Check for updates

SYSTEMATIC REVIEW

REVISED Sero-epidemiological studies of pathogen burden: A scoping review

[version 3; peer review: 3 approved]

Anne Suffel[1], Danny Turton[1], Kathryin E Mansfield[2], Alexander J Mentzer[3,4], Naomi E Allen[5,6], Aisha Babi[7,8], David Burgner[9], Tim Waterboer[7], Charlotte Warren-Gash[1]

[1]Department of Non-Communicable Disease Epidemiology, London School of Hygiene & Tropical Medicine, London, England, UK
[2]School of Health and Care Sciences, University of Lincoln, Lincoln, England, UK
[3]Centre for Human Genetics, University of Oxford, Oxford, England, UK
[4]Chinese Academy of Medical Science (CAMS) Oxford Institute (COI), University of Oxford, Oxford, England, UK
[5]Nuffield Department of Population Health, University of Oxford, Oxford, England, UK
[6]UK Biobank, Stockport, England, UK
[7]German Cancer Research Center, Heidelberg, Baden-Württemberg, Germany
[8]Faculty of Biosciences, Universitat Heidelberg, Heidelberg, Baden-Württemberg, Germany
[9]Royal Children's Hospital, Murdoch Children's Research Institute, Parkville, Victoria, Australia

v3 **First published:** 26 Mar 2026, **11**:192
https://doi.org/10.12688/wellcomeopenres.26070.1
Second version: 09 Jun 2026, **11**:192
https://doi.org/10.12688/wellcomeopenres.26070.2
Latest published: 11 Jul 2026, **11**:192
https://doi.org/10.12688/wellcomeopenres.26070.3

Abstract

Background

Methodological advances in serological testing have made it feasible to quantify antibodies to multiple pathogens simultaneously. This has been used to study the impact of pathogen burden, i.e., cumulative lifetime exposure to persistent pathogens, on health and mortality outcomes. However, definitions for pathogen burden using serological data vary considerably. It is unclear how to best combine serological data for multiple pathogens for exposure/outcome definitions in epidemiological studies.

Methods

Following PRISMA-ScR guidelines, we conducted a scoping review of all relevant publications from MEDLINE, Embase, and Scopus until May 1st, 2025, with no language restrictions. We identified epidemiological studies defining pathogen burden, and which included at least three different pathogens.

Open Peer Review

Approval Status ✓ ✓ ✓

	1	2	3
version 3 (revision) 11 Jul 2026	✓ view	✓ view	✓ view
version 2 (revision) 09 Jun 2026			
version 1 26 Mar 2026	? view	? view	

1. **Gary Lin**, Johns Hopkins University Applied Physics Laboratory, Laurel, USA
2. **Timothy R Julian**, Eawag Swiss Federal Institute of Aquatic Science and Technology, Dübendorf, Switzerland
3. **Andrea Haselbeck**, The University of Edinburgh, Edinburgh, UK

Results

We included a total of 74 studies from 19 countries (published 2000–2025). Serology data from 50 different pathogens were analysed across all studies, with pathogen burden measured by combining a range of 3–17 pathogens across studies into a single metric. Approaches included counts of seropositive result, quantiles of antibody titres, and more complex statistical analyses to group pathogens. No study gave a clinical/scientific rationale for the approaches chosen.

Conclusions

Various analytical approaches have been used to define pathogen burden but there was limited justification for the chosen methodological approaches. More integrated approaches are needed for combining scientific and clinical knowledge of the interaction across different pathogens and epidemiological evidence into a meaningful measure of pathogen burden.

Plain Language summary

Blood tests can show which infections a person has had in the past. These tests can check for many infections at the same time. However, there is no clear agreement on how to combine these results into one overall measure of a person's past exposure to infections (pathogen burden).

We reviewed published studies to see how researchers have done this. We found 74 population health studies published between 2000 and 2025 that combined blood test results for at least three different infections. Across these studies, researchers looked at more than 50 infections and combined between 3 and 17 infections into a single score. Most studies simply counted how many infections a person had, while others used more complicated methods.

None of the studies clearly explained why they chose their method. Better ways are needed to measure exposure to multiple infections. These methods should be based on what we know about how infections affect the body and health, so the results are easier to understand and more useful.

Keywords

pathogen burden; serology, infections

Any reports and responses or comments on the article can be found at the end of the article.

Corresponding author: Anne Suffel (anne.suffel@lshtm.ac.uk)

Author roles: Suffel A: Conceptualization, Data Curation, Formal Analysis, Methodology, Writing – Original Draft Preparation, Writing – Review & Editing; **Turton D**: Data Curation, Writing – Review & Editing; **Mansfield KE**: Conceptualization, Writing – Review & Editing; **Mentzer AJ**: Conceptualization, Writing – Review & Editing; **Allen NE**: Conceptualization, Writing – Review & Editing; **Babi A**: Data Curation, Writing – Review & Editing; **Burgner D**: Conceptualization, Writing – Review & Editing; **Waterboer T**: Conceptualization, Writing – Review & Editing; **Warren-Gash C**: Conceptualization, Funding Acquisition, Methodology, Supervision, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: This work was supported, in part, by Open Philanthropy. CWG is supported by a Wellcome Career Development Award (225868/Z/22/Z). AJM is supported through the Nuffield Department of Medicine Career Development Scheme and receives funding from the Chinese Academy of Medical Sciences Innovation Fund for Medical Science (2024-I2M-2–001–1). AJM is a contributor to intellectual property licensed by Oxford University Innovation to AstraZeneca.
The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Copyright: © 2026 Suffel A *et al*. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Suffel A, Turton D, Mansfield KE *et al.* **Sero-epidemiological studies of pathogen burden: A scoping review [version 3; peer review: 3 approved]** Wellcome Open Research 2026, **11**:192 https://doi.org/10.12688/wellcomeopenres.26070.3

First published: 26 Mar 2026, **11**:192 https://doi.org/10.12688/wellcomeopenres.26070.1

REVISED **Amendments from Version 2**

We have updated the manuscript following the suggestions and comments of the second reviewer. I have updated Table 1 after finding a typo for one the studies.

Any further responses from the reviewers can be found at the end of the article

Introduction

Serological data are a useful tool for assessing an individual's past exposure and immunity to pathogens. Various methods have been developed to examine serum or plasma for the presence of pathogen-specific antibodies. Laboratory techniques include functional assays (e.g., plaque or pseudovirus reduction neutralization assays [PRNT/pVNT]), and binding assays (including enzyme-linked immunosorbent assays [ELISA], fluorescent bead-based assays, or fluorescent micro-array-based assays)[1] or Phage Immunoprecipitation sequencing (PhIP-seq) technology.[2]

Serology data provides a variety of advantages for epidemiological research. They are an objective measure that can detect asymptomatic or subclinical infections, and past infection[3] that does not rely on self-recall. In combination with other data sources, such as electronic health records and questionnaires, serology data can be used to investigate more complex and longitudinal interactions between various pathogens and chronic conditions.[4,5]

With the development of multiplex assays, it is possible to study previous exposure to multiple pathogens simultaneously from small blood samples, including archived dried blood spots. It uses beads coated with specific antibodies of binding sites for the pathogen which can then be read though flow cytometry using the fluorescent signature of the beads.[6] An increasing number of studies have investigated links between infections and various non-communicable diseases.[7–9] Well-described associations include: the causal role of certain strains of human papillomaviruses and cervical and oropharyngeal cancer[10]; *Helicobacter pylori* and gastric cancer[11]; and Epstein Barr virus (EBV) and Burkitt's lymphoma[12]: EBV and multiple sclerosis[13,14] and hepatitis viruses, and liver cirrhosis and hepatocellular carcinoma.[15]

Often, more than one pathogen is linked to a chronic condition. For example, both *Chlamydophilia pneumoniae*, cytomegalovirus (CMV) and other herpes viruses, as well as several oral infections, have putatively been associated with atherosclerosis.[9,16] The reactivation of the varicella zoster virus[17] and previous herpes simplex virus (HSV) 1 and 2, have both been associated with increased dementia risk.[18]

A recent study found evidence for an association between multiple pathogens with non-communicable diseases, including replication of an association between cytomegalovirus and ulcerative colitis.[8] However, more research is needed on how multiple pathogens interact with each other to influence the pathogenesis of other health conditions. Understanding pathogen-disease interactions could help to improve prevention through targeted vaccination or treatment of infections. For example, infections in childhood have already shown to be associated with an increased risk of cardiovascular disease and cardiovascular risk factors in adults, offering opportunities for earlier prevention.[19] An optimal definition of pathogen burden could improve studies investigating the links between infection history and non-communicable disease outcomes. Furthermore, multiplex techniques could be used to detect viral epitopes that might not be required for defining a seropositive exposure status but are potentially associated with disease progression ("serolomics").[20]

We aimed to summarise how definitions of the burden of serologically defined infections have been conceptualised and operationalised in seroepidemiological studies to inform future research using measures of pathogen burden.

Methods

We conducted our scoping review following guidelines outlined by the Preferred Reporting Items for Systematic Reviews and Meta-Analyses for Scoping Reviews (PRISMA-ScR[21]). We sought to describe all methods used in epidemiological research to combine measurements of at least three different pathogens into a concept of pathogen burden (defined as a measure of cumulative exposure to at least three serologically defined pathogens).

Review question

How are multiple serological measures of chronic or latent infections combined into a single concept, such as pathogen burden or infectious burden, for epidemiological research questions?

Inclusion criteria
We included all peer-reviewed studies with an epidemiological research question in a human population that defined a composite measure of antibodies to at least three different pathogens. Previous infection with a pathogen had to be ascertained through serological data generated using any assay and at least three pathogens had to be combined into a single outcome or exposure. We chose studies with at least three pathogens to capture studies with more complex exposure/outcome definitions of multiple infections. There was no restriction on type of infection, country, year, or language of publication.

Exclusion criteria
We excluded reviews, comments, opinion pieces, conference abstracts, study protocols, errata, grey literature, and retracted studies. Further, we excluded studies with less than three pathogens or if the prevalence or incidence was reported separately for each pathogen, rather than combined into a composite measure of pathogen burden. Studies with clinical, or other, definitions of infection but no serological data were excluded.

Search strategy and study selection
For this scoping review, a literature search was conducted in MEDLINE, Embase, and Scopus from database inception to the 1st of May 2025. The search strategy combined the concept of pathogen/infection burden with various kinds of infections and limited the study results to a human population. The specific infections in the search strategy focused on common persistent pathogens measured in studies investigating links between infections and non-communicable diseases (e.g., cardiovascular disease, cancer or dementia). The full search strategy can be found in the Supplement (S1[22]) refer to extended data. Titles were subsequently imported into Rayyan (web-based tool to screen and manage references during systematic reviews) for removal of duplicates, further screening and review. Two authors independently screened article titles and abstracts for inclusion based on the above criteria (AS and DT). All disagreements were resolved between the two authors (third author available if necessary). The full text of articles was subsequently reviewed for inclusion by two independent reviewers (AS and DT). Any discrepancies were discussed between the two authors for resolution.

Data extraction and synthesis
The data from included articles were extracted by AS. Data from a 10% sample of the included studies were extracted by a second author (CWG) to ensure quality of the data extraction process. We extracted general study information, operationalisation of pathogen burden, pathogens included, serology methods used to assess past infection, and rationale for the choice of pathogen burden measure (Supplementary Text S1[22] for more detail refer to extended data). If the location of the study was missing, we supplemented the location with the country of the institution the majority of authors were affiliated with. Extracted data were summarised narratively, and in tables and figures.

We did not conduct a risk of bias assessment of the studies as this review focused on the different conceptualisations of pathogen burden rather than the epidemiological questions addressed in the included studies.

Role of the funding source
The funder had no role in study design, data collection, analysis, interpretation of the data, nor in writing the report and the decision to submit the paper for publication.

Results
We retrieved 6,868 articles across all three databases (Figure 1). Following removal of duplicates, 3,561 articles were included in abstract screening, 100 in full text screening, with 74 studies included in the final review (Table 1, refer to data repository[22]).

Summary of all studies
The 74 included studies were conducted in North America, Europe, or Asia. Most studies were conducted in the US (n studies = 23, 31.1%[23–45]), followed by Germany (n = 6, 8.1%[46–51]) and China (n = 6, 8.1%[52–57]). Six publications did not explicitly state the country where the study was conducted.[58–63] Forty-four studies included cohorts of adults in a community setting (59.5%[23,26–33,35,37–45,51,63–86]), 28 study populations were recruited adults from hospitals (37.8%[24,34,46–50,52,54–60,62,87–98]), and two studies included children (2.7%[99,100]).

