0.8)	66
Saxony	3 649 964 (5.1)	2046 (2.8)	56
Saxony-Anhalt	1 997 063 (2.8)	805 (1.1)	40
Schleswig-Holstein	2 479 796 (3.4)	1711 (2.4)	69
Thuringia	1 905 184 (2.6)	702 (1.0)	37

RESULTS

We found a total of 72 636 individuals with HIV in 2021, which corresponds to a prevalence of 101 per 100 000 individuals (i.e., 0.10%). Of these, 56 895 were males (78%) and 15 741 were females (22%). At federal state level (NUTS-1), the largest number of people with diagnosed HIV was observed in North Rhine-Westphalia ($n = 15\,408$) and Berlin ($n = 13\,215$) (Table 1). At the district level (NUT-S3), the prevalence of diagnosed HIV varied between 13 and 417 per 100 000 individuals (Figure 1a). The highest prevalence was observed in big urban municipalities: Berlin (417), Frankfurt/Main (406), Cologne (389), Hamburg (270), Munich (266), and Stuttgart (257). Sparsely populated rural districts displayed the lowest prevalences. We found 59 districts with a prevalence of HIV ≥0.10% (Table S1). The spatial autocorrelation coefficient was 0.24 ($p < 0.0001$, Global Moran's I). Several spatial clusters of high HIV prevalence of different sizes were identified (Figure 1b). The largest cluster comprised eight districts from two federal states, including the city of Frankfurt/Main (Table S2). The second spatial cluster consisted of four districts, including the cities of Cologne and Düsseldorf. The third cluster, comprising two districts (Mannheim and Ludwigshafen) from two federal states, was observed in the southwest of Germany. All districts identified as high prevalence clusters were either big urban municipalities with a population of at least 100 000 inhabitants (nine districts) or urban districts (five districts, population density of at least 150 inhabitants per 1 km^2). Of note, Berlin, the district with the highest prevalence (417 per 100 000 individuals), was not identified as a high HIV prevalence cluster, as the neighbouring districts displayed a relatively lower prevalence of HIV.

DISCUSSION

Using the nationwide outpatient claims data containing information for about 88% of the total German population, we examined small-area variations and the presence of spatial clusters of high HIV prevalence in Germany. It is known that HIV is distributed unevenly across

