

# Prevalence of HIV in people with potential HIV-indicator conditions in Germany: an analysis of data from statutory health insurances



Frederik Valbert,<sup>a,\*</sup> Georg M. N. Behrens,<sup>b,i</sup> Markus Bickel,<sup>d</sup> Christoph Boesecke,<sup>c,e</sup> Stefan Esser,<sup>f</sup> Patrik Dröge,<sup>g</sup> Thomas Ruhnke,<sup>g</sup> Amrei Krings,<sup>h</sup> Daniel Schmidt,<sup>h</sup> Uwe Koppe,<sup>h</sup> Barbara Gunsenheimer-Bartmeyer,<sup>h</sup> Lea Wienholt,<sup>a</sup> Jürgen Wasem,<sup>a</sup> and Anja Neumann<sup>a</sup>



<sup>a</sup>Institute for Healthcare Management and Research, University of Duisburg-Essen, Thea-Leymann-Str. 9, Essen, Germany

<sup>b</sup>Department of Rheumatology and Immunology, Hannover Medical School, Carl-Neuberg-Str. 1, Hanover, Germany

<sup>c</sup>German Centre for Infection Research (DZIF), Venusberg-Campus 1, Bonn, Germany

<sup>d</sup>Infektiologikum Frankfurt, Stresemannallee 3, Frankfurt, Germany

<sup>e</sup>Department of Medicine I, Bonn University Hospital, Venusberg-Campus 1, Bonn, Germany

<sup>f</sup>Department of Dermatology and Venerology, University Hospital Essen, University Duisburg-Essen, Hufelandstrasse 55, Essen, Germany

<sup>g</sup>AOK Research Institute (WIdO), Rosenthaler Strasse 31, Berlin, Germany

<sup>h</sup>Department of Infectious Disease Epidemiology, Robert Koch Institute, Seestrasse 10, Berlin, Germany

<sup>i</sup>German Centre for Infection Research (DZIF), Partner Site Braunschweig-Hannover, Carl-Neuberg-Str. 1, Hannover, Germany

## Summary

**Background** In care of people living with human immunodeficiency virus (HIV), early diagnosis of infection is one of the greatest challenges remaining. A promising approach to increase early diagnosis could be optimized HIV testing in persons with indicator conditions (ICs). ICs are conditions which are AIDS-defining in people living with HIV, conditions that may have significant adverse consequences for the individual's clinical management if the presence of HIV infection is not detected, and conditions with an (undiagnosed) HIV prevalence of  $\geq 0.1\%$ .

**Methods** In this cohort study, anonymous routine healthcare data of German statutory health insurances from 07/01/2016 to 06/30/2021 based on insured persons with an ICD-10-based diagnosis of selected ICs were analyzed. In a primary analysis, two stratifications (gender and age), and four sensitivity analyses HIV prevalence/incidence were calculated for persons with at least one of 26 IC described in international literature. This study is registered in the German Clinical Trials Register (identifier: DRKS00028743).

**Findings** Routine healthcare data from 513,509 insured persons were selected for analysis. In the primary analysis, only in malignant neoplasm of bronchus and lung a HIV prevalence was observed with a 95%-CI  $< 0.1\%$ . ICs with particularly high HIV prevalence were pneumocystosis (40.33%), oral hairy leukoplakia (36.71%), and Kaposi's sarcoma (29.86%). When stratified by gender, it was observed that in female patients, the 95%-CI of HIV prevalence fell below 0.1% for seven ICs. No such effect was observed in male patients. Stratified by age, among patients aged 30 to  $< 60$  years, the 95%-CI of HIV prevalence were always  $\geq 0.1\%$ , while in the other groups the 95%-CI fell below 0.1% for several ICs.

**Interpretation** In samples of patients with ICs in Germany, HIV prevalences/incidences were found to be  $\geq 0.1\%$  for all ICs except malignant neoplasm of bronchus and lung. This confirms the classification of these conditions as ICs for the German context and emphasizes the importance of HIV testing in these populations.

**Funding** This analysis is part of the HIV testing recommendations in guidelines and practice study (German title of the study: "HIV-Testempfehlungen in Leitlinien und Praxis"; acronym: HeLP), which is funded by the German Federal Joint Committee as part of the Innovationsfonds program to further develop the German healthcare system (funding number 01VSF21050).

**Copyright** © 2024 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

**Keywords:** HIV; Human immunodeficiency virus; AIDS; Acquired immune deficiency syndrome; Indicator conditions

eClinicalMedicine  
2024;73: 102694

Published Online 24 June  
2024  
<https://doi.org/10.1016/j.eclinm.2024.102694>

\*Corresponding author.

E-mail address: [frederik.valbert@medman.uni-due.de](mailto:frederik.valbert@medman.uni-due.de) (F. Valbert).

### Research in context

#### Evidence before this study

Human immunodeficiency virus (HIV) indicator conditions (ICs) have been proclaimed for several years as an opportunity for effective HIV testing. Yet, many potential HIV indicator conditions remain rarely investigated. To review the current published evidence, a systematic literature search was conducted in Pubmed (search terms: (HIV[Title/Abstract] OR HIV[MeSH Terms] OR Aids[Title/Abstract] OR HIV infection [Title/Abstract] OR "HIV infection"[Title/Abstract] OR "human immunodeficiency syndrome" [Title/Abstract] OR "acquired immunodeficiency syndrome"[Title/Abstract] OR human immunodeficiency\* virus[Title/Abstract]) AND (indicator condition[Title/Abstract] OR indicator conditions[Title/Abstract] OR indicator disease[Title/Abstract] OR indicator diseases[Title/Abstract])) and Embase (search terms: ('human immunodeficiency virus':ab,ti OR 'human immunodeficiency virus infection':ab,ti OR 'human immunodeficiency virus'/exp OR 'human immunodeficiency virus' OR 'human immunodeficiency syndrome':ab,ti OR 'acquired immune deficiency syndrome':ab,ti) AND ('indicator condition':ab,ti OR 'indicator conditions':ab,ti OR 'indicator disease':ab,ti OR 'indicator diseases':ab,ti) AND ([english]/lim OR [german]/lim)) in August 2022. The only filter used was "language of publication in English or German". After full-text screening, twelve publications reporting HIV prevalences/incidences in HIV indicator conditions could be included. One of these twelve publications was an erratum to another included publication. The remaining studies included both a review and original research articles. For example, the numbers of indicator conditions examined, the time windows within which an indicator condition is associated with an HIV infection, the age limits for inclusion, the regions of the studies, the numbers of cases reached and, where applicable, the groupings for subgroup analyses vary widely between studies. In consequence, there were some major differences in

the reported prevalence/incidence of HIV in patients with indicator diseases. None of the studies focussed on Germany, although Germany was sometimes considered together with other countries. In addition, expert opinions from clinical and epidemiological experts of importance in German HIV care were taken into account during planning and conduction of the study.

#### Added value of this study

The analysis of statutory health insurance data reported here, provides results that confirm or falsify the classification of health conditions as HIV indicator conditions for German care on the basis of a high number of cases. In addition, as shown by the few publications addressing HIV prevalence/incidence in indicator conditions found, the evidence on HIV prevalence in indicator conditions could be expanded by this study.

#### Implications of all the available evidence

This Germany-specific analysis may be used to raise awareness of the need to offer regular HIV testing for people with confirmed indicator conditions in Germany. In addition, in the further course of the study, German guidelines of confirmed indicator diseases will be examined with regard to HIV testing recommendations. Internationally, findings from this study can also strengthen programs on indicator condition-guided HIV testing through expanded evidence. By adding to the sparse evidence available (often from studies with low number of cases), the presented results can contribute to a higher reliability of summarized results. Moreover, in the presented analysis, different sensitivity analyses and stratifications were calculated, allowing a deeper insight into HIV prevalences/incidences in HIV indicator conditions. Overall, the results of the HIV testing recommendations in guidelines and practice study (HeLP) have the potential to reduce the rate of late HIV diagnoses.

### Introduction

A major concern of care for people living with human immunodeficiency virus (HIV) in Germany, such as reported by other Western countries as well, is the late diagnosis of HIV infection.<sup>1-5</sup> This is reflected in German rates of undiagnosed infections that are above the 5% targeted in the 95-95-95 goals of the World Health Organization/The Joint United Nations Programme on HIV/AIDS (at the end of 2021, around 90,800 people in Germany were living with HIV, of which around 8600 were unaware of their infection).<sup>1,6</sup> In addition, many reports show a high proportion (about 50%) of HIV late diagnosis in Germany and other western countries.<sup>4,7-9</sup> According to the outdated consensus definition, a late diagnosis meant that a person presented for HIV care for the first time with a CD4 count below 350 cells/ $\mu$ L or with an acquired

immune deficiency syndrome (AIDS)-defining event.<sup>10</sup> There is an updated definition, but due to its appearance in the second half of 2022, it is not yet used in many published studies.<sup>11</sup> Because delayed HIV diagnosis is associated with increased morbidity and mortality, more transmissions, and higher costs to the health care system, improvement is needed from a medical, epidemiological, ethical, and health economic perspective.<sup>12-16</sup> Most people living with HIV in Germany are men who have sex with men.<sup>1</sup> However, focusing on this group only would miss a relevant proportion of diagnoses. In addition, sexuality can be a sensitive topic in medical consultations. Therefore, approaches are needed to provide adequate HIV testing independent of sexual behavior.