The reporting of the pathogen burden measurement methods varied across different studies. Some studies (n = 34, 45.9%) reported details of the assays used including information on their sensitivity, specificity, and threshold values of measured antibody levels to define seropositivity. The most commonly used assay was Enzyme-Linked Immunosorbent Assay [ELISA] (n = 46, 62.2%, explicitly stated[23,24,26,27,29–31,33–40,42,44,46–50,54–57,59,62–64,68,72,76–80,82,83,85,89–92,95,101]).

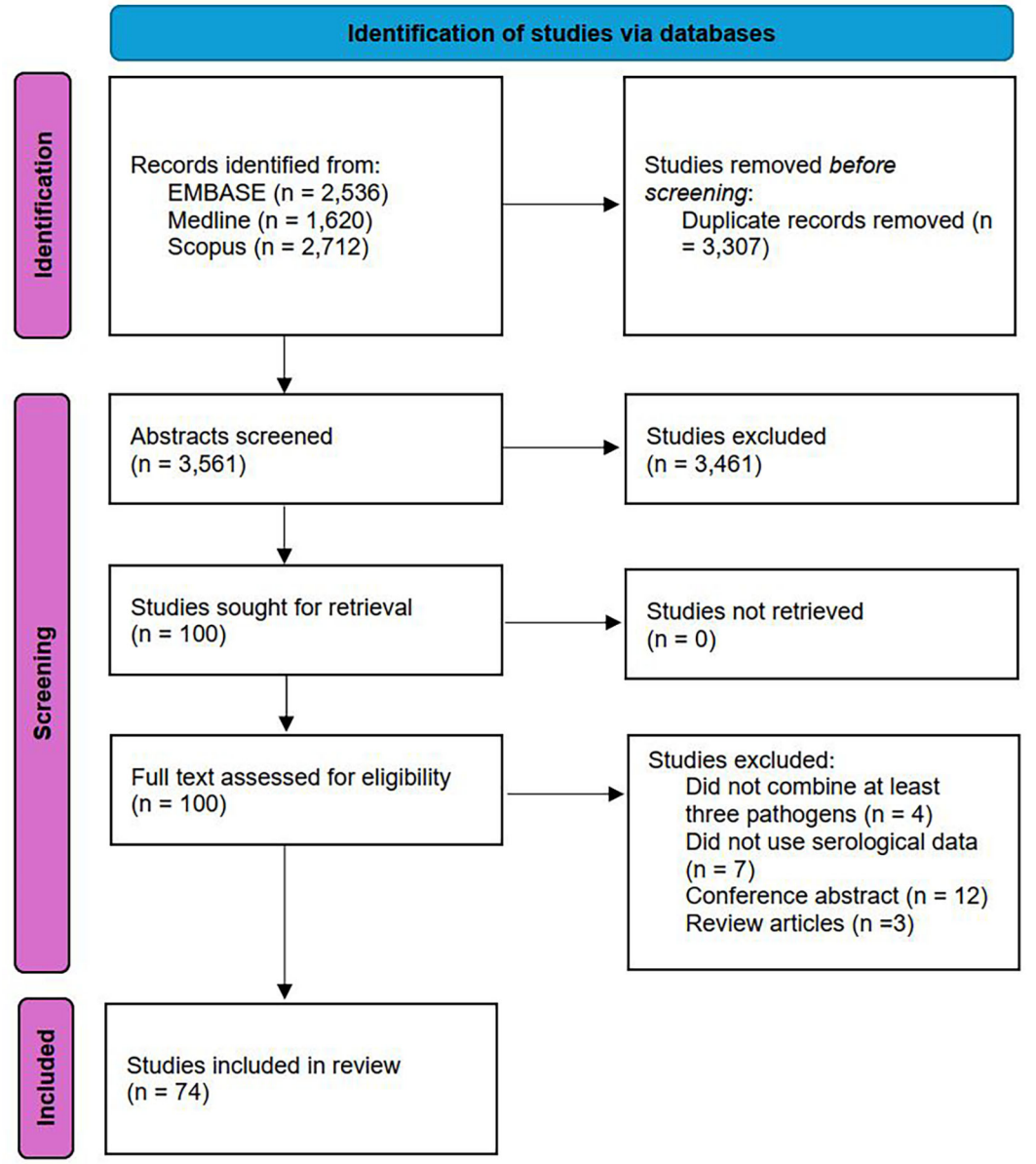


Figure 1. Number of articles retrieved and excluded at different stages of the review.

However, there were some studies that only referred to the commercial test kits (n = 3, 4.1%[25,75,99]) or just stated immunoassays were used (n = 1, 1.4%[66]). A full description of the assays, antibodies measured, population prevalence and cut-off values to define seropositivity can be found in the supplementary material.

Fifty pathogens were included across all studies, with the most studied pathogens being CMV (n = 64, 86.5%[23–31,33–44,46,48,52,54,56–60,62–65,67–70,72,73,75–84,88–92,95,97,98,100,102,103]), herpes simplex virus 1 (HSV-1, n = 63, 85.1%[23–31,33–46,48,54,56,57,59,62–66,68,69,72–74,76–85,87,89,90,92,95,97–99,100,102–104]) and *H. pylori* (Hp, n = 54, 73.0%[24,26–31,33–38,40,42–44,46,48,51–54,56,58–60,62–68,70,72,73,75–77,79,82,83,88–90,92,95,97–99,102–104]) (Figure 2). The majority of pathogens were viruses (n = 268, 62.8% of measured pathogens), followed by bacteria (n = 142, 33.3%); only 17 (4.0%) measurements included other micro-organisms such as alveolates, protozoans, or helminths (Figure 3). The number of pathogens combined into a putative 'pathogen burden' ranged from 3 to 17 across the studies. The pathogens most frequently combined into a single burden measure were CMV, HSV-1, Hp, and *Chlamydophila pneumoniae* (Cp); this combination measure was analysed in nine different studies (Figure 4). Six studies of these nine studies using a combined CMV/HSV-1/Hp/Cp measure also included serological data on herpes simplex virus 2 (HSV-2).

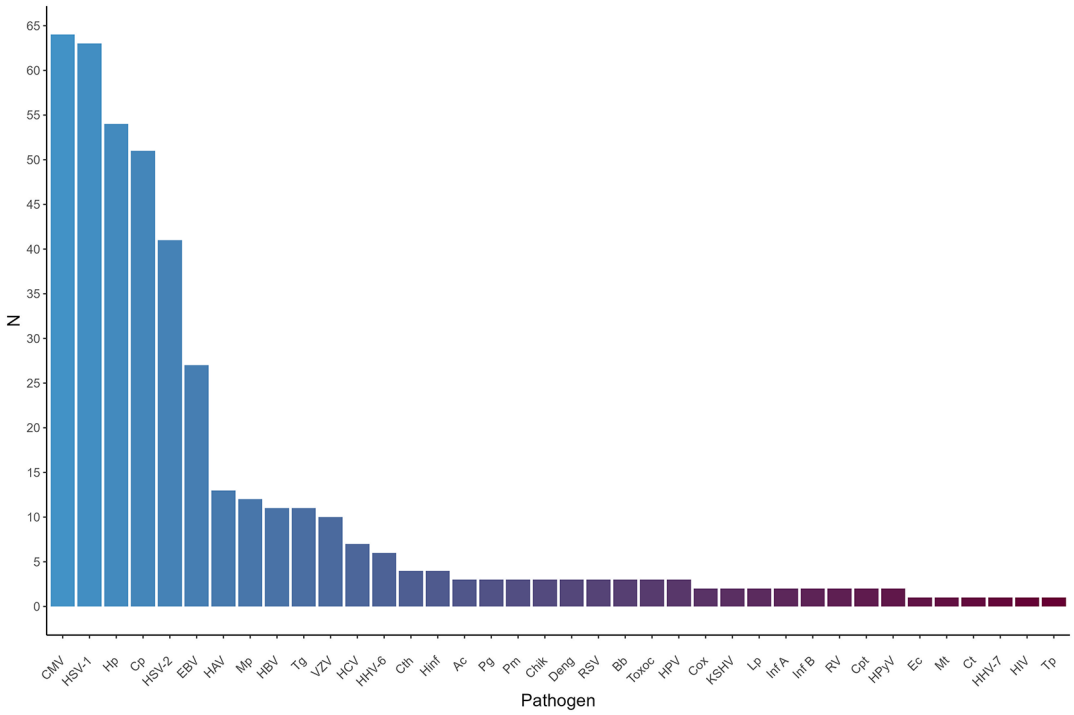


Figure 2. Frequency of different pathogens included in a pathogen burden measure across all studies. CMV: Cytomegalovirus. HSV-1: Herpes simplex virus 1. Hp: *Helicobacter pylori.* Cp: *Chlamydia pneumoniae.* HSV-2: Herpes simplex virus 2. EBV: Epstein-Barr virus. HAV: Hepatitis A virus. Mp: *Mycoplasma pneumoniae.* HBV: Hepatitis B virus. Tg: Toxoplasma gondii. VZV: Varicella zoster virus. HHV-6: Human herpesvirus 6 (A, B). HCV: Hepatitis C virus. Hinf: *Haemophilus influenzae.* Ac: *Aggregatibacter actinomycetemcomitans.* Pg: *Porphyromonas gingivalis.* Inf A: Influenza A virus. Inf B: Influenza B virus. Pm: Plasmodium malariae. HPV: Human papillomavirus (6, 11, 16, 18). Bb: *Borrelia burgdorferi.* Cox: Coxsackievirus B3. Ch t: *Chlamydia trachomatis.* RV: Rubella virus. Clt: *Clostridium tetani.* Chik: Chikungunya virus. Deng: Dengue virus. RSV: Respiratory syncytial virus. Toxoc: Toxocara canis and Toxocara cati. Ec: *Escherichia coli.* Lp: *Legionella pneumophila.* Mt: *Mycobacterium tuberculosis.* HPyV: Human polyomaviruses (6, 7, 9, 10). HHV-7: Human herpesvirus 7. KSHV: Kaposi's Sarcoma-Associated Herpesvirus. HIV: human immunodeficiency virus. Syphilis: *Treponema pallidum.*

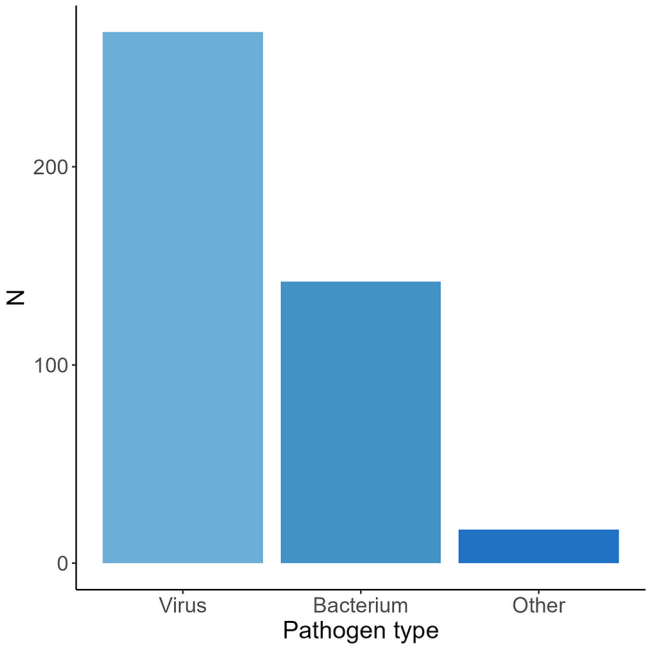


Figure 3. Types of pathogens included in pathogen burden concept across all included studies. The Other category includes: protozoae, helminths, and an alveolate.