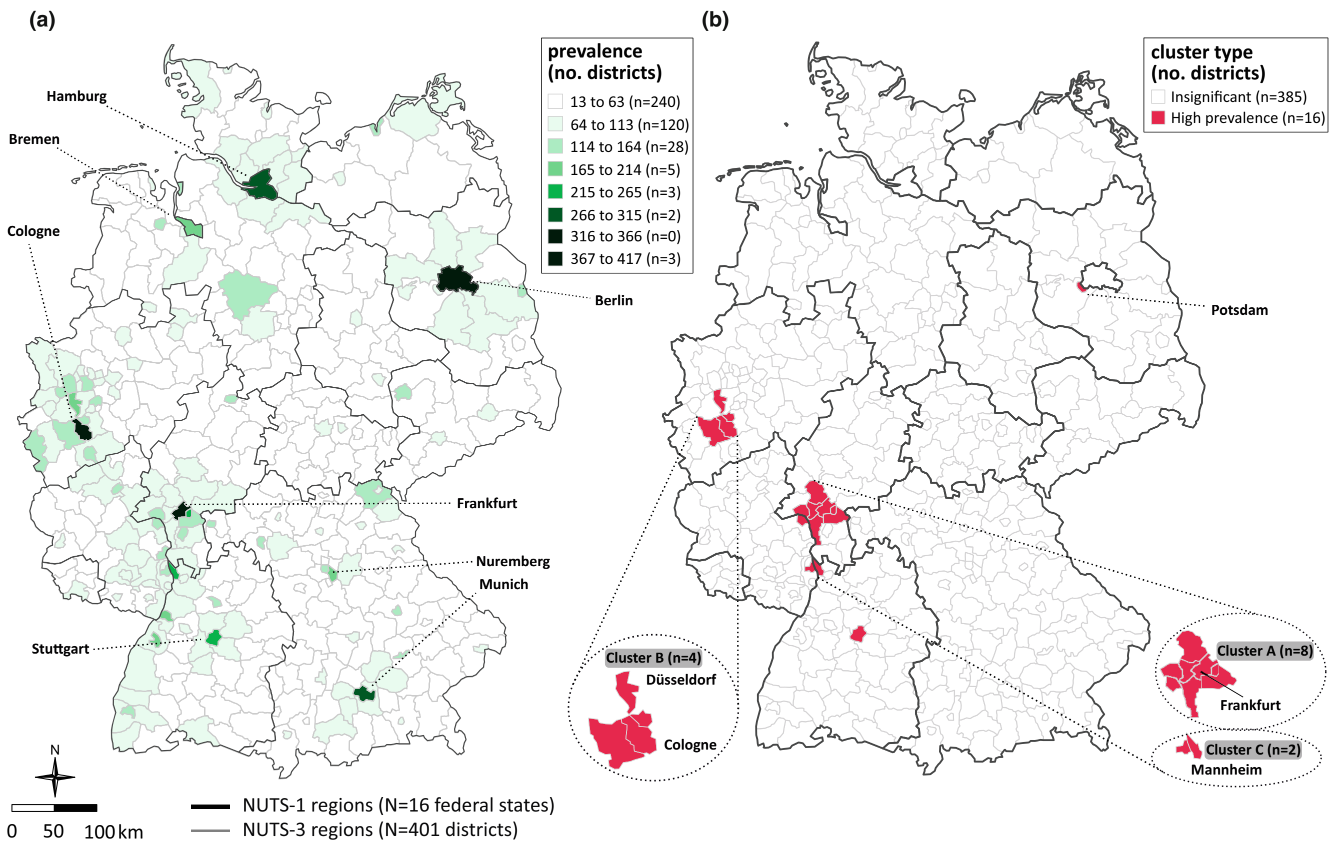

FIGURE 1 District-level prevalence of diagnosed HIV (a) and spatial clusters with a high prevalence (b) in Germany, 2021. In (a), an equal interval was used for group classification. In (b), districts with significant spatial clusters (red) were identified with Local Moran's I [5]. Only the cluster type 'high-high' prevalence was observed. None of the remaining districts showed a spatial pattern. Selected characteristics of districts in the identified clusters (i.e., A, B, and C) can be found in Table S2. NUTS, Nomenclature of Territorial Units for Statistics.

geographic regions [6]. In particular, the prevalence of HIV is higher in urban than in rural areas [6]. Furthermore, it is assumed to be higher in western than in eastern Germany [7]. To the best of our knowledge, small-area data of HIV prevalence in Germany are not yet available because of a lack of regional data on a small scale. In Germany, only estimations from mathematical models of the number of people with HIV in federal states are available [1]. Internationally, previous research on small-area variations in the prevalence of HIV have used small-area estimation techniques. Most of these studies were conducted in African countries [8–11]. One study from the USA estimated the number of people with HIV in 677 counties where HIV case numbers were not previously available [12]. They found that the prevalence of HIV was low in most of the counties. However, counties with a high HIV prevalence were also observed in this study [12]. The nationwide claims data we analysed containing information for ~72 million individuals with SHI yielded a unique opportunity to conduct a spatial analysis at the district level (NUTS-3). Of note, we also observed small case numbers (fewer than 30 people with HIV) in 50 of the 401 districts (12.5%). Since direct estimation of HIV prevalence was not possible in these districts because of data protection regulations, we set the HIV case number to 15 for these districts to calculate prevalence estimates. Thus, the prevalence in these districts may be less accurate than in districts with a higher number of people with HIV. However, it does not impact spatial analysis and identification of spatial clusters since district-level prevalence rather than individual-level data were used in this analysis.

We found small-area variations differing by a factor of 32 between 13 and 417 per 100 000 individuals. Not surprisingly, the highest prevalence was observed in the capital of Germany, Berlin (417), followed by further big urban municipalities such as Frankfurt/Main (406), Cologne (389), Hamburg (270), Munich (266), and Stuttgart (257). The lowest prevalence of HIV was observed in rural areas, particularly those with a low population density (data not shown). This finding could be explained by differences in the distribution of key populations at increased risk of HIV infection such as MSM or intravenous drug users [1]. The latter two groups are the most prevalent population groups among people with HIV in Germany (72% and 11% of all individuals of German ethnicity with HIV, respectively). Furthermore, using spatial analysis, we identified geographical clusters of districts with a relatively high HIV prevalence. This is of particular importance for planning and evaluating ambulatory care structures as well as preventive interventions and has never been done before for Germany.

A few limitations of the study should be mentioned. First, we performed a secondary data analysis of routinely collected medical claims data, which were collected for administrative purposes and not for epidemiological research. More specifically, an external validation of ICD-10 codes of about 72 600 patients with HIV is not possible. Thus, misclassification cannot be ruled out and may impact the validity of the study. In addition, the case definition was based on ICD-10 diagnoses only. Laboratory findings such as CD4 cell count were not part of our dataset. However, we found that both the total number of people with HIV and figures at the federal state level agree very well with the estimates from a mathematical modelling within the yearly national HIV reporting conducted by the Robert Koch Institute [1] (Figure S1), although these figures cannot be directly compared because of the differing methodological approaches applied. The very good agreement can be interpreted as a sign of indirect external validation. Second, because of the nature of the data used, we provide an administrative (also known as diagnostic) prevalence of HIV in the ambulatory setting, which cannot be interpreted as a 'true' epidemiological estimate. Third, the dataset does not include data for the remaining 12% of the German population, who are mostly privately insured. In addition, the dataset only contains data for insurees who visited an SHI-authorized physician in 2021.