One potential but still insufficiently used way to optimize HIV testing (and consequently facilitate an

earlier diagnosis) is to consistently offer HIV testing in case of diagnosed HIV indicator condition (IC).<sup>17–24</sup> ICs are conditions which are (1) AIDS-defining among people living with HIV, (2) conditions which may have significant adverse implications for the individual's clinical management in case the presence of HIV infection is not identified, and (3) conditions with an (undiagnosed) HIV prevalence of >0.1% or ≥0.1%.<sup>17,19,20,22,25–33</sup> Published studies differ in the classification of conditions with a prevalence of exactly 0.1% and in the question how strictly "undiagnosed" is considered.<sup>17,19,20,22,25–33</sup> For Germany, specific evidence regarding HIV prevalences/incidences in potential ICs is missing so far. In addition, international evidence is still rare. The prevalence/incidence of HIV in some ICs has only been examined in a few studies with sometimes inconsistent results between these studies, and the reported number of cases is often low.<sup>17,22,25–33</sup>

Health insurance is mandatory in Germany. There is a choice (depending on factors such as income and previous insurance periods) between numerous statutory and private health insurance funds. At around 89%, the majority of the population is covered by statutory health insurance.<sup>34</sup> Due to the extensive range of benefits, these insurers hold a large amount of health data. This data includes, for example, International Statistical Classification of Diseases and Related Health Problems (Version 10) (ICD-10) diagnoses and billing codes from office-based general practitioners, specialists, and hospitals, and prescribed drugs. Assignment to these ICD-10 codes, which are primarily for billing purposes, is based on the subjective assessment of the coding physician. In addition, unless otherwise specified in the employment contract, employees in Germany are required by law to obtain a medical certificate if they are incapable to work for more than three days. A version of the certificate, including the diagnosis, is sent by the insured person to his or her health insurance fund, so these diagnoses are also included. This data, originally collected for billing purposes – known as routine data – have developed into an established real world data base for research.<sup>35,36</sup>

The HIV testing recommendations in guidelines and practice study (German title of the study: "HIV-Tesempfehlungen in Leitlinien und Praxis"; acronym: HeLP) uses a mixed methods approach to investigate different aspects of IC-guided HIV testing in Germany. This includes an analysis of the frequency of HIV in people with internationally described ICs in Germany, an analysis of relevant medical guidelines, expert interviews with authors/editors of such guidelines focusing on why HIV testing recommendations are (not) included in the guidelines, and a written survey of office-based physicians regarding HIV testing in routine care. The study will then develop and publish strategies to strengthen IC-guided HIV testing.<sup>37</sup> The objective of this presented module of the HeLP study was to answer

the following two research questions: Can international ICs currently published in the literature be confirmed (regarding the prevalence/incidence of HIV) for the German context based on German statutory health insurance data? What is the prevalence of HIV stratified by gender and age for potential ICs in Germany?

## Methods

### Study design and data source

The cohort study presented here is one of the quantitative modules of the HeLP study.<sup>37</sup> It is a retrospective analysis of real-world data (routine data from statutory health insurance funds) without trial-specific interventions.

The selection of the ICs considered here was carried out in a two-stage procedure. First, potential ICs were identified with a systematic literature search. The results were discussed with clinical and epidemiological experts in German HIV care in order to classify their relevance for the national care situation. The next step was to assign ICD-10 codes to all selected diseases.

Accordingly, a list of 48 ICD-10 codes (see [Table 3](#)) was compiled for which samples of people insured with a statutory health insurance were drawn up at the AOK Research Institute. The routine data (07/01/2016–06/30/2021) provided by the AOK Research Institute were derived from statutory health insurance data of eleven regional (covering every region in Germany) AOK insurances covering about 27 million patients in total. Enrolment in the AOK is open to any inhabitant regardless of region, profession, income, age, or health status. All persons who cumulatively fulfilled the following characteristics were eligible for inclusion:

- Persons who had at least one of the selected ICD-10 codes documented as a diagnosis in inpatient or outpatient hospital data or in billing data of outpatient medical care in the date period 07/01/2016–06/30/2021;
- Persons who were at least 18 years of age (reference date: 07/01/2016);
- Persons with continuous statutory health insurance coverage (maximum of five days of absence per quarter until the end of the data period/death to avoid missed diagnoses, since analyses are based on statutory health insurance data).

The first date of the data period was chosen as a reference date, so all people in the data were at least 18 years of age (age of majority in Germany). As there were very few people who met these criteria, persons older than 95 years (reference date: 07/01/2016) or documented with the gender marker "unspecific" or "diverse" were excluded from the sampling to prevent possible re-identification. Gender in the data corresponds to the information provided to the health

insurance funds. If the insured persons do not consider themselves to be male or female, they can enter 'unspecified' or 'diverse' instead. In addition, transgender persons have the option to change the data regarding gender assigned at birth, so "male" might include cisgender and transgender men, and "female" might include cisgender and transgender women. For diseases associated with ICD-10 codes that have 15,000 patients or fewer, the study includes data for all patients identified with those conditions. In case of more than 15,000 patients for a respective ICD-10 code, a random sample of AOK-insured persons was drawn, stratified by 10-year age groups, gender and HIV indicators, so that a total of 15,000 AOK-insured persons were drawn per ICD-10 sample. For each sample, it was checked whether the distribution of age, gender and frequency of HIV indicators in the sample matched the distribution in the AOK ICD-10 population. The disease samples were drawn with replacement, so that the draws are independent of each other and persons could be selected for more than one disease sample. The AOK Research Institute created a list of which anonymous study IDs were to be assigned to which ICD-10 samples.

Anonymous data selected by the AOK Research Institute were edited and analyzed at the Institute for Healthcare Management and Research. First, ICD-10 codes for ICs were identified in outpatient, inpatient, and incapacity for work diagnoses, along with associated dates of treatment initiation/ICD-10 start/admission dates during the period from 07/01/2016 to 12/31/2017 (18 months). Outpatient diagnoses were only considered if they were labelled by the physician as "confirmed". Next, different disease-related ICD-10 codes were grouped into ICs (see Table 3). Within an IC, the chronologically latest date of incidence in the timeline of the 18-month time period was adopted as the date of the IC. To ensure accuracy in identifying HIV diagnoses, we chose a method with higher specificity. This approach aims to minimize the risk of overestimating HIV prevalences/incidences due to potential miscoding. Specifically, if billing codes indicate counselling for HIV pre-exposure prophylaxis (PrEP), initiation of PrEP, or monitoring during PrEP treatment, any prior or simultaneous HIV diagnoses (ICD-10 codes: B20 – B24, O98.7, or Z21) are considered suspect and excluded. These are likely miscodes. In cases where an HIV infection is diagnosed during PrEP therapy, specific HIV billing codes are used instead of PrEP-related codes. After a person is known to be HIV-positive, PrEP is discontinued, and thus, any subsequent HIV diagnoses are not considered miscoded. In this analysis only outpatient diagnoses with the code "confirmed" were considered. As an additional condition, HIV diagnoses were only considered if they were coded in two consecutive quarters (for inpatient or outpatient). To consolidate the conservative approach, this is a procedure to avoid concerning false diagnoses, which are due

to miscoding. The date of the earliest coded HIV diagnosis, that was not falsified along the criteria described above, was considered as the date of the initial HIV diagnosis in further analyses.

### Data analysis

Because only diagnosed HIV infections are coded in the statutory health insurance data, a literature-based assumption was made that HIV infections existed 40 months before their initial diagnosis.<sup>38</sup> Effects of a smaller or longer follow-up period were tested in different sensitivity analyses. In the primary analysis, HIV prevalence per IC was determined by first identifying the IC population (cumulatively fulfilling the characteristics: picked up with IC as described above in the first 18 months of the data period & part of the AOK Research Institute sample for that IC & surviving at least 40 months after the date of the IC) (see Fig. 1). The condition that the person must also be part of the random sample for the specific IC was chosen to ensure representativeness. Otherwise, people from other samples with the same diagnosis would distort the sample and make it dependent on other samples. Survival time of at least 40 months was required in the primary analysis to ensure an appropriate follow-up. Since for some ICs (e.g., malignant neoplasm of bronchus and lung), a rather high proportion of people were excluded due to that condition, a sensitivity analysis was carried out to test the effect of this condition.

HIV prevalence/incidence was calculated as the quotient of the number of persons in each IC population with HIV and the total number of persons in the IC population. An IC was considered confirmed if the 95%-confidence interval (CI) of HIV prevalence/incidence in the respective random sample was  $\geq 0.1\%$ . Ninety-five-percent-CI were calculated using bootstrapping. For each IC, 10,000 iterations were calculated. This number of iterations meets or exceeds common recommendations.<sup>39–42</sup>

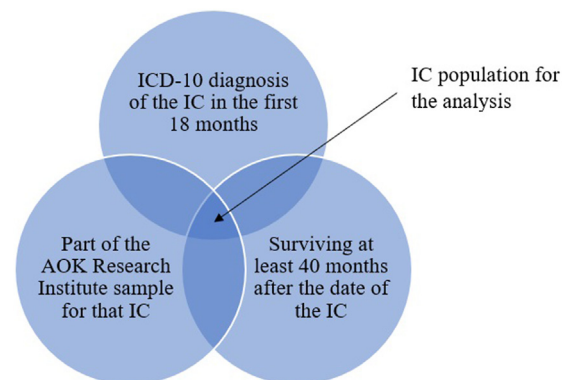


Fig. 1: IC population in the primary analysis.