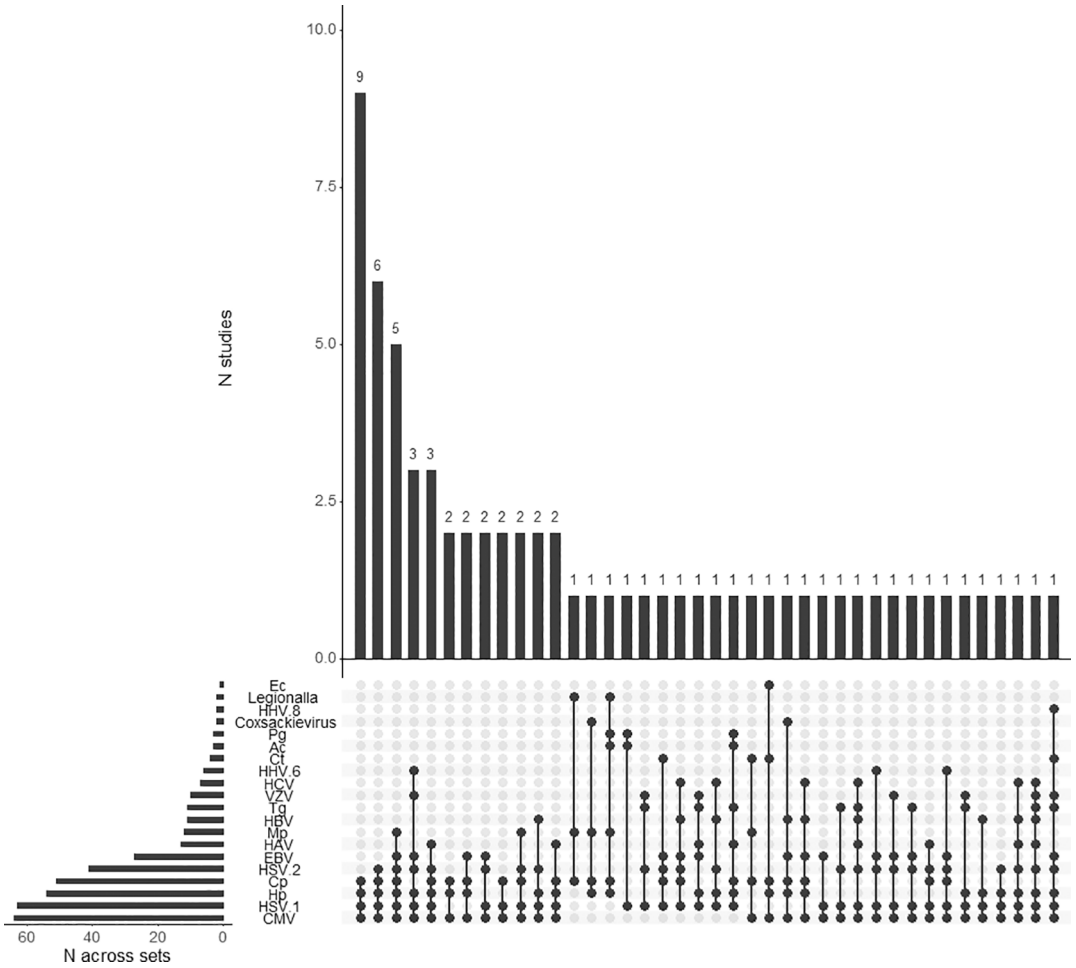


Figure 4. Frequency of combinations for the 20 most common pathogens included into pathogen burden across studies. N combinations states how many studies combined the pathogens under each bar, whereas N across sets indicates the frequency of each single pathogen across different studies. CMV: Cytomegalovirus. HSV-1: Herpes simplex virus 1. Hp: *Helicobacter pylori.* Cp: *Chlamydia pneumoniae.* HSV-2: Herpes simplex virus 2. EBV: Epstein-Barr virus. HAV: Hepatitis A virus. Mp: *Mycoplasma pneumoniae.* HBV: Hepatitis B virus. Tg: Toxoplasma gondii. VZV: Varicella zoster virus. HHV-6: Human herpesvirus 6 (A, B). HCV: Hepatitis C virus. Ct: *Chlamydia trachomatis.* H inf: *Haemophilus influenzae.* RSV: Respiratory syncytial virus. Ac: *Aggregatibacter actinomycetemcomitans.* Pg: *Porphyromonas gingivalis.* Inf A: Influenza A virus. Inf B: Influenza B virus.

The prevalence of each type of pathogen varied widely across different studies. For example, HSV-2 prevalence ranged from 1.5% in a paediatric study population[100] (Greece) to 74.5% in a population of individuals receiving haemodialysis[98] (country not stated).

Most studies (n = 71, 95.9%) included pathogen burden as an exposure, with the most studied outcomes being cognitive function, indicators of cardiovascular health, and inflammatory markers. Only three studies investigated pathogen burden as an outcome, investigating differences in socioeconomic and ethnic disparities in relation to prevalence of infections.[28,41,105]

Definitions of pathogen burden

Pathogen burden was assessed across different studies using various methods including using pathogen counts, combinations of seropositive status across specific pathogens, and antibody levels.

Pathogen count

The majority of papers (n = 54, 73.0%,[23–26,33–35,38,42,43,46,48–50,52–56,58–60,62–64,66–68,74,78,80–84,87–91,95–98,100,106]) counted the number of pathogens each individual tested positive for and used the summed score as an exposure variable

(53/54 studies). Depending on the prevalence of different pathogens, the baseline reference category for statistical analyses varied across studies, with the reference category usually comprised of scores of 0–2 or 0–3 positive counts (e.g.,[102,107]). Some studies classified pathogen burden into low, medium or high exposure based on the absolute score (e.g.,[34,66]). One study dichotomised pathogen score and defined the cut-off value based on its ability to predict the outcome using a receiver operator characteristic curve (ROC) curve analysis.[57] Two studies used the proportion of pathogens each individual tested positive as a proportion of all pathogens assayed.[41,65] One study created a summary pathogen load coefficient that divided the sum of positive test results (pathogen load) by 1.6, which was the average summary pathogen load of the healthy comparator group.[87,93]

Seven studies (n = 7, 9.5%,[35,55,56,65,80,99,100]) differentiated by the types of pathogens and created separate scores such as positive count for herpesviruses and polyomaviruses[100] or using a separate score for neurotropic viruses.[65]

Combinations of seropositive status

A single study with four different pathogens used all potential combinations of positive or negative serology tests across the four pathogens as their exposure variable.[108]

Use of antibody levels

Thirteen studies used antibody titres as an exposure variable (Table 1, refer to data repository[22,27,29,45,51,65,72,73,75–77,85,92,107]). The majority (n = 11, 14.8%,[27,29,45,51,65,72,73,75–77,85,92,107]) of these studies divided antibody levels into thirds or fourths to define a categorical exposure variable for each pathogen separately.[65,109] This individual-pathogen categorical exposure approach was used as an additional analysis for specific pathogens in addition to a composite measure. One study divided antibody levels for each pathogen into fourths and then assigned between 1–4 points for each pathogen depending on the category of each individual's antibody titre level, using the total score as their exposure variable.[85] Another study used the average z-scored IgG antibody level across all five pathogens for all individuals who tested positive for at least one pathogen.[31] No study used a serological correlate of protection against a pathogen.

Other pathogen burden definitions

Six studies (n = 6, 8.1%,[30,37,39,44,86,101]) used weighted regression coefficients as a pathogen score (i.e. all exposure and confounder variables were regressed on the outcome of interest, and a pathogen score calculated by summing the regression coefficients of each positive pathogen).[30,39,101,110] One study used MIMIC models (multiple indicators, multiple causes), a type of structural equation modelling that captures the underlying statistical clusters of pathogens.[28] In this approach, the latent variable representing infection burden was modelled as influencing individual infection outcomes, and pathogen burden was expressed as a function of the observed predictors. Model fit was evaluated using the comparative fit index (CFI) and root mean square error of approximation (RMSEA). One study[40] used latent class analysis to identify common clusters of seropositivity status across different pathogens. They determined the number of classes by comparing different statistical models using a combination of likelihood-ratio value (G2), Akaike information criteria (AIC), Bayesian information criterion (BIC), sample size-adjusted BIC, log-likelihood and entropy. After the optimal number of classes was determined, they assigned a class to each individual based on their posterior probability of most likely class membership.

Rationale behind pathogen scores

There was no scientific rationale described for the approach taken in any study. Some studies (n = 5, 6.8%,[30,37,39,44,101]) referred to a review paper.[16] However, this paper only states that there is no consensus on defining pathogen burden and no clinical consensus about which pathogens should be included for pathogen burden score. A Russian study references a book chapter for their summary pathogen load definition that could not be verified by the authors of this review.[93]

Discussion

Our scoping review and previous related reviews in the literature, demonstrate that there is no consensus on how to define pathogen burden, nor for which pathogens to include in an index or score for pathogen burden.[16] We showed that there are major differences across studies in the choice of pathogens measured and how they are combined into a composite pathogen burden score. It remains unclear whether the wide range of pathogens included in studies reflected hypothesised associations with outcomes or highly prevalent pathogens in the study population, and whether it was also influenced by availability and cost of serological assays.

The approaches taken to define pathogen burden varied widely, ranging from simpler methods such as scores based on pathogen-specific seropositivity, to more complex methods such as latent class analysis and regression-based scores.

Using a summary score for different pathogens means that different pathogens are considered to contribute equally to the pathogenesis of a specific disease in the statistical model. However, this approach ignores potential synergistic effects between different pathogens as well unique pathological mechanisms. Only one study included in our review tried to address this issue by looking at the different combinations of positive serostatus across four pathogens. However, this approach may not always be feasible as the number of serostatus combinations increase exponentially with the number of pathogens.[108]

Approaches such as the weighted pathogen index, which reported an association with the risk of stroke,[111] are difficult to interpret as the presented hazard ratios are in relation to the increase by one standard deviation of the index. The clinical significance of such a measure is unclear (e.g., what serostatus in an individual does this increased stroke risk correspond to?). Similar to using a simpler index, this approach also cannot differentiate potential synergistic and differing effects of different pathogens.

Latent class analysis provides a more nuanced approach to classifying different pathogens into exposure groups as it accounts for their relationships with the outcome, and other confounders. However, this method requires a sufficiently large sample size and may be limited when dealing with low prevalence infections. For example, in one study, HIV status did not contribute to class membership attribution owing to its low prevalence,[40] meaning that information from a potentially important infection was lost. This loss of information is particularly pertinent for HIV as it has been associated with higher risk of several non-communicable diseases including cardiovascular diseases, cancer, and chronic respiratory diseases.[112,113] Furthermore, these methods do not acknowledge underlying biological hypotheses, e.g., tropism of the pathogen, and direct or indirect effects of the pathogens.

Using data on antibody titres may reflect the severity of the past infection, viral load, immunogenetics of the host and time since infection. However, this is highly dependent on the pathogen and antibody measured. Higher levels of haemagglutinin inhibition antibodies for influenza have been associated with febrile illness from influenza[114] and the antibody responses one year after SARS-CoV-2 infection was higher in individuals with more severe infections.[115] However, researchers found no association between the severity of mononucleosis and levels of specific EBV protein and neutralising antibodies after infection.[116] Higher antibody levels may also reflect a reactivation of a virus, for example, higher HSV-1 titres in individuals with more frequent cold sores.[117] Other pathogens such as *Neisseria gonorrhoeae* are difficult to prove due to great variability in antigens and often weak antibody response to infections.[118] It remains unclear how different levels of antibody responses are linked to the mechanisms involved in the pathogenesis of non-communicable diseases.

With the increasing availability of higher-volume and lower-cost testing assays for multiple pathogens, a challenge arises regarding how best to use these rich data to study combined effects of multiple pathogens on multiple disease outcomes. The choice of included pathogens in any measure of pathogen burden is highly dependent on the research question, study population of interest, and available laboratory technology for testing. Even within a multiplex assay, there may be different validities for different pathogens, which may be challenging for overall pathogen index validation. The study population of interest presents an additional challenge as children may not show a serological correlate of infection despite symptomatic disease.[119]

Our scoping review included a wide range of different studies with no exclusion criteria related to country of study or publication language. Our search terms covered a wide range of pathogens often linked to non-communicable diseases, including commonly studied viruses and general terms for infections. However, our approach might have overlooked some rare or neglected pathogens. We also focused out search strategy on specified pathogens that have been commonly studies in relationship to non-communicable diseases and measures in large population-based cohorts such as the UK Biobank to inform future research using such resources. We did not assess risk of bias in the included studies as the main interest of this review was to assess the scope of methods used to conceptualise pathogen burden rather than to assess the quality of the epidemiological research.

Overall, a better integration of knowledge of the infection pathogenicity, transmission route, and epidemiology, in relation to underlying cellular and genetic mechanisms of the disease (outcome) of interest, are essential for better operationalisation of pathogen burden in epidemiological research. No universal indicator of pathogen burden will be feasible as long there is no agreed set of pathogens that should be tested for in epidemiological research. As a consequence, some findings of studies using their own definition of pathogen burden may not be reproducible unless the same definition and method of defining pathogen burden is used. Future studies should include a clear justification for the measure of pathogen burden used as well as a clear description of both serological assays and statistical methods (see key recommendations).