The current study provided small-area estimates of HIV prevalence and identified spatial clusters of high HIV prevalence in Germany, showing the general population HIV prevalence across 401 districts for the first time. These data may serve as a basis for regionally targeted prevention and intervention measures as well as for planning of medical care for individuals with HIV.

AUTHOR CONTRIBUTIONS

Akmatov MK designed the study, analysed the data, and wrote the manuscript. Hu E extracted the data and critically reviewed the manuscript. Rüsenberg R, Kollan C, Schmidt D, Bickel M, and Bätzing J participated in designing the study and critically reviewed the manuscript. Kohring C and Holstiege J critically reviewed the manuscript. Bätzing J supervised the project. All authors read and approved the final version of the manuscript.

ACKNOWLEDGEMENTS

The authors thank the 17 regional Associations of Statutory Health Insurance Physicians (ASHIPs) in Germany for providing the datasets.

CONFLICT OF INTEREST STATEMENT

The authors reported no financial interests or potential conflicts of interest related to this study.

ETHICS STATEMENT

The use of claims data for scientific research is regulated by the Social Code Book (SGB V) in Germany. Ethical approval and informed consent were not required as this study used routinely collected anonymized data. The research was conducted in accordance with the Helsinki Declaration (in its current revised form: 64th WMA General Assembly, Fortaleza, Brazil, October 2013).

ORCID

M. K. Akmatov https://orcid.org/0000-0003-2310-3179
C. Kollan https://orcid.org/0000-0002-1483-5194
C. Kohring https://orcid.org/0000-0002-1865-1923
J. Bätzing https://orcid.org/0000-0001-7609-8559

REFERENCES

1. an der Heiden M, Marcus U, Kollan C, et al. Estimation of the number of HIV new infections in 2021 and of the total number of individuals living with HIV in Germany. *Epid Bull*. 2022;47:3-18.
2. Kurch-Bek D, Gallowitz C, Baumgarten A, Tenckhoff B, an der Heiden M. Prevalence of patients with HIV receiving outpatient medical care in Germany-an analysis based on secondary data. *Bundesgesundheitsbl*. 2020;63:1250-1261.
3. Finkenstädt V, Wild F. HIV infection in the private insurance. Prevalence, incidence and pharmacological care. Available at: https://www.wip-pkv.de/fileadmin/DATEN/Dokumente/Veroeffentlichungen/HIV-Infektion_in_der_PKV.pdf
4. Moran P. The interpretation of statistical maps. *J R Stat Soc Series B Stat Methodol*. 1948;10:243-251.
5. Anselin L. Local indicators of spatial association - LISA. *Geogr Anal*. 1995;27:93-115.
6. Hamouda O, Marcus U, Voß L, Kollan C. Epidemiology of HIV infections in Germany. *Bundesgesundheitsbl*. 2007;50:399-411.
7. an der Heiden M, Marcus U, Kollan C, et al. HIV/AIDS: Key data and trends for Germany and the federal states. Available at: www.rki.de/DE/Content/InfAZ/H/HIVAIDS/Eckdaten/Eckdaten.html [Accessed 5 September 2023]
8. Eaton JW, Dwyer-Lindgren L, Gutreuter S, et al. Naomi: a new modelling tool for estimating HIV epidemic indicators at the district level in sub-Saharan Africa. *J Int AIDS Soc*. 2021;24:e25788.
9. Ouma J, Jeffery C, Awor CA, et al. Model-based small area estimation methods and precise district-level HIV prevalence estimates in Uganda. *PloS One*. 2021;16:e0253375.
10. O'Brien-Carelli C, Steuben K, Stafford KA, et al. Mapping HIV prevalence in Nigeria using small area estimates to develop a targeted HIV intervention strategy. *PloS One*. 2022;17:e0268892.
11. Saldarriaga EM. HIV-prevalence mapping using small area estimation in Kenya, Tanzania, and Mozambique at the first sub-National Level. *Ann Glob Health*. 2021;87:93.
12. Khan SS, McLain AC, Olatosi BA, Torres ME, Eberth JM. Small area estimation of county-level U.S. HIV-prevalent cases. *Ann Epidemiol*. 2020;48:30-35.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Akmatov MK, Hu E, Rüsenberg R, et al. Areas with high HIV prevalence: A spatial analysis of nationwide claims data in Germany. *HIV Med*. 2024;25(4):498-503. doi:10.1111/hiv.13601