To test the influence of the assumptions/conditions regarding different follow-up scenarios, three sensitivity analyses were performed: Sensitivity analysis 1: How do the results change if those persons who die within 40 months after IC are also considered when creating the IC populations? Sensitivity analysis 2: How do the results change if only HIV infections that are already contemporaneous with IC are considered? Sensitivity analysis 3: How do the results change if all HIV infections up to the end of the data period are considered? In addition, a fourth sensitivity analysis was calculated, following the different approaches of defining ICs: here, HIV incidence were calculated similar to the primary analysis, but this time only individuals who were without HIV diagnosis at the time of IC but had an incident HIV infection in the following 40 months were included in the IC population (sensitivity analysis 4) (see [Table 1](#)).

To test whether HIV prevalences in the primary analysis are substantially impacted by individual sub-populations, stratifications by gender (male/female (as registered at the statutory health insurance)) and age group (younger than 30 years/30 to <60 years/60 years and older) are calculated for the primary analysis. Because some ICs were expected to have very low number of cases in individual strata (e.g., cancers in persons <30 years of age), we reported stratification results only for strata with  $\geq 50$  individuals.

At the AOK Research Institute, Oracle Database 19c and Python version 3.10 were used for data management and SPSS Statistics version 29 was used at the Institute for Healthcare Management and Research for the analyses outlined above.

### Ethics statement

In the presented analysis, only anonymous data were used, so no informed consent was required. Nevertheless, a positive ethical vote was obtained for the overall study from the Ethics Committee of the Medical Faculty of the University of Duisburg-Essen (identifier: 22-10908-BO). Furthermore, the study was registered in the German Clinical Trials Register (identifier: DRKS00028743) and the study protocol was published open access.<sup>37</sup>

### Role of the funding source

The HeLP study is publicly funded by the German Federal Joint Committee (G-BA, “Gemeinsamer Bundesausschuss”) as part of the Innovationsfonds program to further develop the German healthcare system based on the standards and principles of evidence-based healthcare (funding number 01VSF21050). The funder of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report.

### Results

Data of 513,509 insured persons were transmitted from the AOK Research Institute to the Institute for Healthcare Management and Research. Sample sizes ranged from 353 to 15,000 individuals per ICD-10 code and from 353 (oral hairy leukoplakia) to 119,148 (pneumonias) per grouped IC. Ninety-two percent of persons only occurred in one ICD-10 sample, 7% occurred in two samples, the remaining 1% occurred in three to seven samples. Regarding the grouped ICs, 95% of persons only occurred in one IC, and 4% occurred in two ICs. Due to rounding deviations, there are only 0.3% of persons in more than two ICs. A short description of patient characteristics in each identified IC population can be found in [Table 2](#).

The HIV prevalences and corresponding 95%-CIs per IC are shown in [Table 3](#). All but one IC could be confirmed in the primary analysis (as stated in the methods section, an IC was considered confirmed if the 95%-confidence interval of HIV prevalence/incidence in the respective sample was  $\geq 0.1\%$ ). Only for lung carcinoma, the lower limit of the 95%-CI was below 0.1%. However, HIV prevalences varied widely. The highest prevalences of HIV infections were seen in patients with pneumocystosis (HIV prevalence: 40.33%, 95%-CI: 36.42–44.23%), hairy leukoplakia (HIV prevalence: 36.71%, 95%-CI: 29.11–44.3%), Kaposi’s sarcoma (HIV prevalence: 29.86%, 95%-CI: 27.05–32.77%), syphilis (HIV prevalence: 11.49%, 95%-CI: 10.95–12.05%), and gonococcal infection (HIV prevalence: 6.92%, 95%-CI: 6.31–7.55%). Considerably lower prevalences were observed, for example, in patients with seborrheic

	Duration of follow-up	Inclusion of patients who died within follow-up	Consideration of HIV diagnoses
Primary analysis	40 months	no	prevalent and incidental diagnoses
Sensitivity analysis 1	40 months	yes	prevalent and incidental diagnoses
Sensitivity analysis 2	no follow-up; only HIV infections that are already contemporaneous with IC are considered	no	prevalent and incidental diagnoses
Sensitivity analysis 3	from IC diagnosis until the end of data period	no	prevalent and incidental diagnoses
Sensitivity analysis 4	40 months	no	Only incidental diagnoses of individuals, who were without HIV diagnosis at the time of IC

**Table 1:** Overview regarding the primary and the sensitivity analyses.



Indicator condition	Sample size	Age <sup>a</sup> (reference date: 07/01/2016)	Gender <sup>b</sup>	
		Mean (standard deviation)	Male (in%)	Female (in%)
Pneumocystosis	615	56 (15.2)	64.6	35.4
Oral hairy leukoplakia	158	54 (13.6)	54.4	45.6
Kaposi's sarcoma	998	60 (15.2)	70.0	30.0
Syphilis	12,939	57 (17.6)	59.2	40.8
Gonococcal infection	6305	42 (17.3)	56.0	44.0
Acute hepatitis C	9128	52 (12.8)	59.0	41.0
Malignant neoplasm of anus and anal canal	4791	64 (12.4)	37.7	62.3
Chronic viral hepatitis	9620	51 (13.5)	56.5	43.5
Anogenital (venereal) warts	6261	38 (13.1)	55.7	44.3
Chlamydial infection	6020	31 (13.5)	30.8	69.2
Acute hepatitis B	9975	51 (14)	50.6	49.4
Papillomavirus as the cause of diseases	3887	37 (12.5)	4.3	95.7
Tuberculosis	17,107	57 (17.3)	48.3	51.7
Hodgkin's lymphoma	7981	54 (17.1)	50.2	49.8
Acute hepatitis A	9090	57 (15)	47.3	52.7
Pneumonia	24,013	60 (17.5)	52.2	47.8
Non-Hodgkin's lymphoma	23,271	64 (14.4)	49.2	50.8
Infectious mononucleosis	6858	37 (16.3)	38.7	61.3
Herpes simplex infections	5877	50 (18.4)	30.0	70.0
Abnormal weight loss	4262	56 (19.6)	36.9	63.1
Candidiasis	5146	49 (19.1)	16.5	83.5
Herpes zoster	4635	60 (17.4)	36.4	63.6
Dysplasia/malignant neoplasm of cervix uteri	14,508	48 (16.5)	0.1 <sup>c</sup>	99.9
Trichomoniasis	6512	42 (15.1)	14.9	85.1
Malignant neoplasm of bronchus and lung	2467	66 (11.2)	57.6	42.4
Seborrheic dermatitis	6229	55 (18.6)	51.7	48.3

<sup>a</sup>The first date of the data period was chosen as a reference date, so all people in the data were at least 18 years of age (age of majority in Germany). <sup>b</sup>As registered at the insurance institution, the patient has the possibility to change the data. <sup>c</sup>This group might include transgender men with a cervix uteri.

**Table 2: Distribution of patient characteristics by indicator conditions (primary analysis).**

dermatitis (HIV prevalence: 0.22%, 95%-CI: 0.11–0.35%), trichomoniasis (HIV prevalence: 0.26%, 95%-CI: 0.14–0.4%), dysplasia/malignant neoplasm of cervix uteri (HIV prevalence: 0.26%, 95%-CI: 0.18–0.35%), herpes zoster (HIV prevalence: 0.35%, 95%-CI: 0.19–0.52%), or candidiasis (HIV prevalence: 0.35%, 95%-CI: 0.19–0.52%).

The results of the primary analysis were very robust. All of the three sensitivity analyses which tested the assumptions/conditions from the primary analysis confirmed a 95%-CI overlapping 0.1% for lung cancer and 95%-CIs  $\geq 0.1\%$  for all other ICs (see [Supplementary Table S1](#)). If, as in sensitivity analysis 4, we did not consider the random sample of an IC, but only the individuals in the random sample who were without HIV diagnosis at the time of the IC, the HIV incidences were significantly lower with respect to the following 40 months. Only in six of the 26 ICs considered, the 95%-CI were never overlapping 0.1% (pneumocystosis (HIV incidence: 1.34%, 95%-CI: 0.27–2.69%), gonococcal infection (HIV incidence: 1.13%, 95%-CI: 0.86–1.4%), Kaposi's sarcoma (HIV incidence: 0.85%, 95%-CI: 0.28–1.56%), syphilis (HIV

incidence: 0.67%, 95%-CI: 0.52–0.82%), anogenital (venereal) warts (HIV incidence: 0.23%, 95%-CI: 0.11–0.36%), and chlamydial infection (HIV incidence: 0.2%, 95%-CI: 0.1–0.32%)) (see [Supplementary Table S1](#)).