Key recommendations for reporting pathogen burden in epidemiological studies:

- Clearly defined research question and explanation of the role of pathogen burden (e.g., outcome, exposure, covariate)
- List of pathogens included
- Justification for choice of pathogens included (e.g., relevance to study population)
- Reporting of type of assays, and thresholds used to define sero-positivity
- Any other supplementary methods used to measure pathogens (e.g., clinical data, surveys)
- Statistical methods used to aggregate serological data
- Justifications for the methods used (e.g., choice of assays, thresholds, statistical approach)
- Careful interpretations of findings in the context of the selected pathogens and research question
- Comment on the generalisability of findings

Ethics and consent

Ethical approval and consent were not required.

Data availability statement

Extended data and study protocol can be found in the supplementary material: https://doi.org/10.5281/zenodo.18669395.[22]

Data are available under the terms of the MIT License.

Acknowledgements

We thank Yu Wu for the screening and translation of the Chinese language paper.

References

1. Haselbeck AH, Im J, Prifti K, *et al.*: **Serology as a Tool to Assess Infectious Disease Landscapes and Guide Public Health Policy.** *Pathogens.* 2022; **11**(7). PubMed Abstract | Publisher Full Text | Free Full Text | Reference Source
2. Xu GJ, Kula T, Xu Q, *et al.*: **Comprehensive serological profiling of human populations using a synthetic human virome.** *Science (1979).* 2015; **348**(6239).
3. Menezes A, Razafimahatratra SL, Wariri O, *et al.*: **Strengthening serological studies: the need for greater geographical diversity, biobanking, and data-accessibility.** *Trends Microbiol.* 2025; **33**: 1155–1162. PubMed Abstract | Publisher Full Text
4. Elkind MSV: **Infectious burden: A new risk factor and treatment target for atherosclerosis.** *Infect. Disord. Drug Targets.* 2010; **10**(2): 84–90. Reference Source
5. Mentzer AJ, Brenner N, Allen N, *et al.*: **Identification of host–pathogen-disease relationships using a scalable multiplex serology platform in UK Biobank.** *Nat. Commun.* 2022; **13**(1): 1818. PubMed Abstract | Publisher Full Text | Free Full Text
6. Waterboer T, Sehr P, Michael KM, *et al.*: **Multiplex Human Papillomavirus Serology Based on In Situ–Purified Glutathione S-Transferase Fusion Proteins.** *Clin. Chem.* 2005; **51**(10): 1845–1853. PubMed Abstract | Publisher Full Text
7. Jartti T, Gern JE: **Role of viral infections in the development and exacerbation of asthma in children.** *J. Allergy Clin. Immunol.* 2017; **140**(4): 895–906. PubMed Abstract | Publisher Full Text | Free Full Text
8. Lape M, Schnell D, Parameswaran S, *et al.*: **A survey of pathogenic involvement in non-communicable human diseases.** *Commun. Med.* 2025; **5**(1): 242. Publisher Full Text
9. Ogoina D, Onyemelukwe GC: **The role of infections in the emergence of non-communicable diseases (NCDs): Compelling needs for novel strategies in the developing world.** *J. Infect. Public Health.* 2009; **2**(1): 14–29. PubMed Abstract | Publisher Full Text | Free Full Text
10. Saeed M, Faisal SM, Akhtar F, *et al.*: **Human Papillomavirus Induced Cervical and Oropharyngeal Cancers: From Mechanisms to Potential Immuno-therapeutic Strategies.** *Curr. Drug Metab.* 2020; **21**(3): 167–177. Publisher Full Text
11. Amieva M, Peek RM: **Pathobiology of Helicobacter pylori–Induced Gastric Cancer.** *Gastroenterology.* 2016; **150**(1): 64–78. PubMed Abstract | Publisher Full Text | Free Full Text
12. Shannon-Lowe C, Rickinson A: **The Global Landscape of EBV-Associated Tumors.** *Front. Oncol.* 2019; **9**.
13. Lanz TV, Brewer RC, Ho PP, *et al.*: **Clonally expanded B cells in multiple sclerosis bind EBV EBNA1 and GlialCAM.** *Nature.* 2022;

603(7900): 321–327.
PubMed Abstract | Publisher Full Text | Free Full Text

14. Bjornevik K, Cortese M, Healy BC, *et al.*: **Longitudinal analysis reveals high prevalence of Epstein-Barr virus associated with multiple sclerosis.** *Science (1979).* 2022; **375**(6578): 296–301.

15. Usuda D, Kaneoka Y, Ono R, *et al.*: **Current perspectives of viral hepatitis.** *World J. Gastroenterol.* 2024; **30**(18): 2402–2417.
PubMed Abstract | Publisher Full Text | Free Full Text

16. Elkind MSV: **Infectious burden: A new risk factor and treatment target for atherosclerosis.** *Infect. Disord. Drug Targets.* 2010; **10**(2): 84–90. Netherlands: Bentham Science Publishers.
Reference Source

17. Polisky V, Littmann M, Triastcyn A, *et al.*: **Varicella-zoster virus reactivation and the risk of dementia.** *Nat. Med.* 2025; **31**: 4172–4179.
PubMed Abstract | Publisher Full Text | Free Full Text

18. Araya K, Watson R, Khanipov K, *et al.*: **Increased risk of dementia associated with herpes simplex virus infections: Evidence from a retrospective cohort study using U.S. electronic health records.** *J. Alzheimer's Dis.* 2025; **104**(2): 393–402.
PubMed Abstract | Publisher Full Text | Free Full Text

19. Qanitha A, de Mol BA, Pabittei DR, *et al.*: **Infections in early life and premature acute coronary syndrome: A case-control study.** *Eur. J. Prev. Cardiol.* 2016; **23**(15): 1640–1648.

20. Butt J, Murugan R, Hippchen T, *et al.*: **From Multiplex Serology to Serolomics—A Novel Approach to the Antibody Response against the SARS-CoV-2 Proteome.** *Viruses.* 2021; **13**(5): 749.
PubMed Abstract | Publisher Full Text | Free Full Text

21. Tricco AC, Lillie E, Zarin W, *et al.*: **PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation.** *Ann. Intern. Med.* 2018; **169**(7): 467–473.
Publisher Full Text

22. Suffel A: **Supplementary material for "Sero-epidemiological studies of pathogen burden: A scoping review".** *Zenodo.* 2026 [cited 2026 Feb 19].
Reference Source

23. Zhu J, Quyyumi AA, Norman JE, *et al.*: **Effects of total pathogen burden on coronary artery disease risk and C-reactive protein levels.** *Am. J. Cardiol.* 2000; **85**(2): 140–146. United States.
Reference Source

24. Zhu J, Nieto FJ, Horne BD, *et al.*: **Prospective study of pathogen burden and risk of myocardial infarction or death.** *Circulation.* 2001; **103**(1): 45–51. United States.
Reference Source

25. Miller GE, Freedland KE, Duntley S, *et al.*: **Relation of depressive symptoms to C-reactive protein and pathogen burden (cytomegalovirus, herpes simplex virus, Epstein-Barr virus) in patients with earlier acute coronary syndromes.** *Am. J. Cardiol.* 2005; **95**(3): 317–321.
Publisher Full Text | Reference Source

26. Szklo M, Ding J, Tsai MY, *et al.*: **Individual pathogens, pathogen burden and markers of subclinical atherosclerosis: the Multi-Ethnic Study of Atherosclerosis.** *J. Cardiovasc. Med. (Hagerstown).* 2009; **10**(10): 747–751. United States.
Reference Source

27. Aiello AE, Diez-Roux A, Noone A-M, *et al.*: **Socioeconomic and psychosocial gradients in cardiovascular pathogen burden and immune response: the multi-ethnic study of atherosclerosis.** *Brain Behav. Immun.* 2009; **23**(5): 663–671. Netherlands.
Reference Source

28. Zajacova A, Dowd JB, Aiello AE: **Socioeconomic and race/ethnic patterns in persistent infection burden among U.S. adults.** *J. Gerontol. A Biol. Sci. Med. Sci.* 2009; **64**(2): 272–279. United States.
Reference Source

29. Nazmi A, Diez-Roux AV, Jenny NS, *et al.*: **The influence of persistent pathogens on circulating levels of inflammatory markers: a cross-sectional analysis from the Multi-Ethnic Study of Atherosclerosis.** *BMC Public Health.* 2010; **10**: 706. England.
Reference Source

30. Elkind MSV, Luna JM, Moon YP, *et al.*: **Infectious burden and carotid plaque thickness: the northern Manhattan study.** *Stroke.* 2010; **41**(3): e117–e122. United States.
Reference Source

31. Sealy-Jefferson S, Gillespie BW, Aiello AE, *et al.*: **Antibody levels to persistent pathogens and incident stroke in Mexican Americans.** *PLoS One.* 2013; **8**(6): e65959. United States.
Reference Source

32. Katan M, Moon YP, Paik MC, *et al.*: **Infectious burden and cognition: The northern manhattan study.** *Stroke.* 2011; **42**(3): e163. Lippincott Williams and Wilkins.
Reference Source

33. Simanek AM, Dowd JB, Zajacova A, *et al.*: **Unpacking the "black box" of total pathogen burden: is number or type of pathogens most predictive of all-cause mortality in the United States?.** *Epidemiol. Infect.* 2015; **143**(12): 2624–2634. England.
Reference Source

34. Miller PS, Evangelista LS, Giger JN, *et al.*: **Exhaustion, immuno-inflammation, and pathogen burden after cardiac surgery: an exploratory study.** *Eur. J. Cardiovasc. Nurs.* 2014; **13**(3): 211–220. England.
Reference Source

35. Rubicz R, Yolken R, Drigalenko E, *et al.*: **Genome-wide genetic investigation of serological measures of common infections.** *Eur. J. Hum. Genet.* 2015; **23**(11): 1544–1548.
PubMed Abstract | Publisher Full Text | Free Full Text | Reference Source

36. Wright CB, Gardener H, Dong C, *et al.*: **Infectious burden and cognitive performance: The northern Manhattan study.** *Stroke.* 2014; **45**(SUPPL. 1). Lippincott Williams and Wilkins.
Reference Source

37. Gale SD, Erickson LD, Berrett A, *et al.*: **Infectious disease burden and cognitive function in young to middle-aged adults.** *Brain Behav. Immun.* 2016; **52**: 161–168.
PubMed Abstract | Publisher Full Text | Reference Source

38. Aiello AE, Jayabalasingham B, Simanek AM, *et al.*: **The impact of pathogen burden on leukocyte telomere length in the Multi-Ethnic Study of Atherosclerosis.** *Epidemiol. Infect.* 2017; **145**(14): 3076–3084. England.
Reference Source

39. Hedges DW, Berrett AN, Erickson LD, *et al.*: **Association between infection burden and adult height.** *Econ. Hum. Biol.* 2017; **27**(Pt A): 275–280. Netherlands.
Reference Source

40. Noppert GA, Aiello AE, O'Rand AM, *et al.*: **Investigating pathogen burden in relation to a cumulative deficits index in a representative sample of US adults.** *Epidemiol. Infect.* 2018; **146**(15): 1968–1976. England.
Reference Source

41. Stebbins RC, Noppert GA, Aiello AE, *et al.*: **Persistent socioeconomic and racial and ethnic disparities in pathogen burden in the United States, 1999–2014.** *Epidemiol. Infect.* 2019; **147**: e301. England.
Reference Source

42. Noppert GA, Feinstein L, Dowd JB, *et al.*: **Pathogen burden and leukocyte telomere length in the United States.** *Immun. Ageing.* 2020; **17**.
Publisher Full Text | Reference Source

43. Pearce BD, Bracher A, Jones JL, *et al.*: **Viral and parasitic pathogen burden and the association with stroke in a population-based cohort.** *Int. J. Stroke.* 2018; **13**(5): 481–495. United Kingdom: SAGE Publications Inc.
Reference Source

44. Zhao C, Strobino K, Moon YP, *et al.*: **APOE epsilon4 modifies the relationship between infectious burden and poor cognition.** *Neurol Genet.* 2020; **6**(4): e462. United Kingdom: Lippincott Williams and Wilkins.
Reference Source

45. Beydoun MA, Beydoun HA, Hedges DW, *et al.*: **Infection burden, periodontal pathogens, and their interactive association with incident all-cause and Alzheimer's disease dementia in a large national survey.** *Alzheimers Dement.* 2024; **20**(9): 6468–6485.
PubMed Abstract | Publisher Full Text | Free Full Text

46. Rupprecht HJ, Blankenberg S, Bickel C, *et al.*: **Impact of viral and bacterial infectious burden on long-term prognosis in patients with coronary artery disease.** *Circulation.* 2001; **104**(1): 25–31. United States. Investigators A, editor.
Reference Source