[Table 4](#) provides an overview regarding the results of stratification of the primary analysis by gender. Significantly higher prevalences of HIV in male patients are noticeable. This is particularly the case for some AIDS-defining conditions and sexual transmitted infections (STI). While there were no changes with regard to the threshold of 0.1% in male patients, the 95%-CIs in female patients fell below the threshold for seven conditions that were  $\geq$  the threshold in the primary analysis: chlamydial infection, infectious mononucleosis, herpes simplex infections, abnormal weight loss, candidiasis, herpes zoster, and seborrheic dermatitis. It can be observed that the site of candidiasis differs between the genders. While candidiasis of the vulva and vagina is by far the most common diagnosis in female patients with candidiasis, candidal stomatitis is the most common diagnosis in male patients with candidiasis.

HIV indicator condition	ICD-10 codes	Sample size	HIV prevalence in % (95%-CI)
Pneumocystosis	B48.5, B59	615	40.33 (36.42–44.23)
Oral hairy leukoplakia	K13.3	158	36.71 (29.11–44.3)
Kaposi's sarcoma	C46	998	29.86 (27.05–32.77)
Syphilis	A50, A51, A52, A53	12,939	11.49 (10.95–12.05)
Gonococcal infection	A54	6305	6.92 (6.31–7.55)
Acute hepatitis C	B17.1	9128	3.13 (2.78–3.49)
Malignant neoplasm of anus and anal canal	C21	4791	2.94 (2.46–3.42)
Chronic viral hepatitis	B18	9620	2.93 (2.6–3.26)
Anogenital (venereal) warts	A63.0	6261	2.86 (2.46–3.27)
Chlamydial infection	A55, A56	6020	2.24 (1.88–2.62)
Acute hepatitis B	B16	9975	1.66 (1.41–1.91)
Papillomavirus as the cause of diseases	B97.7	3887	1.65 (1.26–2.06)
Tuberculosis	A15, A16, A17, A18, A19	17,107	1.57 (1.39–1.76)
Hodgkin's lymphoma	C81	7981	1.13 (0.9–1.37)
Acute hepatitis A	B15	9090	1.08 (0.87–1.29)
Pneumonia	J10.0, J11.0, J12, J13, J14, J15, J16, J17, J18	24,013	0.91 (0.79–1.03)
Non-Hodgkin's lymphoma	C82, C83, C84, C85, C86	23,271	0.53 (0.44–0.63)
Infectious mononucleosis	B27	6858	0.39 (0.26–0.55)
Herpes simplex infections	B00	5877	0.39 (0.24–0.56)
Abnormal weight loss	R63.4	4262	0.38 (0.21–0.56)
Candidiasis	B37	5146	0.35 (0.19–0.52)
Herpes zoster	B02	4635	0.35 (0.19–0.52)
Dysplasia/malignant neoplasm of cervix uteri	C53, N87	14,508	0.26 (0.18–0.35)
Trichomoniasis	A59	6512	0.26 (0.14–0.4)
Malignant neoplasm of bronchus and lung	C34	2467	0.24 (0.08–0.45)
Seborrheic dermatitis	L21	6229	0.22 (0.11–0.35)

**Table 3: HIV prevalences in HIV indicator conditions in the primary analysis.**

Differences between the groups are also apparent when stratified by age. Predominantly, the highest HIV prevalences are found in patients between 30 and 59 years of age (see Table 5). While the 95%-CIs for this group were  $\geq 0.1\%$  for each IC, the 95%-CIs fell below the 0.1% thresholds for some conditions regarding the other groups. In patients younger than 30 years, the 95%-CIs fell below the threshold in the following conditions: Seborrheic dermatitis, trichomoniasis, dysplasia/malignant neoplasm of cervix uteri, herpes zoster, candidiasis, abnormal weight loss, non-Hodgkin's lymphoma, and Hodgkin's lymphoma. In the group of patients 60 years and older, this is the case for seborrheic dermatitis, malignant neoplasm of bronchus and lung, trichomoniasis, dysplasia/malignant neoplasm of cervix uteri, herpes zoster, candidiasis, abnormal weight loss, herpes simplex, and mononucleosis. As shown in Table 5, due to the small number of cases in the stratum of those under 30 years of age, no results were considered for five diseases. These are predominantly malignant diseases.

## Discussion

In the HeLP study, large samples of patients diagnosed with internationally described ICs were examined

regarding the occurrence of HIV, in order to investigate if these ICs are useable as an indicator for a HIV infection in Germany (HIV prevalence is  $\geq 0.1\%$ ). Most of the IC populations show HIV prevalence rates where HIV testing would be effective. As mentioned in the introduction, there are other possible reasons besides HIV prevalence for classifying a disease as an HIV indicator condition. Due to the methodology of the analysis, which was focused on the frequency of HIV infections, these additional reasons are not considered here. The results should be viewed from this perspective. Recommendations for other reasons should explicitly not be disproved. They just may not be additionally supported by prevalence/incidence presented here. As described above, HIV testing recommendations can thus be made for HIV indicator conditions.

The recommendation for testing depends on the level of differentiation regarding patient characteristics. If only the disease status is considered, HIV testing would be recommended for all IC examined except malignant neoplasm of bronchus and lung. A more complex picture emerges when patients are differentiated based on their sociodemographic characteristics such as gender and age. In order to avoid potential stigmatization of groups and complicated testing recommendations in guidelines, it seems reasonable that HIV testing recommendations may be based on the

HIV indicator condition	Male <sup>a</sup>		Female <sup>a</sup>	
	Sample size	HIV prevalence in % (95%-CI)	Sample size	HIV prevalence in % (95%-CI)
Pneumocystosis	397	51.13 (46.35-55.92)	218	20.64 (15.6-26.15)
Oral hairy leukoplakia	86	52.33 (41.86-62.79)	72	18.06 (9.72-27.78)
Kaposi's sarcoma	699	39.34 (35.77-43.06)	299	7.69 (4.68-11.04)
Syphilis	7658	18.7 (17.85-19.56)	5281	1.04 (0.78-1.33)
Gonococcal infection	3532	11.66 (10.62-12.74)	2773	0.87 (0.54-1.23)
Acute hepatitis C	5383	3.9 (3.4-4.42)	3745	2.03 (1.58-2.48)
Malignant neoplasm of anus and anal canal	1806	7.14 (5.98-8.31)	2985	0.4 (0.2-0.64)
Chronic viral hepatitis	5431	3.81 (3.31-4.33)	4189	1.79 (1.41-2.22)
Anogenital (venereal) warts	3487	4.59 (3.9-5.28)	2774	0.68 (0.4-1.01)
Chlamydial infection	1855	7.06 (5.88-8.25)	4165	0.1 (0.02-0.19)
Acute hepatitis B	5049	2.32 (1.92-2.75)	4926	0.99 (0.73-1.28)
Papillomavirus as the cause of diseases	166	21.69 (15.66-28.31)	3721	0.75 (0.48-1.05)
Tuberculosis	8271	2.01 (1.72-2.31)	8836	1.15 (0.94-1.38)
Hodgkin's lymphoma	4010	1.87 (1.47-2.29)	3971	0.38 (0.2-0.58)
Acute hepatitis A	4296	1.61 (1.23-2)	4794	0.6 (0.4-0.83)
Pneumonia	12,533	1.22 (1.03-1.42)	11,480	0.57 (0.44-0.71)
Non-Hodgkin's lymphoma	11,444	0.81 (0.66-0.98)	11,827	0.26 (0.18-0.36)
Infectious mononucleosis	2654	0.79 (0.49-1.13)	4204	0.14 (0.05-0.26)
Herpes simplex infections	1764	0.91 (0.51-1.36)	4113	0.17 (0.05-0.32)
Abnormal weight loss	1571	0.76 (0.38-1.21)	2691	0.15 (0.04-0.3)
Candidiasis	848	1.53 (0.71-2.36)	4298	0.12 (0.02-0.23)
Herpes zoster	1686	0.59 (0.24-1.01)	2949	0.2 (0.07-0.37)
Dysplasia/malignant neoplasm of cervix uteri	<50 <sup>b</sup>	-	14,490	0.26 (0.18-0.35)
Trichomoniasis	968	0.52 (0.1-1.03)	5544	0.22 (0.11-0.34)
Malignant neoplasm of bronchus and lung	1421	0.35 (0.07-0.7)	1046	0.1 (0-0.29)
Seborrheic dermatitis	3222	0.4 (0.22-0.62)	3007	0.03 (0-0.1)

<sup>a</sup>As registered at the insurance institution, the patient has the possibility to change the data. <sup>b</sup>This group might include transgender men with a cervix uteri. However, the sample size was too small for an investigation of HIV prevalence.

**Table 4: HIV prevalences in HIV indicator conditions stratified by gender.**

non-stratified analysis. However, to ensure transparency and specificity, a stratified presentation by age and sex has been taken into account.

In patients diagnosed with syphilis, gonococcal infection, acute hepatitis C, chronic viral hepatitis, anogenital (venereal) warts, acute hepatitis B, papillomavirus as the cause of diseases, tuberculosis, acute hepatitis A, and pneumonia, ICs never overlapped the threshold value of 0.1% in the primary analysis and in all stratifications of the primary analysis. Being AIDS-defining conditions, tuberculosis and recurrent pneumonias are also generally considered as ICs regardless of HIV prevalence.<sup>43</sup>

There are gaps in evidence for some strata regarding oral hairy leukoplakia, Kaposi's sarcoma, and malignant neoplasm of anus and anal canal. In these ICs, the stratified results regarding HIV prevalences in under 30 years old were not considered, due to the low number of cases. However, in the primary analysis and all considered strata, the lower ends of the 95%-CI were always  $\geq 0.1\%$ . In addition, pneumocystosis and Kaposi's sarcoma a generally regarded as ICs, because these are AIDS-defining conditions.<sup>43</sup>

In the following ICs, the 95%-CI was  $\geq$  the threshold in the primary analysis, but fell below the threshold in some stratifications regarding age or gender: Chlamydial infection (below the threshold in women), Hodgkin's lymphoma (below the threshold in patients under 30 year old), Non-Hodgkin's lymphoma (below the threshold in patients under 30 years old), infectious mononucleosis (below the threshold in women or patients older than 60 years), herpes simplex infections (below the threshold in women or patients older than 60 years), abnormal weight loss (below the threshold in women or patients younger than 30 or patients older than 60 years), candidiasis (below the threshold in women or patients younger than 30 or patients older than 60 years), herpes zoster (below the threshold in patients younger than 30 or patients older than 60 years), dysplasia/malignant neoplasm of cervix uteri (below the threshold in patients younger than 30 or patients older than 60 years), trichomoniasis (below the threshold in patients younger than 30 or patients older than 60 years), and seborrheic dermatitis (below the threshold in women or patients younger than 30 or patients older than 60 years). It should be emphasized



HIV indicator condition	<30 years of age		30–< 60 years of age		From 60 years of age	
	Sample size	HIV prevalence in % (95%-CI)	Sample size	HIV prevalence in % (95%-CI)	Sample size	HIV prevalence in % (95%-CI)
Pneumocystosis	<50	–	314	63.38 (57.97–68.79)	268	13.06 (8.96–17.16)
Oral hairy leukoplakia	<50	–	97	48.45 (38.14–58.76)	57	15.79 (7.02–26.32)
Kaposi's sarcoma	<50	–	477	51.78 (47.18–56.18)	497	9.26 (6.84–11.87)
Syphilis	996	14.56 (12.35–16.77)	5775	20.85 (19.83–21.9)	6168	2.24 (1.88–2.61)
Gonococcal infection	1885	4.3 (3.4–5.25)	3248	10.13 (9.08–11.18)	1172	2.22 (1.37–3.16)
Acute hepatitis C	230	3.48 (1.3–6.09)	6608	3.92 (3.45–4.4)	2290	0.83 (0.48–1.22)
Malignant neoplasm of anus and anal canal	<50	–	1763	6.24 (5.1–7.37)	3014	1 (0.66–1.36)
Chronic viral hepatitis	470	1.91 (0.85–3.19)	6811	3.7 (3.26–4.16)	2339	0.9 (0.56–1.33)
Anogenital (venereal) warts	1845	1.19 (0.7–1.73)	3958	3.74 (3.16–4.32)	458	1.97 (0.87–3.28)
Chlamydial infection	3584	0.7 (0.45–0.98)	2112	4.88 (3.98–5.82)	324	2.16 (0.62–3.7)
Acute hepatitis B	672	0.74 (0.15–1.49)	6652	2.04 (1.71–2.39)	2651	0.94 (0.6–1.32)
Papillomavirus as the cause of diseases	1230	0.49 (0.16–0.89)	2417	2.19 (1.61–2.81)	240	2.08 (0.42–4.17)
Tuberculosis	1430	0.84 (0.42–1.33)	7841	2.88 (2.51–3.26)	7836	0.38 (0.26–0.52)
Hodgkin's lymphoma	805	0.25 (0–0.62)	4005	1.87 (1.47–2.3)	3171	0.41 (0.19–0.63)
Acute hepatitis A	367	1.09 (0.27–2.18)	4725	1.71 (1.35–2.1)	3998	0.33 (0.15–0.53)
Pneumonia	1576	0.82 (0.38–1.33)	9181	1.87 (1.6–2.16)	13,256	0.25 (0.17–0.34)
Non-Hodgkin's lymphoma	619	0.48 (0–1.13)	7330	1.3 (1.04–1.57)	15,322	0.17 (0.1–0.23)
Infectious mononucleosis	2942	0.27 (0.1–0.48)	3164	0.57 (0.32–0.85)	752	0.13 (0–0.4)
Herpes simplex infections	1036	0.39 (0.1–0.77)	3006	0.5 (0.27–0.77)	1835	0.22 (0.05–0.44)
Abnormal weight loss	550	0.18 (0–0.55)	1705	0.82 (0.41–1.29)	2007	0.05 (0–0.15)
Candidiasis	1040	0 (0–0)	2541	0.55 (0.28–0.87)	1565	0.26 (0.06–0.51)
Herpes zoster	312	0.64 (0–1.6)	1713	0.76 (0.35–1.23)	2610	0.04 (0–0.11)
Dysplasia/malignant neoplasm of cervix uteri	2121	0.09 (0–0.24)	8709	0.39 (0.26–0.53)	3678	0.05 (0–0.14)
Trichomoniasis	1696	0 (0–0)	4021	0.4 (0.22–0.6)	795	0.13 (0–0.38)
Malignant neoplasm of bronchus and lung	<50	–	655	0.61 (0.15–1.22)	1798	0.11 (0–0.28)
Seborrheic dermatitis	779	0.13 (0–0.39)	2666	0.41 (0.19–0.68)	2784	0.07 (0–0.18)

Table 5: HIV prevalences in HIV indicator conditions stratified by age.

that AIDS-defining conditions like Burkitt lymphoma, immunoblastic lymphoma, primary lymphoma of brain, herpes simplex (in some locations/duration), wasting syndrome attributed to HIV, candidiasis of bronchi, trachea, lung or esophagus, and invasive cervical cancer are generally considered to be ICs regardless of HIV prevalence.<sup>43</sup>

In the group of malignant neoplasm of bronchus and lung, the 95%-CI of HIV prevalence was  $\geq$  the threshold only in one stratum (patients 30–< 60 years of age).

The integration of the results of the HeLP study into published evidence is only possible with limitations. The occurrence of HIV and ICs differs between different regions (e.g., with regard to the overall prevalence/incidence and the predominantly affected groups of people), so in order to keep the bias low, the comparison of prevalences/incidences is only made with publications that refer to Europe, North America, Australia, New Zealand, or Japan. However, there may be differences in the epidemiology of HIV and ICs between these regions. Another important point is that the study designs differ (e.g., how the HIV prevalences/incidences are determined, which time period is taken into account and which patients are included). Besides

differences in study design, there are also differences in HIV testing strategies in usual care.<sup>44</sup> In addition, many studies examine only a relatively small number of patients. These samples are often selective, since certain groups of people are more likely to be offered an HIV test than others, which means that published HIV prevalences/incidences relating to selective subgroups probably overestimate HIV prevalence/incidence. Since our sensitivity analyses 1 to 3 include also patients with known HIV diagnoses, the results cannot be extrapolated to reflect a hypothetical positivity rate among people with unknown HIV status tested for HIV. Moreover, people living with HIV may be more likely to be screened for some diseases, e.g., STI, and thus have a higher chance of being selected for an indicator disease sample. This would result in an overrepresentation of people with HIV among those indicator conditions and an overestimate of HIV prevalence in these sensitivity analyses. However, only individuals who were without HIV diagnosis at the time of IC but had an incident HIV infection in the following 40 months were included in sensitivity analysis 4.

STIs are commonly studied ICs in the corresponding literature.<sup>17,25,27,30</sup> In terms of summarized investigation

of STIs, the published results vary. Agusti et al. and Matulionytė et al. did not confirm "unspecified STI" as IC, but Sullivan et al. did.<sup>17,25,27</sup> In Omland et al., 2016, the appropriateness examining "other STI" as an IC depended on sex (higher for male than for female).<sup>30</sup> The difference by sex reported by Omland et al. fits well to our presented results. Additionally, presented HIV prevalences/incidences vary among different STIs, despite the commonality in the mode of transmission. One possible explanation for this can be found in sociodemographic differences. For example, the data used show that the majority of persons diagnosed with trichomoniasis are women, whereas men make up the majority for other STIs. Agusti et al. observed the highest proportions of HIV within STIs in syphilis, followed by gonococcal infections and, at a clear distance, chlamydial infections, which corresponds to the results of the primary analysis in HeLP.<sup>25</sup> In contrast, in sensitivity analysis 4, HIV proportions are lower for syphilis than for gonococcal infections. Moreover, contrary to the findings of Omland et al., in HeLP syphilis is also a suitable indicator disease in females.<sup>30</sup>

Viral hepatitis were also considered in a number of studies.<sup>17,22,25,28,30,32,33</sup> As in the HeLP study, acute hepatitis infections are included in Girardi et al. and HIV proportions considerably higher than 0.1% are observed in men.<sup>28</sup> In contrast, HIV proportion found in this analysis are  $\geq 0.1\%$  for women, too. Published HIV proportions of other studies are difficult to compare because, for example, both acute and chronic hepatitis diseases have been reported together or the authors summarized, e.g., in "non-A hepatitis" or "hepatitis B or C".<sup>17,22,25,30,32,33</sup> In the vast majority of publications, hepatitis diseases are confirmed as ICs which fits to the results of the presented primary analysis, but differs to the results of sensitivity analysis 4.<sup>17,22,25,28,30,32,33</sup> In addition, results presented here are in line with reports about higher HIV proportions in patients with hepatitis C than in patients with hepatitis B.<sup>22,25</sup>

In terms of HIV proportions in malignant disease, cervical dysplasia/carcinoma is the most commonly investigated.<sup>17,22,25,26,32,33</sup> However, definitions of ICs vary between publications. For example, it is reported together with anal dysplasia/cancer in Sullivan et al.<sup>17</sup> The published HIV proportions vary accordingly. In this regard a trend cannot be identified.<sup>17,22,25,26,32,33</sup>

It is notable that Kaposi's sarcoma has a very high HIV prevalence in our primary analysis and could also be confirmed as one of few ICs in sensitivity analysis 4, but previously reported data are scarce. Agusti et al. reported a proportion of zero for HIV diagnoses in this IC.<sup>25</sup> However, this may be due to the small number of cases in the cited study.<sup>25</sup>

In contrast to the presented primary analysis, anal carcinoma was not confirmed as IC in different definitions and subgroups by Raben et al. and Sullivan et al.<sup>17,32,33</sup> Nevertheless, their findings are in line with

our sensitivity analysis 4. Strikingly, in our primary analysis the HIV prevalence in anal carcinoma is  $\geq 0.1\%$  in all subgroups considered.