47. Espinola-Klein C, Rupprecht H-J, Blankenberg S, *et al.*: **Impact of infectious burden on progression of carotid atherosclerosis.** *Stroke.* 2002; **33**(11): 2581–2586. United States.
Reference Source

48. Georges JL, Rupprecht HJ, Blankenberg S, *et al.*: **Impact of pathogen burden in patients with coronary artery disease in relation to systemic inflammation and variation in genes encoding cytokines. Group A, editor.** *Am. J. Cardiol.* 2003; **92**(5): 515–521. United States.
Reference Source

49. Rothenbacher D, Brenner H, Hoffmeister A, *et al.*: **Relationship between infectious burden, systemic inflammatory response, and risk of stable coronary artery disease: role of confounding and reference group.** *Atherosclerosis.* 2003; **170**(2): 339–345. Ireland.
Reference Source

50. Espinola-Klein C, Rupprecht HJ, Blankenberg S, *et al.*: **Impact of infectious burden on extent and long-term prognosis of atherosclerosis.** *Circulation.* 2002; **105**(1): 15–21. United States: Lippincott Williams and Wilkins (530 Walnut Street, Philadelphia,

United States).
Reference Source

51. Palm F, Pussinen PJ, Aigner A, *et al.*: **Association between infectious burden, socioeconomic status, and ischemic stroke.** *Atherosclerosis.* 2016; **254**: 117–123. Ireland.
Reference Source

52. Niu YH, Xu CF, Shi JH, *et al.*: **Relationship between infection burden and atherosclerosis and plaque feature.** *Zhonghua xin xue guan bing za zhi [Chinese journal of cardiovascular diseases].* 2005; **33**(4): 303–306. China.
Reference Source

53. Liu L, Liu Y, Tong W, *et al.*: **Pathogen burden in essential hypertension.** *Circ. J.* 2007; **71**(11): 1761–1764. Japan.
Reference Source

54. Bu X-L, Yao X-Q, Jiao S-S, *et al.*: **A study on the association between infectious burden and Alzheimer's disease.** *Eur. J. Neurol.* 2015; **22**(12): 1519–1525. England.
Reference Source

55. Bu X-L, Wang X, Xiang Y, *et al.*: **The association between infectious burden and Parkinson's disease: Acase-control study.** *Parkinsonism Relat. Disord.* 2015; **21**(8): 877–881.
Publisher Full Text | Reference Source

56. Fan F, Yang C, Zhu X, *et al.*: **Association between infectious burden and cerebral microbleeds: a pilot cross-sectional study.** *Ann. Clin. Transl. Neurol.* 2021; **8**(2): 395–405.
Reference Source

57. Zeng C, Ge P, Yin Z, *et al.*: **RNF213 Variant and Infectious Burden Associated With Intracranial Artery Stenosis in Moyamoya Disease.** *J. Am. Heart Assoc.* 2025; **14**(5): e036830.
PubMed Abstract | Publisher Full Text | Free Full Text | Reference Source

58. Horne BD, Muhlestein JB, Strobel GG, *et al.*: **Greater pathogen burden but not elevated C-reactive protein increases the risk of clinical restenosis after percutaneous coronary intervention. Group IHC (IHC) S, editor.** *Am. Heart J.* 2002; **144**(3): 491–500. United States.
Reference Source

59. Prasad A, Zhu J, Halcox JPJ, *et al.*: **Predisposition to atherosclerosis by infections: role of endothelial dysfunction.** *Circulation.* 2002; **106**(2): 184–190. United States.
Reference Source

60. Lanza GA, Sestito A, Cammarota G, *et al.*: **Assessment of systemic inflammation and infective pathogen burden in patients with cardiac syndrome X.** *Am. J. Cardiol.* 2004; **94**(1): 40–44. United States.
Reference Source

61. Tsirpanlis G, Chatzipanagiotou S, Ioannidis A, *et al.*: **Serum and peripheral blood mononuclear cells infectious burden: Correlation to inflammation and atherosclerosis in haemodialysis patients.** *Nephrology.* 2005; **10**(3): 256–263. Australia: Blackwell Publishing (550 Swanston Street, Carlton South, Australia).
Reference Source

62. Saleh N, Svane B, Jensen J, *et al.*: **Stent implantation, but not pathogen burden, is associated with plasma C-reactive protein and interleukin-6 levels after percutaneous coronary intervention in patients with stable angina pectoris.** *Am. Heart J.* 2005; **149**(5): 876–882. United States.
Reference Source

63. Lutsey PL, Pankow JS, Bertoni AG, *et al.*: **Serological evidence of infections and Type 2 diabetes: the MultiEthnic Study of Atherosclerosis.** *Diabet. Med.* 2009; **26**(2): 149–152. England.
Reference Source

64. Ebrahimi A, Nabipour I, Vahdat K, *et al.*: **High sensitivity C-reactive protein is associated with the metabolic syndrome independent to viral and bacterial pathogen burden.** *Diabetes Res. Clin. Pract.* 2009; **84**(3): 296–302. Ireland.
Reference Source

65. Green RE, Sudre CH, Warren-Gash C, *et al.*: **Common infections and neuroimaging markers of dementia in three UK cohort studies.** *Alzheimers Dement.* 2024; **20**(3): 2128–2142. United States.
Reference Source

66. Hartog L, van Rooijen MS, Ujcic-Voortman J, *et al.*: **Ethnic differences in infectious burden and the association with metabolic risk factors for cardiovascular disease: a cross-sectional analysis.** *BMC Public Health.* 2018; **18**(1): 276. England.
Reference Source

67. Howard BV, Best L, Comuzzie A, *et al.*: **C-Reactive protein, insulin resistance, and metabolic syndrome in a population with a high burden of subclinical infection: insights from the Genetics of Coronary Artery Disease in Alaska Natives (GOCADAN) study.** *Diabetes Care.* 2008; **31**(12): 2312–2314. United States.
Reference Source

68. Janson C, Asbjornsdottir H, Birgisdottir A, *et al.*: **The effect of infectious burden on the prevalence of atopy and respiratory allergies in Iceland, Estonia, and Sweden.** *J. Allergy Clin. Immunol.* 2007; **120**(3): 673–679. United States: Mosby Inc. (11830 Westline Industrial Drive, St. Louis, United States).

69. Jonker I, Schoevers R, Klein H, *et al.*: **The association between herpes virus infections and functional somatic symptoms in a general population of adolescents. The TRAILS study.** *PLoS One.* 2017; **12**(10): e0185608. United States.

70. Khairy AM, Awad A, El-Akel W, *et al.*: **Using data mining techniques to predict sustained virological response in Egyptian patients with chronic Hepatitis C Virus.** *J. Gastroenterol. Hepatol.* 2012; **27**(SUPPL. 5): 251–252. Blackwell Publishing.

71. Liu L, Liu Y, Tong W, *et al.*: **Pathogen burden in essential hypertension.** *Circ. J.* 2007; **71**(11): 1761–1764.
Publisher Full Text | Reference Source

72. Lu Y, Liu B-P, Tan CT, *et al.*: **Lifetime pathogen burden, inflammatory markers, and depression in community-dwelling older adults.** *Brain Behav. Immun.* 2022; **102**: 124–134. Netherlands.

73. Lu Y, Tan CTY, Gwee X, *et al.*: **Pathogen Burden, Blood Biomarkers, and Functional Aging in Community-Dwelling Older Adults.** *J. Gerontol. A Biol. Sci. Med. Sci.* 2021; **76**(10): 1864–1873. United States.

74. Marchesi S, Lupattelli G, Sensini A, *et al.*: **Racial difference in endothelial function: role of the infective burden.** *Atherosclerosis.* 2007; **191**(1): 227–234. Ireland.

75. Mayr M, Kiechl S, Tsimikas S, *et al.*: **Oxidized low-density lipoprotein autoantibodies, chronic infections, and carotid atherosclerosis in a population-based study.** *J. Am. Coll. Cardiol.* 2006; **47**(12): 2436–2443. United States.

76. Mundkur LA, Rao VS, Hebbagudi S, *et al.*: **Pathogen burden, cytomegalovirus infection and inflammatory markers in the risk of premature coronary artery disease in individuals of Indian origin.** *Exp. Clin. Cardiol.* 2012; **17**(2): 63–68. Canada.

77. Shi R, Yu S, Larbi A, *et al.*: **Specific and cumulative infection burden and mild cognitive impairment and dementia: A population-based study.** *Brain Behav. Immun.* 2024; **121**: 155–164. Netherlands.

78. Steptoe A, Shamaei-Tousi A, Gylfe A, *et al.*: **Socioeconomic status, pathogen burden and cardiovascular disease risk.** *Heart.* 2007; **93**(12): 1567–1570. England.

79. Steptoe A, Gylfe A, Shamaei-Tousi A, *et al.*: **Pathogen burden and cortisol profiles over the day.** *Epidemiol. Infect.* 2009; **137**(12): 1816–1824. England.
Reference Source

80. Strandberg TE, Pitkala KH, Linnavuori KH, *et al.*: **Impact of viral and bacterial burden on cognitive impairment in elderly persons with cardiovascular diseases.** *Stroke.* 2003; **34**(9): 2126–2131. United States.
Reference Source

81. Strandberg TE, Pitkala KH, Linnavuori K, *et al.*: **Cognitive impairment and infectious burden in the elderly.** *Arch. Gerontol. Geriatr.* 2004; **38**(SUPPL): 419–423.
Reference Source

82. Vahdat K, Pourbehi MR, Ostovar A, *et al.*: **Association of pathogen burden and hypertension: the Persian Gulf Healthy Heart Study.** *Am. J. Hypertens.* 2013; **26**(9): 1140–1147. United States.
Reference Source

83. Vahdat K, Azizi F, Nabipour I, *et al.*: **Chronic inflammation is correlated with percentage body fat independent to burden of infection.** *J. Diabetes.* 2011; **3**(SUPPL. 1): 246. Wiley-Blackwell.
Reference Source

84. van der Ven A, van Diest R, Hamulyak K, *et al.*: **Herpes viruses, cytokines, and altered hemostasis in vital exhaustion.** *Psychosom. Med.* 2003; **65**(2): 194–200. United States.
Reference Source

85. Vilkuna-Rautiainen T, Pussinen PJ, Roivainen M, *et al.*: **Serum antibody response to periodontal pathogens and herpes simplex virus in relation to classic risk factors of cardiovascular disease.** *Int. J. Epidemiol.* 2006; **35**(6): 1486–1494. England.
Reference Source

86. Wright CB, Gardener H, Dong C, *et al.*: **Infectious Burden and Cognitive Decline in the Northern Manhattan Study.** *J. Am. Geriatr. Soc.* 2015; **63**(8): 1540–1545. United States.
Reference Source

87. Basinkevich AB, Shakhnovich RM, Martynova VR, *et al.*: **Role of Chlamydia, mycoplasma and cytomegalovirus infection in the development of coronary artery disease.** *Kardiologiia.* 2003; **43**(11): 4–9. Russia (Federation).
Reference Source

88. Bloemenkamp DGM, Mali WPTM, Tanis BC, *et al.*: **Chlamydia pneumoniae, Helicobacter pylori and cytomegalovirus infections and the risk of peripheral arterial disease in young**

women. *Atherosclerosis.* 2002; **163**(1): 149–156.
PubMed Abstract | Publisher Full Text | Reference Source

89. Dai D-F, Lin J-W, Kao J-H, *et al.*: **The effects of metabolic syndrome versus infectious burden on inflammation, severity of coronary atherosclerosis, and major adverse cardiovascular events.** *J. Clin. Endocrinol. Metab.* 2007; **92**(7): 2532–2537. United States.
Reference Source

90. Hagiwara N, Toyoda K, Inoue T, *et al.*: **Lack of association between infectious burden and carotid atherosclerosis in Japanese patients.** *J. Stroke Cerebrovasc. Dis.* 2007; **16**(4): 145–152. United States.

91. Jafarzadeh A, Nemati M, Tahmasbi M, *et al.*: **The association between infection burden in Iranian patients with acute myocardial infarction and unstable angina.** *Acta Med. Indones.* 2011; **43**(2): 105–111. Indonesia.

92. Kis Z, Sas K, Gyulai Z, *et al.*: **Chronic infections and genetic factors in the development of ischemic stroke.** *New Microbiol.* 2007; **30**(3): 213–220. Italy.

93. Lobzin IV, Boïtsov SA, Filippov AE, *et al.*: **Effect of respiratory infections on the clinical course of coronary artery disease.** *Klin Med (Mosk).* 2005; **83**(11): 22–26.