The evidence regarding HIV prevalence/incidence in Hodgkin's and non-Hodgkin's lymphoma is still lacking. Sullivan et al. report a HIV proportion of less than 0.1% for malignant lymphoma in general, while in a study by Bogers et al. 0.7% of all patients with malignant lymphoma tested for HIV within 3 months before or after diagnosis were HIV-positive.<sup>17,45</sup> Agusti et al. report no HIV cases for "lymphoma other than non-Hodgkin's". For non-Hodgkin's lymphoma, they report an 95%-CI above 0.1% only for males.<sup>25</sup> For non-Hodgkin's lymphoma, Raben et al. report varying results depending on audit.<sup>32</sup> These weak conformations fit with rather low prevalences as reported in the primary analysis of HeLP, where prevalences  $\geq 0.1\%$  are nevertheless achieved in most stratifications.

In line with our study, Raben et al. ("primary lung cancer") and Omland et al. ("respiratory tract cancer") did not find proportions of HIV in malignant neoplasm of bronchus and lung to classify them as an IC.<sup>30,33</sup>

The proportion of HIV in patients with pneumocystosis or oral hairy cell leukoplakia have not been well investigated before, but the reported high numbers in HeLP are clinically plausible.

Reported HIV proportions in patients with tuberculosis vary widely.<sup>22,25,27,30,32</sup> While Bogers et al. report a proportion of about 5%, the proportion reported by Matulionytė et al. is less than 1% and even below the threshold according to Agusti et al.<sup>22,25,27</sup> Omland et al. and Raben et al. come to different results depending on stratum/audit.<sup>30,32</sup> In the primary analysis of HeLP, HIV testing can be recommended due to HIV prevalence in all strata considered, but with rather intermediate to low HIV rates in the lower single-digit percentage range.

The published evidence regarding HIV proportions in pneumonia is also very heterogeneous.<sup>25,30,31,33</sup> Peck et al. report (at a low case number) an HIV proportion of 21% (including already diagnosed infections) for "all types of pneumonia," whereas Agusti et al. report zero HIV cases for "recurrent pneumonia".<sup>25,31</sup> The results by Omland et al. and Raben et al. show proportions between these results.<sup>30,33</sup> While Raben et al. confirmed pneumonia with at least 24 h of hospitalization as IC, Omland et al. found this for different types of pneumonia to be dependent on e.g., age or gender.<sup>30,33</sup> The results of HeLP indicate pneumonia being an IC in the primary analysis, although HIV prevalences are rather low.

Concerning herpes zoster, several publications report HIV proportions in the single-digit percentage range.<sup>17,25,27</sup> However, in contrast, Menacho et al. report zero HIV cases, and in results presented by Omland et al. it is age-dependent in females whether an HIV proportion is above the threshold.<sup>29,30</sup> In the primary analysis of HeLP, no testing recommendation due to HIV prevalence for women and specific age groups can be confirmed.

The situation is similar for herpes simplex. While the published evidence is rather weak (few publications in some of which only certain herpes simplex infections are considered) and does not offer a consistent conclusion, a recommendation for testing due to HIV prevalence limited to specific groups of age or gender can also be made regarding the results of HeLP.<sup>25,27,30</sup>

In terms of candidiasis, more specific cohorts than those used here are included in most of the published studies.<sup>25,27,30,32</sup> To ensure transparency, a more detailed presentation of HIV prevalences in patients with candidiasis is given in [Supplementary Table S2](#). This more detailed breakdown is provided in the appendix, as these results cannot be interpreted reliably. This is partly due to the often small number of cases and partly to the rather unspecific subcodes specified by ICD-10. Furthermore, although the cumulative candidiasis sample corresponds to the ICD-10 population of the health insurance funds, this is not necessarily the case for the more specific subcodes. Agusti et al. reported HIV proportions above 0.1% in women with non-pulmonary candidiasis, Matulionytė et al. reported an HIV proportion of zero in oral candidiasis, and Raben et al. confirmed esophageal candidiasis as an IC.<sup>25,27,32</sup> The results by Omland et al. depend on age and sex.<sup>30</sup> Once again, based on the results shown here, a recommendation for testing due to HIV prevalence limited to specific groups of age or gender can be made. In the analysis presented, candidiasis of the vulva and vagina accounted for the largest proportion of diagnoses in female patients with candidiasis, while candidal stomatitis was most common in men. Differences in HIV prevalence between the genders may therefore also be due to differences in the type of candidiasis.

Results of HeLP cannot confirm the relatively high HIV proportions reported for mononucleosis and mononucleosis-like diseases in some studies.<sup>17,25,29,30,33</sup> According to the German data in the primary analysis, testing recommendation could be rethink for women and persons 60 years of age and older.

For seborrheic dermatitis, most reported HIV proportions are higher than presented here. However, the basic message that seborrheic dermatitis is an appropriate IC is consistent between most publications and fits to the results in the primary analysis of HeLP.<sup>17,25,27,29,30,33</sup> Consistent with Omland et al., this statement depends on age and sex.<sup>30</sup>

The significantly higher HIV prevalences/incidences among male persons with ICs reported here seem plausible, since in Germany, as in many western countries, men account for the majority of HIV-positive persons.<sup>1,5</sup> The very significant differences in HIV prevalences for chlamydial infection and papillomavirus as the cause of diseases also reflect the fact that in Germany, women are comparatively often routinely tested for these STIs, whereas in men this is predominantly only the case in the event of symptoms or known

risk behavior. It can be assumed that there is a high gap in the diagnosis of infections with chlamydia and papillomavirus in the German male population. Patients were grouped according to their registered gender, not their sex at birth. This may lead to different rates, as, for example, transgender women probably show higher HIV prevalence rates than cisgender women.<sup>46</sup> However, as the samples are unlikely to include many transgender persons, the effect is estimated to be minor. The general differences between age strata also seem plausible.

One potential limitation of our analysis relates to the composition of some ICs from individual ICD-10 codes. Hence, the random samples within ICD-10 codes are representative for the respective ICD-10 population of the cooperating statutory health insurances, but since a maximum of 15,000 patients per ICD-10 code were drawn, the ratio of individual ICD-10 codes within a grouped IC does not necessarily correspond to the actual distribution. However, this only affects six of the 26 ICs as the effect is only seen in ICs that are composed of multiple ICD-10 codes and in which individual ICD-10 codes are very common (syphilis, chlamydial infection, tuberculosis, pneumonia, non-Hodgkin's lymphoma, dysplasia/malignant neoplasm of cervix uteri).

In addition, the statutory health insurances involved do not represent the entire German population. As mentioned above, the AOK is open to all German inhabitants. Nevertheless, in a study published in 2017, Hoffman & Koller showed that, compared to other health insurance funds, people insured by the AOK are more likely to have a low socio-economic status and to be migrants. They are also more often obese and smokers.<sup>47</sup> However, it should also be noted that study populations are generally subject to selection, and the high number of cases from all over Germany may give a more representative result than, for example, recruitment from hospitals in large cities.

With around 27 million people, samples are drawn from a very high number of cases. The number of cases of each IC group far exceeds what can be achieved in most studies. Nevertheless, especially in the stratifications, it cannot be ruled out that 95%-CIs are not above the threshold due to a too small number of cases. Another limitation is that ICD-10 codes in routine data are primarily intended for billing and not for scientific evaluations and may contain miscoding.<sup>48,49</sup> Nevertheless, they are considered to be an established data source for scientific research.<sup>35,36</sup> A strength of routine data is that it is real evidence from everyday care, not from an artificial, often selective clinical trial population. Because of the observation of everyday care without intervention, there was no trial-specific HIV testing. It can be assumed that especially people with an increased risk of HIV are offered/willing to take an HIV test. Therefore, studies that focus on patients with HIV testing may

overestimate the prevalence of HIV in patients with IC. Vice versa, without trial-specific HIV testing, it is possible that some HIV infections remain undetected despite several years of follow-up, which would lead to an underestimation of HIV prevalence.

As shown in sensitivity analyses 1 to 3, the influence of other potential limitations, such as the fact that the exact time interval between HIV infection and initial HIV diagnosis is unknown and literature based estimated, is minor. The time period in the primary analysis does not lead to fundamentally different results than in the sensitivity analyses 1 to 3. Sensitivity analysis 4, however, shows a strong effect of whether HIV prevalence/incidence is calculated in an IC random sample or in a sample without people already living with HIV.

In conclusion, the results presented showed that a majority of the internationally described ICs examined could be confirmed for the German context in our primary analysis based on German statutory health insurance data, with remarkable differences regarding age and gender. So, the module of the HeLP study was able to fill an important evidence gap based on an exceptionally large and less selective cohort. The results are specific to Germany, but they can also be used to establish testing recommendations in other countries with a similar epidemiology of ICs and HIV. This is particularly useful if there are no studies for these regions or only a few with low case numbers. In the further course of the study, approaches will be developed to strengthen HIV testing recommendations regarding confirmed IC in corresponding medical guidelines.

#### Contributors

FV, GB, MB, CB, SE, PD, TR, AK, DS, UK, BGB, LW, JW, AN: Substantial contributions to the conception or design of the work; or the acquisition, analysis, or interpretation of data for the work.

FV: Drafting the work.

FV, AN: Accessed and verified the data.

GB, MB, CB, SE, PD, TR, AK, DS, UK, BGB, LW, JW, AN: Revising the work critically for important intellectual content.

FV, GB, MB, CB, SE, PD, TR, AK, DS, UK, BGB, LW, JW, AN: Final approval of the version to be submitted.

#### Data sharing statement

The authors confirm that the data utilized in this study cannot be made available in the manuscript, the supplemental files, or in a public repository due to German data protection laws ('Bundesdatenschutzgesetz', BDSG). Therefore, they are stored on a secure drive in the AOK Research Institute, to facilitate replication of the results. Generally, access to data of statutory health insurance funds for research purposes is possible only under the conditions defined in German Social Law (SGB V § 287). Requests for data access can be sent as a formal proposal specifying the recipient and purpose of the data transfer to the appropriate data protection agency. Access to the data used in this study can only be provided to external parties under the conditions of the cooperation contract of this research project and after written approval by the sickness fund. For assistance in obtaining access to the data, please contact [wido@wido.bv.aok.de](mailto:wido@wido.bv.aok.de).

#### Declaration of interests

GB: Grants or contracts from Niedersächsisches Ministerium für Wissenschaft und Kultur, COFONI Network, European Regional

Development Fund, German Center for Infection Research, consulting fees Gilead, ViiV, MSD, Virology Education, Janssen, payment or honoraria for lectures, presentations, speakers bureaus, manuscript writing or educational events from Gilead, ViiV, MSD, Virology Education, Janssen, support for attending meetings and/or travel from Gilead, ViiV, MSD, Janssen, Participation on a Data Safety Monitoring Board or Advisory Board of TherVacB\_Phase1a.

CB: Grants or contracts from DZIF, DFG and NEAT ID, consulting fees from Abbvie, JnJ, MSD, Gilead, Pfizer and ViiV, payment or honoraria for lectures, presentations, speakers bureaus, manuscript writing or educational events from Abbvie, AZ, BN, JnJ, Gilead, MSD, Pfizer and ViiV, support for attending meetings and/or travel from Abbvie, AZ, JnJ, Gilead, MSD, Pfizer and ViiV, leadership or fiduciary role in other board, society, committee or advocacy group, paid or unpaid in EACS.

SE: Grants or contracts from Gilead, Janssen, MSD, and ViiV consulting fees from Gilead, GSK, Janssen, MSD, and ViiV, payment or honoraria for lectures, presentations, speakers bureaus, manuscript writing or educational events from Gilead, Janssen, MSD, and ViiV, support for attending meetings and/or travel from Gilead, Janssen, MSD, and ViiV, Leadership or fiduciary role in other board, society, committee or advocacy group, paid or unpaid in DAIG, DAGNÄ, DSTIG, DDG, DGI, Regional commission AIDS NRW.

MB: Grants or contracts from Merck & Co, GSK plc, Gilead Science, leadership or fiduciary role in other board, society, committee or advocacy group, paid or unpaid in Deutsche Arbeitsgemeinschaft ambulant tätiger Ärztinnen und Ärzte für Infektionskrankheiten und HIV-Medizin e. V. (DAGNÄ e. V.).

UK: Support for attending meetings and/or travel (travel and accommodation costs for participation in meetings/conferences from ECDC, German AIDS Society, German STI Society, and European AIDS Society), other financial or non-financial interests (UK owns 'Exchange-traded fund' (ETF) shares, which includes stocks of companies that are involved in health care).

The remaining authors have no competing interests to declare.

#### Acknowledgements

The HeLP study is publicly funded by the German Federal Joint Committee (G-BA, "Gemeinsamer Bundesausschuss") as part of the Innovationsfonds program to further develop the German healthcare system based on the standards and principles of evidence-based healthcare (funding number 01VVSF21050).

During the preparation of this work the author(s) used DeepL in order to improve translations. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the publication.

We acknowledge support by the Open Access Publication Fund of the University of Duisburg-Essen.

#### Appendix A. Supplementary data

Supplementary data related to this article can be found at <https://doi.org/10.1016/j.eclinm.2024.102694>.

#### References

- 1 an der Heiden M, Marcus U, Kollan C, et al. *Epidemiologisches Bulletin* 47\_2022–47\_22: Aktuelle Daten und Informationen zu Infektionskrankheiten und Public Health; 2023. Available from: [https://www.rki.de/DE/Content/Infekt/EpidBull/Archiv/2022/Ausgaben/47\\_22.html](https://www.rki.de/DE/Content/Infekt/EpidBull/Archiv/2022/Ausgaben/47_22.html).
- 2 Johnson AS, Song R, Hall HI. Estimated HIV incidence, prevalence, and undiagnosed infections in US States and Washington, DC, 2010–2014. *J Acquir Immune Defic Syndr*. 2017. <https://doi.org/10.1097/QAI.0000000000001495>.
- 3 Pharris A, Quinten C, Noori T, Amato-Gauci AJ, van Sighem A. Estimating HIV incidence and number of undiagnosed individuals living with HIV in the European Union/European Economic Area, 2015. *Euro Surveill*. 2016. <https://doi.org/10.2807/1560-7917.ES.2016.21.48.30417>.
- 4 Schäfer G, Kreuels B, Schmiedel S, et al. High proportion of HIV late presenters at an academic tertiary care center in northern