94. Miller GE, Freedland KE, Duntley S, *et al.*: **Relation of depressive symptoms to C-reactive protein and pathogen burden (cytomegalovirus, herpes simplex virus, Epstein-Barr virus) in patients with earlier acute coronary syndromes.** *Am. J. Cardiol.* 2005; **95**(3): 317–321. United States.

95. Moretti E, Figura N, Campagna MS, *et al.*: **Infectious Burden and Semen Parameters.** *Urology.* 2017; **100**: 90–96. United States.

96. Ngeh J, Goodbourn C: **Chlamydia pneumoniae, Mycoplasma pneumoniae, and Legionella pneumophila in elderly patients with stroke (C-PEPS, M-PEPS, L-PEPS): a case-control study on the infectious burden of atypical respiratory pathogens in elderly patients with acute cerebrovascular disease.** *Stroke.* 2005; **36**(2): 259–265. United States.

97. Nyberg A, Skagius E, Englund E, *et al.*: **Abdominal aortic aneurysm and the impact of infectious burden.** *Eur. J. Vasc. Endovasc. Surg.* 2008; **36**(3): 292–296. England.

98. Tsirpanlis G, Chatzipanagiotou S, Ioannidis A, *et al.*: **Serum and peripheral blood mononuclear cells infectious burden: correlation to inflammation and atherosclerosis in haemodialysis patients.** *Nephrology (Carlton).* 2005; **10**(3): 256–263. Australia.

99. Alcantara-Neves NM, Veiga RV, Dattoli VC, *et al.*: **The effect of single and multiple infections on atopy and wheezing in children.** *J. Allergy Clin. Immunol.* 2012; **129**(2): 359–367. United States.

100. Karachaliou M, de Sanjose S, Waterboer T, *et al.*: **Is early life exposure to polyomaviruses and herpesviruses associated with obesity indices and metabolic traits in childhood?.** *Int. J. Obes.* 2018; **42**(9): 1590–1601.

101. Katan M, Elkind MSV: **Infectious burden and its role in cerebrovascular disease and cognitive impairment.** *Future Virol.* 2013; **8**(9): 833–836. United Kingdom: Future Medicine Ltd. (2nd Albert Place, Finchley Central, London, United Kingdom).

102. Bu X-L, Wang X, Xiang Y, *et al.*: **The association between infectious burden and Parkinson's disease: A case-control study.** *Parkinsonism Relat. Disord.* 2015; **21**(8): 877–881. England.

103. Katan M, Moon YP, Paik MC, *et al.*: **Infectious burden and cognitive function: The Northern Manhattan Study.** *Neurology.* 2013; **80**(13): 1209–1215. United States: Lippincott Williams and Wilkins (530 Walnut Street,P O Box 327, Philadelphia, United States).

104. von Hertzen LC, Laatikainen T, Makela MJ, *et al.*: **Infectious burden as a determinant of atopy-- a comparison between adults in Finnish and Russian Karelia.** *Int. Arch. Allergy Immunol.* 2006; **140**(2): 89–95. Switzerland.

105. Aiello AE, Diez-Roux A, Noone A-M, *et al.*: **Socioeconomic and psychosocial gradients in cardiovascular pathogen burden and immune response: the multi-ethnic study of atherosclerosis.** *Brain Behav. Immun.* 2009; **23**(5): 663–671. Netherlands.

106. Khairy P, Rinfret S, Tardif J-C, *et al.*: **Absence of association between infectious agents and endothelial function in healthy young men.** *Circulation.* 2003; **107**(15): 1966–1971. United States.

107. Espinola-Klein C, H.-J. R, Blankenberg S, *et al.*: **Manifestations of atherosclerosis in various vascular regions: Similarities and differences with regard to epidemiology, pathogenesis, and prognosis.** *Med. Klin.* 2002; **97**(4): 221–228. Germany: Urban und Vogel GmbH (Neumarkter Strasse 43, Munchen, Germany).

108. Simanek AM, Dowd JB, Zajacova A, *et al.*: **Unpacking the "black box" of total pathogen burden: Is number or type of pathogens most predictive of all-cause mortality in the United States?.** *Epidemiol. Infect.* 2015; **143**(12): 2624–2634.
PubMed Abstract | Publisher Full Text | Free Full Text | Reference Source

109. Shi R, Yu S, Larbi A, *et al.*: **Specific and cumulative infection burden and mild cognitive impairment and dementia: A population-based study.** *Brain Behav. Immun.* 2024; **121**: 155–164.
PubMed Abstract | Publisher Full Text | Reference Source

110. Zhao C, Strobino K, Moon YP, *et al.*: **APOE ϵ 4 modifies the relationship between infectious burden and poor cognition.** *Neurol Genet.* 2020; **6**(4).
Publisher Full Text

111. Elkind MSV, Ramakrishnan P, Moon YP, *et al.*: **Infectious burden and risk of stroke: the northern Manhattan study.** *Arch. Neurol.* 2010; **67**(1): 33–38. United States.
Reference Source

112. Hadavandsiri F, Shafaati M, Mohammad Nejad S, *et al.*: **Non-communicable disease comorbidities in HIV patients: diabetes, hypertension, heart disease, and obstructive sleep apnea as a neglected issue.** *Sci. Rep.* 2023; **13**(1): 12730.
PubMed Abstract | Publisher Full Text | Free Full Text

113. Moyo-Chilufya M, Maluleke K, Kgarosi K, *et al.*: **The burden of non-communicable diseases among people living with HIV in Sub-Saharan Africa: a systematic review and meta-analysis.** *EClinicalMedicine.* 2023; **65**: 102255.
Publisher Full Text

114. Tricoche AD, Wagner AL, Balmaseda A, *et al.*: **Symptoms, Infection Duration, and Hemagglutinin Inhibition Antibody Response in Influenza A Infections.** *J. Infect. Dis.* 2021; **223**(5): 838–842.
PubMed Abstract | Publisher Full Text

115. Kallaste A, Kisand K, Aart A, *et al.*: **Antibody levels remain high to one-year's follow-up after moderate and severe COVID-19, but not after mild cases.** *Infect Dis (Lond).* 2022; **54**(5): 345–355.
PubMed Abstract | Publisher Full Text | Free Full Text

116. Bu W, Hayes GM, Liu H, *et al.*: **Kinetics of Epstein-Barr Virus (EBV) Neutralizing and Virus-Specific Antibodies after Primary Infection with EBV.** *Clin. Vaccine Immunol.* 2016; **23**(4): 363–369.
PubMed Abstract | Publisher Full Text | Free Full Text

117. Spruance SL, Evans TG, McKeough MB, *et al.*: **Th1Th2-like immunity and resistance to herpes simplex labialis.** *Antivir. Res.* 1995; **28**(1): 39–55.
Publisher Full Text

118. Russell MW: **Immune Responses to Neisseria gonorrhoeae: Challenges and Opportunities With Respect to Pelvic Inflammatory Disease.** *J. Infect. Dis.* 2021; **224**(12 Suppl 2): S96–S102.
PubMed Abstract | Publisher Full Text | Free Full Text

119. Tosif S, Neeland MR, Sutton P, *et al.*: **Immune responses to SARS-CoV-2 in three children of parents with symptomatic COVID-19.** *Nat. Commun.* 2020; **11**(1): 5703.
PubMed Abstract | Publisher Full Text | Free Full Text

Open Peer Review

Current Peer Review Status: ✔ ✔ ✔

Version 3

Reviewer Report 23 October 2026

https://doi.org/10.21956/wellcomeopenres.29880.r160643

© **2026 Lin G.** This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

✔ **Gary Lin**
Johns Hopkins University Applied Physics Laboratory, Laurel, Maryland, USA

I have reviewed the response and the paper has sufficiently addressed my concerns.

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Computational Epidemiology, Public Health, and Systems Modeling and Simulation

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Reviewer Report 22 July 2026

https://doi.org/10.21956/wellcomeopenres.29880.r161004

© **2026 Haselbeck A.** This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

✔ **Andrea Haselbeck**
The University of Edinburgh, Edinburgh, Scotland, UK

This scoping review, conducted according to PRISMA-ScR, characterises how epidemiological studies have combined serological data on three or more pathogens into a composite measure of "pathogen burden." Searching MEDLINE, Embase and Scopus to 1 May 2025 without language restriction, the authors screened 3,561 records and included 74 studies published between 2000 and 2025, spanning 19 countries and 50 pathogens.
This is a useful and well-motivated review addressing a real and under-recognised methodological

problem.
My only reservation would be that this review seems to focus on non-communicable disease, as mentioned briefly in the search strategy "links between infections and NCDs. This should be acknowledge in title or in the background for clarity. A stratified presentation by pathogen (viral vs bacterial vs others) is done, but then not further followed through. I would be interested if definitions of pathogen burden are more aligned within one pathogen grouping as compared to just overall. Has it been considered to distinguish between pathogens for which seropositivity is a lifetime marker versus for which it vanishes? There is no mention of this in methods or the paragraph on use of antibody levels in the results.

Are the rationale for, and objectives of, the Systematic Review clearly stated?
Yes

Are sufficient details of the methods and analysis provided to allow replication by others?
Partly

Is the statistical analysis and its interpretation appropriate?
Yes

Are the conclusions drawn adequately supported by the results presented in the review?
Yes

If this is a Living Systematic Review, is the 'living' method appropriate and is the search schedule clearly defined and justified? ('Living Systematic Review' or a variation of this term should be included in the title.)
Not applicable

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Epipdemiological studies, disease surveillance

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Author Response 13 Aug 2026

Anne Suffel

Comment 1: This scoping review, conducted according to PRISMA-ScR, characterises how epidemiological studies have combined serological data on three or more pathogens into a composite measure of "pathogen burden." Searching MEDLINE, Embase and Scopus to 1 May 2025 without language restriction, the authors screened 3,561 records and included 74 studies published between 2000 and 2025, spanning 19 countries and 50 pathogens.
This is a useful and well-motivated review addressing a real and under-recognised methodological problem.
My only reservation would be that this review seems to focus on non-communicable disease, as mentioned briefly in the search strategy "links between infections and NCDs.

This should be acknowledge in title or in the background for clarity.
Author's response: We added a sentence about the focus on NCDs in Version 3 of the manuscript which states the following: "An optimal definition of pathogen burden could improve studies investigating the links between infection history and non-communicable disease outcomes." We also stated this more explicitly in the aims of the study: "We aimed to summarise how definitions of the burden of serologically defined infections have been conceptualised and operationalised in seroepidemiological studies to inform future research using measures of pathogen burden in relation to non-communicable diseases."

Comment 2: A stratified presentation by pathogen (viral vs bacterial vs others) is done, but then not further followed through. I would be interested if definitions of pathogen burden are more aligned within one pathogen grouping as compared to just overall.
Author's response: We included an additional paragraph differentiating between the types of pathogens included in the definition of pathogen count. "***Pathogen burden by types of pathogens*** The majority of studies included a mix of bacteria and viruses in their definition of pathogen burden independent of their approach how to define pathogen burden. Five studies using pathogen counts included viruses only (Jafarzadeh et al., 2011; Jonker et al., 2017; Karachaliou et al., 2018; Miller et al., 2014; Steptoe et al., 2007) and one study using counts included bacteria only (Ngeh & Goodbourn, 2005). Out of the studies using a mixture of viruses, bacteria and other pathogens, six used a pathogen count (Alcantara-Neves et al., 2012; Janson et al., 2007; Lobzin et al., 2005; Pearce et al., 2018; Rubicz et al., 2015; van der Ven et al., 2003), five used the titers of pathogens for their index (Beydoun et al., 2023; Green et al., 2024; Lu et al., 2021, 2022; Shi et al., 2024) and five studies used other statistical methods (Gale et al., 2016; Hedges et al., 2017; Sealy-Jefferson et al., 2013; Stebbins et al., 2019; Zeng et al., 2025)."

Comment 3: Has it been considered to distinguish between pathogens for which seropositivity is a lifetime marker versus for which it vanishes? There is no mention of this in methods or the paragraph on use of antibody levels in the results.
Author's response: The primary focus of this review was to understand different methodological approaches taken to define pathogen burden in epidemiological research. We agree that it would be interesting to distinguish between pathogens with serological lifetime markers versus those with waning antibody levels. To differentiate between pathogens with lifetime antibody markers and those without is quite challenging. For some pathogens such as Helicobacter pylori this is highly dependent on the antigen itself, some pathogens with more local infections such as HPV tend to have lower antibody titres in general. However, we acknowledge that especially for acute respiratory and gastrointestinal pathogens, quickly waning antibody responses are often observed. Hence, we decided to add this to the result sections. "Three studies including pathogens with shorter-lived antibody responses such as RSV used antibody titres to define pathogen burden (Lu et al., 2021, 2022; Shi et al., 2024). In contrast, two studies used counts of seropositivity for influenza viruses (Niu et al., 2019; Rubicz et al., 2015)."