- Germany confirms the results of several cohorts in Germany: time to put better HIV screening efforts on the national agenda? *Infection*. 2016;44(3):347–352. <https://doi.org/10.1007/s15010-016-0880-4>.
- 5 ECDC/WHO. HIV/AIDS surveillance in Europe 2022. 2021 data [June 13, 2023]; Available from: <https://www.who.int/europe/publications/i/item/9789289058636>.
  - 6 Joint United Nations Programme on HIV/AIDS. *Fast-track – Ending the AIDS epidemic by 2030* 2014.
  - 7 Bickel M, Hoffmann C, Wolf E, et al. High effectiveness of recommended first-line antiretroviral therapies in Germany: a nationwide, prospective cohort study. *Infection*. 2020;48:453–461. <https://doi.org/10.1007/s15010-020-01428-1>.
  - 8 Valbert F, Wolf E, Preis S, et al. HIV-Epidemiologie in Deutschland: Späte Diagnostik. *Dtsch Arztebl*. 2021;118(43):A1994–A1998.
  - 9 Kittner JM, Bialy L von, Wiltink J, et al. Lack of awareness in both patients and physicians contributes to a high rate of late presentation in a South West German HIV patient cohort. *Infection*. 2015;43(3):299–305. <https://doi.org/10.1007/s15010-014-0719-9>.
  - 10 Antinori A, Coenen T, Costagiola D, et al. Late presentation of HIV infection: a consensus definition. *HIV Med*. 2011;12(1):61–64. <https://doi.org/10.1111/j.1468-1293.2010.00857.x>.
  - 11 Croxford S, Rinder Stengaard A, Brännström J, et al. Late diagnosis of HIV: an updated consensus definition. *HIV Med*. 2022. <https://doi.org/10.1111/hiv.13425>.
  - 12 Sobrino-Vegas P, Moreno S, Rubio R, et al. Impact of late presentation of HIV infection on short-, mid- and long-term mortality and causes of death in a multicenter national cohort: 2004-2013. *J Infect*. 2016;72(5):587–596. <https://doi.org/10.1016/j.jinf.2016.01.017>.
  - 13 Sabin CA, Smith CJ, Gumley H, et al. Late presenters in the era of highly active antiretroviral therapy: uptake of and responses to antiretroviral therapy. *AIDS*. 2004;18(16):2145–2151. <https://doi.org/10.1097/00002030-200411050-00006>.
  - 14 Rodger AJ, Cambiano V, Bruun T, et al. Sexual activity without condoms and risk of HIV transmission in serodifferent couples when the HIV-positive partner is using suppressive antiretroviral therapy. *JAMA*. 2016;316(2):171–181. <https://doi.org/10.1001/jama.2016.5148>.
  - 15 Cohen MS, Chen YQ, McCauley M, et al. Antiretroviral therapy for the prevention of HIV-1 transmission. *N Engl J Med*. 2016;375(9):830–839. <https://doi.org/10.1056/NEJMoa1600693>.
  - 16 Valbert F, Wolf E, Schewe K, et al. Cost of human immunodeficiency virus (HIV) and determinants of healthcare costs in HIV-infected treatment-naïve patients initiated on antiretroviral therapy in Germany: experiences of the PROPHET study. *Value Health*. 2020;23(10):1324–1331. <https://doi.org/10.1016/j.jval.2020.04.1836>.
  - 17 Sullivan AK, Raben D, Reekie J, et al. Feasibility and effectiveness of indicator condition-guided testing for HIV: results from HIDES I (HIV indicator diseases across Europe study). *PLoS One*. 2013;8(1):e52845. <https://doi.org/10.1371/journal.pone.0052845>.
  - 18 European Centre for Disease Prevention and Control. *Public health guidance on HIV, hepatitis B and C test-ing in the EU/EFA – an integrated approach*. 2018.
  - 19 Jordans CCE, Vasylyev M, Rae C, et al. National medical specialty guidelines of HIV indicator conditions in Europe lack adequate HIV testing recommendations: a systematic guideline review. *Euro Surveill*. 2022;27(48). <https://doi.org/10.2807/1560-7917.ES.2022.27.48.2200338>.
  - 20 HIV in Europe. *HIV indicator conditions: guidance for implementing HIV testing in adults in health care settings*; 2023. Available from: <https://www.eurotest.org/media/0ymdzdvu/guidancepdf.pdf>.
  - 21 Raben D, Sullivan A. Implementation of indicator condition guided HIV testing still lagging behind the evidence. *eClinicalMedicine*. 2021;36:100918. <https://doi.org/10.1016/j.eclinm.2021.100918>.
  - 22 Bogers SJ, Hulstein SH, Schim van der Loeff MF, et al. Current evidence on the adoption of indicator condition guided testing for HIV in western countries: a systematic review and meta-analysis. *eClinicalMedicine*. 2021;35:100877. <https://doi.org/10.1016/j.eclinm.2021.100877>.
  - 23 Barbanotti D, Tincati C, Tavelli A, et al. HIV-indicator condition guided testing in a hospital setting. *Life*. 2023;13(4). <https://doi.org/10.3390/life13041014>.
  - 24 Vito A de, Colpani A, Mameli MS, et al. HIV infection indicator disease-based active case finding in a University hospital: results from the SHOT project. *Infect Dis Rep*. 2023;15(1):94–101. <https://doi.org/10.3390/idr15010010>.
  - 25 Agustí C, Montoliu A, Mascort J, et al. Missed opportunities for HIV testing of patients diagnosed with an indicator condition in primary care in Catalonia, Spain. *Sex Transm Infect*. 2016;92(5):387–392. <https://doi.org/10.1136/sextrans-2015-052328>.
  - 26 Carllander C, Marrone G, Brännström J, et al. Assessing cervical intraepithelial neoplasia as an indicator disease for HIV in a low endemic setting: a population-based register study. *BJOG An Int J Obstet Gynaecol*. 2017;124(11):1680–1687. <https://doi.org/10.1111/1471-0528.14614>.
  - 27 Matulionytė R, Jakobsen ML, Grecu VI, et al. Increased integrated testing for HIV, hepatitis C and sexually transmitted infections in health care facilities: results from the INTEGRATE Joint Action pilots in Lithuania, Romania and Spain. *BMC Infect Dis*. 2021;21(Suppl 2):845. <https://doi.org/10.1186/s12879-021-06537-2>.
  - 28 Girardi E, Scognamiglio P, Sciarrone MR, et al. High HIV prevalence in male patients with acute hepatitis A in the Rome metropolitan area, Italy 2002-2008. *J Hepatol*. 2011;54(6):1102–1106. <https://doi.org/10.1016/j.jhep.2010.09.024>.
  - 29 Menacho I, Sequeira E, Muns M, et al. Comparison of two HIV testing strategies in primary care centres: indicator-condition-guided testing vs. testing of those with non-indicator conditions. *HIV Med*. 2013;14(Suppl 3):33–37. <https://doi.org/10.1111/hiv.12064>.
  - 30 Omland LH, Legarth R, Ahlström MG, Sørensen HT, Obel N. Five-year risk of HIV diagnosis subsequent to 147 hospital-based indicator diseases: a Danish nationwide population-based cohort study. *Clin Epidemiol*. 2016;8:333–340. <https://doi.org/10.2147/CLEP.S101288>.
  - 31 Peck L, Ferenczi E, Burns F, Cosgrove C, Brown M. Barriers to targeted HIV testing on an acute admissions unit: evaluation of the UK guideline. *QJM*. 2010;103(3):147–151. <https://doi.org/10.1093/qjmed/hcp185>.
  - 32 Raben D, Mocroft A, Rayment M, et al. Auditing HIV testing rates across Europe: results from the HIDES 2 study. *PLoS One*. 2015;10(11):e0140845. <https://doi.org/10.1371/journal.pone.0140845>.
  - 33 Raben D, Sullivan AK, Mocroft A, et al. Improving the evidence for indicator condition guided HIV testing in Europe: results from the HIDES II Study–2012–2015. *PLoS One*. 2019;14(8):e0220108. <https://doi.org/10.1371/journal.pone.0220108>.
  - 34 Statista Research Department. Anzahl der Mitglieder und Versicherten in der GKV und PKV bis 2023 [February 21, 2024]; Available from: <https://de.statista.com/statistik/daten/studie/155823/umfrage/gkv-pkv-mitglieder-und-versichertenzahl-im-vergleich/>.
  - 35 Ganssen FM. Health economic evaluations based on routine data in Germany: a systematic review. *BMC Health Serv Res*. 2018;18(1):268. <https://doi.org/10.1186/s12913-018-3080-3>.
  - 36 Schubert I, Köster I, Küpper-Nybelen J, Ihle P. Versorgungsforschung mit GKV-Routinedaten. Nutzungsmöglichkeiten versichertenbezogener Krankenkassendaten für Fragestellungen der Versorgungsforschung. *Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz*. 2008;51(10):1095–1105. <https://doi.org/10.1007/s00103-008-0644-0>.
  - 37 Valbert F, Koppe U, Schmidt D, et al. Optimization of HIV testing services in Germany using HIV indicator diseases: study protocol of the HeLP study. *Arch Public Health*. 2023;81(159). <https://doi.org/10.1186/s13690-023-01161-9>.
  - 38 Crepaz N, Song R, Lyss SB, Hall HI. Estimated time from HIV infection to diagnosis and diagnosis to first viral suppression during 2014–2018. *AIDS*. 2021;35(13):2181–2190.
  - 39 Davidson R, MacKinnon JG. *Econometric reviews: bootstrap tests: how many bootstraps?* Taylor & Francis online 2007. <https://doi.org/10.1080/07474930008800459>.
  - 40 IBM. IBM SPSS bootstrapping 29 [June 13, 2023]; Available from: [https://www.ibm.com/docs/en/SSLVMB\\_29.0.0/pdf/IBM\\_SPSS\\_Bootstrapping.pdf](https://www.ibm.com/docs/en/SSLVMB_29.0.0/pdf/IBM_SPSS_Bootstrapping.pdf).
  - 41 Preacher KJ, Hayes AF. Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behav Res Methods*. 2008;40(3):879–891. <https://doi.org/10.3758/brm.40.3.879>.
  - 42 Wilcox RR. *Fundamentals of modern statistical methods: substantially improving power and accuracy*. 2nd ed. New York, NY, Heidelberg: Springer; 2010.
  - 43 Centers for Disease Control and Prevention. *Appendix A–AIDS-defining conditions*; 2024. Available from: <https://www.cdc.gov/mmwr/preview/mmwrhtml/rr5710a2.htm>.
  - 44 Vliegthart-Jongbloed KJ, Vasylyev M, Jordans CCE, et al. Systematic review: strategies for improving HIV testing and detection

- rates in European hospitals. *Microorganisms*. 2024;12(2). <https://doi.org/10.3390/microorganisms12020254>.
- 45 Bogers S, Zimmermann H, Ndong A, et al. Mapping hematologists' HIV testing behavior among lymphoma patients-A mixed-methods study. *PLoS One*. 2023;18(1):e0279958. <https://doi.org/10.1371/journal.pone.0279958>.
- 46 Pöge K, Spurgat C, Hamm J, Koppe U. *Forschungsbericht zum Projekt "Sexuelle Gesundheit und HIV/STI in trans und nicht-binären Communities"*. 2023.
- 47 Hoffmann F, Koller D. Verschiedene Regionen, verschiedene Versichertenpopulationen? Soziodemografische und gesundheitsbezogene Unterschiede zwischen Krankenkassen. *Gesundheitswesen*. 2017;79(1):e1–e9. <https://doi.org/10.1055/s-0035-1564074>.
- 48 Hoffmann F, Andersohn F, Giersiepen K, Scharnetzky E, Garbe E. Validierung von Sekundärdaten. Grenzen und Möglichkeiten. *Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz*. 2008;51(10):1118–1126. <https://doi.org/10.1007/s00103-008-0646-y>.
- 49 Schubert I, Ihle P, Köster I. Interne Validierung von Diagnosen in GKV-Routinedaten: Konzeption mit Beispielen und Falldefinition. *Gesundheitswesen*. 2010;72(6):316–322. <https://doi.org/10.1055/s-0030-1249688>.