Competing Interests: No competing interests were disclosed.

Reviewer Report 21 July 2026

https://doi.org/10.21956/wellcomeopenres.29880.r160644

© **2026 Julian T.** This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Timothy R Julian
Eawag Swiss Federal Institute of Aquatic Science and Technology, Dübendorf, Zurich, Switzerland

The authors have included discussion points in response to the previous comments, addressing the priority points for improving clarity.

One issue remains, which is the disconnect between the very specific focus on "persistent pathogens associated with non-communicable diseases" and the very general title and framing of the review as a "scoping review of Sero-epidemiological studies of pathogen burden".

Although some language in the literature search as the author notes is general, there is also a substantial focus on the subset of persistent pathogens that are explicitly searched for. This manifests, for example, in the substantial representation of CMV.

A more accurate review article would highlight the focus of the search and the results on persistent pathogens associated with noncommunicable diseases in the title, abstract, introduction, and discussion.

This minor framing for clarify and specificity is the only outstanding concern in an otherwise well written and well documented assessment of pathogen burden.

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: environmental microbiology, pathogen surveillance

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Version 1

Reviewer Report 06 June 2026

https://doi.org/10.21956/wellcomeopenres.28710.r152401

© **2026 Julian T.** This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

? **Timothy R Julian**
Eawag Swiss Federal Institute of Aquatic Science and Technology, Dübendorf, Zurich, Switzerland

Overall: The authors provide an interesting review of published literature that reports use of composite variables derived from aggregating multiple pathogen-specific serology results to inform "pathogen" or "infectious disease burden" indicator. The study is clear, with well described methodology, and initiates an important conversation on how to aggregate serology data in studies investigating multiple pathogens.

The authors should explicitly list the pathogens included in the search criteria, and motivate why they chose this subset of pathogens. The list included appears narrow, but the overall manuscript reflects a more general set of conclusions. This should be resolved by narrowing the manuscript narrative to the subset of pathogens chosen. See detailed comments.

Also, it is not clear to me that serology data from multiple distinct pathogens should be aggregated without a clear mechanistic model or motivation for the development of a composite indicator. Authors could clarify this in the discussion section: under which conditions should analyses be done on composite variables, as oppose to on individual pathogen data?

Detailed Comments:

Abstract background: Its unclear to me what pathogen burden is in the abstract, could it also be defined conceptually in the background?

Abstract results: Why 3-17? Are these the number of pathogens tested? This could be clarified here. With pathogen burden measured by combining the analyzed pathogens (from 3-17 across studies) into a single metric.

Conceptual comment: pathogen burden will be defined by the subset of pathogens investigated; within that context, no universal indicator is feasible until and unless every study follows the same set of pathogens tested.

Data extraction: It may be helpful to have a second author review a subset of the extracted data to ensure consistency and accuracy.

Introduction:
I don't understand the sentence about serolomics. Can authors expand on this concept? If viral epitopes are detected, why is this not considered seropositive?

Methods:
I am missing an explicit definition of pathogen burden or infectious burden, and I am also missing the indicator values (what is measured) for pathogen or infecctious burden. Can you provide definitions and indicators here, not only by you but also what is included in the publications assessed?

Summary of all studies: have you considered author affiliations for insight on the location of the studies, or contacting authors? This would be helpful information to obtain.
Of the forty-four adult studies and two child studies: did the rest not report or were they mixed

age studies?

Title / Content mismatch: There is a significant mismatch between the general title topic of "pathogen burden" and the vary narrow set of pathogens included in the search strategy (See Supplemental). This manifests in a very skewed interpretation of relevant pathogens, as CMV appears in 86% of all included studies. Missing are studies on, for example, respiratory or gastrointestinal diseases. Authors should consider refining the title and indeed the overall framing to focus specifically on this subset of investigated pathogens (oncogenic and ulcerative pathogens?). It is also not explicit why these specific pathogens were chosen and others were neglected. In my opinion, not limiting search criteria to a set of pathogens would result in a more comprehensive analyses, though I understand this would be substantially more work.

Discussion: "unique pathological mechanisms": also, using this approach will bias findings to the specific (often limited) set of pathogens chosen, as a consequence the findings may not be meaningfully reproducible unless others use the exact same set of pathogens and develop the composite score in the same way.

Using data on antibody titres: to what extent is this also biased by recency of infection? How does this hidden component of time since infection shift study findings?

"The choice of included pathogens" - I would argue that aggregation of serology data from multiple pathogens should be done in careful consideration of mechanisms, or otherwise should not be done at all. That is, authors should analyse data per pathogen; if multiple pathogen burden is the research question, then the specific choice of pathogens included and the development of the composite variable should be grounded in theory of the pathogen interactions.

"Our search terms covered a wide range... " see previous comment; I would argue this is a quite limited set of pathogens. Missing are common endemic diseases.

Are the rationale for, and objectives of, the Systematic Review clearly stated?

Yes

Are sufficient details of the methods and analysis provided to allow replication by others?

Yes

Is the statistical analysis and its interpretation appropriate?

Yes

Are the conclusions drawn adequately supported by the results presented in the review?

Partly

If this is a Living Systematic Review, is the 'living' method appropriate and is the search schedule clearly defined and justified? ('Living Systematic Review' or a variation of this term should be included in the title.)

Not applicable

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: infectious disease transmission, environmental engineering, environmental microbiology

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 24 Jun 2026

Anne Suffel

Comment 1: Abstract background: Its unclear to me what pathogen burden is in the abstract, could it also be defined conceptually in the background?
Author's response: We added a definition of pathogen burden in the background of the abstract. “Background Methodological advances in serological testing have made it feasible to quantify antibodies to multiple pathogens simultaneously. This has been used to study the impact of pathogen burden, i.e. the cumulative lifetime exposure to persistent pathogens, on health and mortality outcomes. However, definitions for pathogen burden using serological data vary considerably. It is unclear how to best combine serological data for multiple pathogens for exposure/outcome definitions in epidemiological studies.”

Comment 2: Abstract results: Why 3-17? Are these the number of pathogens tested? This could be clarified here. With pathogen burden measured by combining the analyzed pathogens (from 3-17 across studies) into a single metric.
Author's response: We clarified this in the abstract: “Serology data from 50 different pathogens were analysed across all studies, with pathogen burden measured by combining a range of 3–17 pathogens across studies into a single metric.”

Comment 3: Conceptual comment: pathogen burden will be defined by the subset of pathogens investigated; within that context, no universal indicator is feasible until and unless every study follows the same set of pathogens tested.
Author's response: We agree with the reviewer and added this consideration as a sentence into the discussion of the paper. “Overall, a better integration of knowledge of the infection pathogenicity, transmission route, and epidemiology, in relation to underlying cellular and genetic mechanisms of the disease (outcome) of interest, are essential for better operationalisation of pathogen burden in epidemiological research. No universal indicator of pathogen burden will be feasible as long there is no agreed set of pathogens that should be tested for in epidemiological research. As a consequence, some findings of studies using their own definition of pathogen burden may not be reproducible unless the same definition and method of defining pathogen burden is used. Future studies should include a clear justification for the measure of pathogen burden used as well as a clear description of both serological assays and statistical methods (see key recommendations).”
Comment 4: Data extraction: It may be helpful to have a second author review a subset of the extracted data to ensure consistency and accuracy.
Author's response: We had a second author reviewing a 10% subset of the extracted data.

We added this to the description of the methods as well. “The data from included articles were extracted by AS. Data from a 10% sample of the included studies were extracted by a second author (CWG) to ensure quality of the data extraction process. We extracted general study information, operationalisation of pathogen burden, pathogens included, serology methods used to assess past infection, and rationale for the choice of pathogen burden measure (**Supplementary Text S1** [22] for more detail refer to extended data). Extracted data were summarised narratively, and in tables and figures.”

Comment 5: Methods:
I am missing an explicit definition of pathogen burden or infectious burden, and I am also missing the indicator values (what is measured) for pathogen or infectious burden. Can you provide definitions and indicators here, not only by you but also what is included in the publications assessed?

Author's response: We have now added text to clarify the definition of pathogen burden in the methods and the inclusion criteria. In addition, we present the individual assays included in each pathogen burden measure by pathogen for each study, along with cut off values ad prevalence where available in supplementary material (https://doi.org/10.5281/zenodo.18669395: Supplementary_table_antibody_assays.xlsx) “We sought to describe all methods used in epidemiological research to combine measurements of at least three different pathogens into an instrument of pathogen burden (defined as a measure of cumulative exposure to at least three serologically defined pathogens)” “We included all peer-reviewed studies with an epidemiological research question in a human population that defined a composite measure of antibodies to at least three different pathogens. Previous infection with a pathogen had to be ascertained through serological data generated using any assay and at least three pathogens had to be combined into a single outcome or exposure.

Comment 6: Summary of all studies: have you considered author affiliations for insight on the location of the studies, or contacting authors? This would be helpful information to obtain.

Author's response: 6 studies had missing information on the study’s location. We have now used the author’s affiliations to supplement this information in Table 1 and also edited the manuscript accordingly. As all authors of these papers were affiliated to institutions from the same country, this approach seemed appropriate. “The data from included articles were extracted by AS. Data from a 10% sample of the included studies were extracted by a second author (CWG) to ensure quality of the data extraction process. We extracted general study information, operationalisation of pathogen burden, pathogens included, serology methods used to assess past infection, and rationale for the choice of pathogen burden measure (**Supplementary Text S1** [22] for more detail refer to extended data). If the location of the study was missing, we supplemented the location with the country of the institution the majority of authors were affiliated with. Extracted data were summarised narratively, and in tables and figures.”

Comment 7: Of the forty-four adult studies and two child studies: did the rest not report or were they mixed age studies?

Author's response: We clarified the description of the study population, now explicitly stating the study populations in hospitals were also adults. We abstained from

differentiating more by age in the adults population, as the age spans in the adult populations varied widely and we did not want to impose a definition of "senior" population. **"**Forty-four studies included cohorts of adults in a community setting (59.5%, [23, 26– 33, 35, 37 – 45, 51, 63– 86]), 28 study populations recruited adults from hospitals (37.8% [24, 34, 46– 50, 52, 54– 60, 62, 87– 98]), and two studies included children (2.7% [99, 100])."

Comment 8: Title / Content mismatch: There is a significant mismatch between the general title topic of "pathogen burden" and the vary narrow set of pathogens included in the search strategy (See Supplemental). This manifests in a very skewed interpretation of relevant pathogens, as CMV appears in 86% of all included studies. Missing are studies on, for example, respiratory or gastrointestinal diseases. Authors should consider refining the title and indeed the overall framing to focus specifically on this subset of investigated pathogens (oncogenic and ulcerative pathogens?). It is also not explicit why these specific pathogens were chosen and others were neglected. In my opinion, not limiting search criteria to a set of pathogens would result in a more comprehensive analyses, though I understand this would be substantially more work.
Author's response: This scoping review aimed to inform studies investigating the link between multiple persistent pathogens defined serologically and non-communicable diseases, with a focus on pathogens that have been tested in large population-based cohorts such as UK Biobank. Acute respiratory and gastrointestinal infections are rarely included in multiplex serology panels in these cohorts. Nevertheless, we also searched for more general terms such as 'pathogen or infection burden', which would capture studies including a wider variety of pathogens. We added some explanation for the pathogens explicitly included in the search terms. "For this scoping review, a literature search was conducted in MEDLINE, Embase, and Scopus from database inception to the 1 st of May 2025. The search strategy combined the concept of pathogen/infection burden with various kinds of infections and limited the study results to a human population. The specific infections in the search strategy focused on common persistent pathogens measured in studies investigating links between infections and non-communicable diseases (e.g., cardiovascular disease, cancer or dementia). The full search strategy can be found in the Supplement (**S1** [22]) refer to extended data." We also added this as a limitation to our discussion section: "We also focused our search strategy on specified pathogens that have been commonly studied in relationship to non-communicable diseases and measured in large population-based cohorts such as UK Biobank to inform future research using such resources"

Comment 9: "unique pathological mechanisms": also, using this approach will bias findings to the specific (often limited) set of pathogens chosen, as a consequence the findings may not be meaningfully reproducible unless others use the exact same set of pathogens and develop the composite score in the same way.
Author's response: We agree with the reviewer and added this to the limitations of our study in the discussion section: "Overall, a better integration of knowledge of the infection pathogenicity, transmission route, and epidemiology, in relation to underlying cellular and genetic mechanisms of the disease (outcome) of interest, are essential for better operationalisation of pathogen burden in epidemiological research. No universal indicator of pathogen will be feasible as long there is no agreed set of pathogens that should be

tested for in epidemiological research. As a consequence, some findings of studies using their own definition of pathogen burden may not be reproducible unless the same definition and method of defining pathogen burden is used."

Comment 10: Using data on antibody titres: to what extent is this also biased by recency of infection? How does this hidden component of time since infection shift study findings?
Author's response: We discuss these limitations in the discussion but improved the clarity of this point: **"**Using data on antibody titres may reflect the severity of the past infection, viral load, immunogenetics of the host and time since infection. However, this is highly dependent on the pathogen and antibody measured. Higher levels of haemagglutinin inhibition antibodies for influenza have been associated with febrile illness from influenza [114] and the antibody responses one year after SARS-CoV-2 infection was higher in individuals with more severe infections. [115] However, researchers found no association between the severity of mononucleosis and levels of specific EBV protein and neutralising antibodies after infection. [116] Higher antibody levels may also reflect a reactivation of a virus, for example, higher HSV-1 titres in individuals with more frequent cold sores. [117] Other pathogens such as *Neisseria gonorrhoeae* are difficult to prove due to great variability in antigens and often weak antibody response to infections. [118] It remains unclear how different levels of antibody responses are linked to the mechanisms involved in the pathogenesis of non-communicable diseases."

Comment 11: "The choice of included pathogens" - I would argue that aggregation of serology data from multiple pathogens should be done in careful consideration of mechanisms, or otherwise should not be done at all. That is, authors should analyse data per pathogen; if multiple pathogen burden is the research question, then the specific choice of pathogens included and the development of the composite variable should be grounded in theory of the pathogen interactions.
Author's response: We share the reviewer's concern around missing theories around pathogen interactions. However, we also believe that creating a composite measure of multiple pathogens can help to develop hypotheses for how pathogens interact with each other and investigate potential mechanisms. For some of the included pathogens in our study, the detailed mechanisms of interaction with the human host, other pathogens or their tropism are not fully understood. Whereas many studies have investigated relationships between single pathogens and health outcomes, we believe that examining health outcomes in relation to multiple pathogens can add knowledge about potential interactions between multiple pathogens. We added another key recommendation to reflect this: "Key recommendations for reporting pathogen burden in epidemiological studies:

- Clearly defined research question and explanation of the role of pathogen burden (e.g., outcome, exposure, covariate)
- - List of pathogens included
- - Justification for choice of pathogens included (e.g., relevance to study population)
- - Reporting of type of assays, and thresholds used to define sero-positivity
- - Any other supplementary methods used to measure pathogens (e.g., clinical data, surveys)

- - Statistical methods used to aggregate serological data
- - Justifications for the methods used (e.g., choice of assays, thresholds, statistical approach)
- Careful interpretations of findings in the context of the selected pathogens and research question
- Comment on the generalisability of findings

”

Comment 12: "Our search terms covered a wide range... " see previous comment; I would argue this is a quite limited set of pathogens. Missing are common endemic diseases.
Author's response: We edited our strengths and limitations section in the discussion to better reflect what kind of pathogens we included in our search criteria. “Our scoping review included a wide range of different studies with no exclusion criteria related to country of study or publication language. Our search terms covered a wide range of pathogens often linked to non-communicable diseases, including commonly studied viruses and general terms for infections. However, our approach might have overlooked some rare or neglected pathogens. We also focused our search strategy on specified pathogens that have been commonly studied in relationship to non-communicable diseases and measured in large population-based cohorts such as UK Biobank to inform future research using such resources”

Competing Interests: No competing interests were disclosed.

Reviewer Report 05 May 2026

https://doi.org/10.21956/wellcomeopenres.28710.r151934

© 2026 Lin G. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

?

Gary Lin
Johns Hopkins University Applied Physics Laboratory, Laurel, Maryland, USA

Summary:

The systematic scoping review analyzed serological studies and assessed their ability to quantify pathogen burden, which the authors defined as a person’s past exposure to infections, and can inform future research on pathogen burden and disease outcomes. The review included 50 pathogens, which were analyzed across 74 studies extracted from three major scientific databases. The study concluded that pathogen burden was not explicitly quantified in most studies, as only binary outcomes (infected or not) were reported. Additionally, there was no strong scientific or clinical rationale for methodological selection.

Main Comments:

Overall, the authors attempt to answer an important question about the pathogen burden on the individual and whether understanding past infections could help mitigate the risk of non-communicable diseases caused by pathogens. The study is comprehensive and rigorous in its exploration of serological studies and includes a variety of pathogens. However, I have a few comments that I would like to see addressed before indexing:

- The study selection process appears rigorous, but I would like to see more documentation of the process itself. For example, was there documentation of disagreement among reviewers regarding inclusion, or a PRISMA-ScR checklist associated with this study (See https://www.prisma-statement.org/scoping)?
- Not really clear on what the section "Definitions of pathogen burden" on page 7 is contributing. It seems there should be a table of potential or already-defined definitions.
- Was there an assessment of study methodologies to measure serology? Will test accuracy be an issue for studies trying to understand pathogen burden?
- In the "Rationale behind pathogen scores" section, what was the rationale exactly? I'm not sure I can trace where each study lists its rationale. I would like to see more critical appraisal of the study rationale. Examples would be good in this section.
- The discussion could be strengthened by more clearly outlining how the findings inform clinical or public health practice, especially changes to health surveillance that could improve understanding of pathogen burden.

Minor Comments:

- Is there a justification for including studies with >3 pathogens? Why were studies with two pathogens not included?

Are the rationale for, and objectives of, the Systematic Review clearly stated?

Yes

Are sufficient details of the methods and analysis provided to allow replication by others?

Partly

Is the statistical analysis and its interpretation appropriate?

Yes

Are the conclusions drawn adequately supported by the results presented in the review?

Yes

If this is a Living Systematic Review, is the 'living' method appropriate and is the search schedule clearly defined and justified? ('Living Systematic Review' or a variation of this term should be included in the title.)

Not applicable

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Computational Epidemiology, Public Health, and Systems Modeling and Simulation

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 18 May 2026

Anne Suffel

Comment 1: The study selection process appears rigorous, but I would like to see more documentation of the process itself. For example, was there documentation of disagreement among reviewers regarding inclusion, or a PRISMA-ScR checklist associated with this study (See https://www.prisma-statement.org/scoping)?
Authors' response: In addition to the PRISMA flow chart, we have now added a completed PRISMA-ScR checklist in the supplementary information.

Comment 2: Not really clear on what the section "Definitions of pathogen burden" on page 7 is contributing. It seems there should be a table of potential or already-defined definitions. **Authors' response**: The section "other pathogen burden definitions" summarises statistical approaches by various studies which go beyond scores based on pathogen counts or categorisations based on antibody titre.
We believe this section is necessary to better describe more complex approaches such as weighted regression indices or latent class analysis which don't fit into the other categories, i.e., counting the number of seropositive pathogens or counts based on antibody titres. Table 1 summarises all included studies ordered by different approaches taken, e.g., Studies with pathogen index as count, Studies using a count as pathogen burden and titres, other approaches for pathogen burden.

Comment 3: Was there an assessment of study methodologies to measure serology? Will test accuracy be an issue for studies trying to understand pathogen burden?
Authors' response: We did not critically assess the methodologies to measure serology as the focus of this study was to identify how pathogen burden has been defined in epidemiological research studies. Scoping reviews do not typically assess risk of bias formally as they aim to map the evidence rather than to provide a definitive answer to a research question (Munn et al. (2018)). Nevertheless, we agree that test accuracy is an important consideration for future studies and systematic reviews of studies of pathogen burden. We have added the following text to the discussion to highlight this issue: "We did not assess risk of bias in the included studies as the main interest of this review was to assess the scope of methods used to conceptualise pathogen burden rather than to assess the quality of the epidemiological research. *The quality of serological assays evolves over time; and sensitivity and specificity vary by the assays used. Test accuracy should be considered in future epidemiological studies using pathogen burden and when critically appraising the findings of such studies in systematic reviews*." Munn Z, Peters MDJ, Stern C, Tufanaru C, McArthur A, Aromataris E. Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. BMC Med Res Methodol. 2018 Nov 19;18(1):143. doi: 10.1186/s12874-018-0611-x. PMID: 30453902; PMCID: PMC6245623.

Comment 4: In the "Rationale behind pathogen scores" section, what was the rationale

exactly? I'm not sure I can trace where each study lists its rationale. I would like to see more critical appraisal of the study rationale. Examples would be good in this section.
Authors' response: We agree with the reviewer that a more detailed description and critical appraisal for the rational of defining pathogen burden would be crucial. However, none of the included studies provided a scientific rationale for the approaches taken and neither was this available in the review paper they referred to. The cited review paper states the following: "*There is at present no standard definition of infectious burden. There is no consensus as to which organisms or clinical infections should be considered, or whether optimal measures should be based on serological evidence of past exposure, or on a history of past exposures derived from questionnaires, or from independent sources of information, such as chart review or administrative databases.*" From: Elkind MS. Infectious burden: a new risk factor and treatment target for atherosclerosis. Infect Disord Drug Targets. 2010 Apr;10(2):84-90. doi: 10.2174/187152610790963519. PMID: 20166973; PMCID: PMC2891124. Hence, we concluded that referencing the Elkind review from 2010 is not sufficient justification for any of the approaches taken.

Comment 5: The discussion could be strengthened by more clearly outlining how the findings inform clinical or public health practice, especially changes to health surveillance that could improve understanding of pathogen burden.
Authors' response: We agree with the reviewer that research on the relationships between pathogen burden and other health outcomes could improve clinical practice and public health. We added the following sentence to the discussion: "Identifying specific patterns of pathogens and infection history in relation to other health outcomes can improve the prevention of both infectious diseases and associated long-term health outcomes. Enhanced surveillance of relevant pathogens, combining repeated serological assays with clinical data, may support risk assessment and guide vaccination and other preventive strategies. "

Comment 6: Is there a justification for including studies with >3 pathogens? Why were studies with two pathogens not included?
Author's response: This paper aimed to provide an overview of definitions for pathogen burden arising from the challenge of testing several pathogens at the same time with multiplex assays. We selected ≥3 pathogens pragmatically to reflect more complex definitions. This is stated in the methods section: "We chose studies with at least three pathogens to capture studies with more complex exposure/outcome definitions of multiple infections."

Competing Interests: No competing interests were disclosed.

Comments on this article

Version 3

Reader Comment 11 Sep 2026

Marianna Karachaliou, Erasmus MC, Rotterdam, The Netherlands

Thank you to the authors for this valuable scoping review of sero-epidemiological studies and approaches to assessing pathogen exposure and burden. I would like to draw attention to a study that appears not to have been identified/included in the review and that may be relevant to the methodological framework discussed: Karachaliou M, de Sanjose S, Roumeliotaki T, et al. *Heterogeneous associations of polyomaviruses and herpesviruses with allergy-related phenotypes in childhood.* Ann Allergy Asthma Immunol. 2021;127(2):191–199.e3. doi:10.1016/j.anai.2021.04.019. In this study, we used multiplex serology to assess exposure to ten polyomaviruses and three herpesviruses in children. Importantly, in addition to examining overall polyomavirus and herpesvirus burden, we constructed polyomavirus burden measures according to the viruses' predominant sites of latency/persistence. Specifically, polyomaviruses were grouped into skin, respiratory, and other sites (urinary/gastrointestinal tract). This biologically informed grouping was relevant to the findings. Overall polyomavirus burden was not associated with the allergy-related phenotypes examined, whereas associations emerged when pathogen burden was considered according to site of latency: higher skin polyomavirus burden was associated with lower odds of eczema, while respiratory polyomavirus burden was associated with higher odds of asthma. I therefore wonder whether this study met the eligibility criteria for the scoping review and, if so, whether its omission might warrant consideration in a future version/update. More broadly, I believe it illustrates an additional approach to defining "pathogen burden": rather than considering burden solely as the cumulative number of pathogens to which an individual is seropositive, pathogens can also be grouped according to biologically meaningful characteristics such as their predominant site of latency or persistence. Such grouping may reveal associations that are obscured when heterogeneous pathogens are combined into a single cumulative measure. I would be interested in the authors' perspective on whether this type of biologically informed pathogen-burden definition could be incorporated into the framework emerging from the review.

Competing Interests: No competing interests were disclosed.