



# HHS Public Access

Author manuscript

*Neuron*. Author manuscript; available in PMC 2024 April 05.

Published in final edited form as:

*Neuron*. 2023 April 05; 111(7): 1086–1093.e2. doi:10.1016/j.neuron.2022.12.029.

## Virus exposure and neurodegenerative disease risk across national biobanks

**Kristin S. Levine, M.S.<sup>1,2</sup>, Hampton L. Leonard, M.S.<sup>1,2,3,4</sup>, Cornelis Blauwendraat, Ph.D<sup>1,3</sup>, Hirotaka Iwaki, M.D.<sup>1,2,3</sup>, Nicholas Johnson, BS<sup>1,2</sup>, Sara Bandres-Ciga, Ph.D<sup>1</sup>, Luigi Ferrucci, Ph.D<sup>5</sup>, Faraz Faghri, Ph.D<sup>1,2,3</sup>, Andrew B. Singleton, Ph.D<sup>1,3</sup>, Mike A. Nalls, Ph.D<sup>1,2,3</sup>**

<sup>1</sup>Center for Alzheimer's and Related Dementias (CARD), National Institute on Aging and National Institute of Neurological Disorders and Stroke, National Institutes of Health, Bethesda, MD, USA

<sup>2</sup>Data Tecnica International LLC, Washington DC, USA

<sup>3</sup>Laboratory of Neurogenetics, National Institute on Aging, National Institutes of Health, Bethesda, MD, USA

<sup>4</sup>University of Tuebingen, Tuebingen, Germany

<sup>5</sup>Longitudinal Studies Section, National Institute on Aging, National Institutes of Health, Baltimore, MD, USA

### SUMMARY

With recent findings connecting the Epstein-Barr virus to an increased risk of multiple sclerosis and growing concerns regarding the neurological impact of the coronavirus pandemic, we examined potential links between viral exposures and neurodegenerative disease risk. Using time series data from FinnGen for discovery and cross-sectional data from the UK Biobank for replication, we identified 45 viral exposures significantly associated with increased risk of neurodegenerative disease and replicated 22 of these associations. The largest effect association was between viral encephalitis exposure and Alzheimer's disease. Influenza with pneumonia was significantly associated with five of the six neurodegenerative diseases studied. We also replicated the Epstein-Barr/multiple sclerosis association. Some of these exposures were associated with

---

**Correspondence:** Dr. Nalls can be contacted at nallsm@nih.gov or at the Center for Alzheimer's and Related Dementias, Building T44, Room 102, NIH Campus, Bethesda MD, USA 20892.

**Equal contributions:** Ms Levine and Ms Leonard and contributed equally to this paper.

#### AUTHOR CONTRIBUTIONS

K.L., H.L.L., F.F., M.N., and H.I. designed the study; K.L., H.L.L., C.B., and M.N. acquired and processed data; K.L., M.N., and H.L.L. analyzed data; all authors contributed to interpreting data and writing and editing of the manuscript.

#### DECLARATIONS OF INTEREST

K.L., H.L.L., H.I., N.J., F.F. and M.A.N.'s participation in this project was part of a competitive contract awarded to Data Tecnica International LLC by the National Institutes of Health to support open science research. M.A.N. also currently serves on the scientific advisory board for Clover Therapeutics and is an advisor to Neuron23 Inc.

#### INCLUSION AND DIVERSITY

We support inclusive, diverse, and equitable conduct of research.

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

an increased risk of neurodegeneration up to 15 years after infection. As vaccines are currently available for some of the associated viruses, vaccination may be a way to reduce some risk of neurodegenerative disease.

## eTOC blurb

We identified 45 viral exposures significantly associated with increased risk of neurodegenerative disease and replicated 22 of these associations, including the association between the Epstein-Barr virus and multiple sclerosis. As vaccines are available for some of the associated viruses, vaccination may be a way to reduce risk of neurodegenerative disease.

## Keywords

Alzheimer's disease; amyotrophic lateral sclerosis; generalized dementia; multiple sclerosis; Parkinson's disease; vascular dementia; viral exposure; influenza; encephalitis; varicella-zoster

---

## INTRODUCTION

Recent research has shown a definitive association between an increased risk of multiple sclerosis and prior infection with the Epstein-Barr virus <sup>1</sup>. Additional concerns regarding the potential short and long-term cognitive impact of the current coronavirus pandemic have raised the priority of investigating the potential connection between viral exposures and neuroinflammation and/or neurodegeneration <sup>2</sup>. The impact of bacterial infections has become an area of interest as well. While previous literature surveys have presented somewhat inconclusive results, they suggest possible associations between microbial exposures and increased risk of neurodegeneration <sup>3,4</sup>. With the current growth of biobank scale data accelerating systematic research in an open science context, we aimed to evaluate putative associations between viral exposures preceding neurodegenerative disease (NDD) onset using real-world data in a hypothesis-free manner.

In this report, we queried resources from the FinnGen project and the UK Biobank (UKB) to mine potential associations between viral exposures and a variety of common neurodegenerative diseases including Alzheimer's disease (AD), amyotrophic lateral sclerosis (ALS), generalized dementia (DEM), vascular dementia (VAS), Parkinson's disease (PD) and multiple sclerosis (MS) <sup>5,6</sup>. This report aims to survey longitudinal and cross-sectional associations between viral exposures and neurodegenerative diseases in an unbiased manner. Additionally, we aimed to shed light on the relationship in timing between pre-, post-, and peri-diagnostic viral exposures and how they relate to NDD risk.

## RESULTS

### DATA ANALYSIS IN FINNGEN AND THE UK BIOBANK TO SURVEY RISK FACTORS

Our discovery cohort came from FinnGen, a nationwide Finnish biobank with genotyping data available for over 300,000 individuals. Hazard ratios for viral endpoints were downloaded through FinnGen's [Ristey's portal](#). Replication data were applied for and downloaded from the UK Biobank, which hosts genotyping data from nearly 500,000

individuals from the UK. For controls in our replication cohort, we used a subset of 96,390 age-matched (baseline age greater than 60 years) unrelated individuals of European ancestry who did not have a neurodegenerative disease of any kind. NDD cases in the UKB were also filtered to include only unrelated individuals of European ancestry. See Table 1 for case numbers and a breakdown by sex for each of the six NDD and Supplementary Table 1 for information about the endpoints and viral groupings analyzed in this study.

## **22 REPLICATED VIRUSES ASSOCIATED WITH INCREASED RISK OF NEURODEGENERATIVE DISEASES**

We found 45 significant NDD/virus associations in FinnGen, and replicated 22 of these associations in UKB. These 22 associations have been summarized in Table 2. The highest hazard ratio, 30.72 (discovery phase HR CI 11.84 – 79.68, uncorrected p-value 1.89E-12; replication phase OR 22.06, CI 5.47 – 88.94, uncorrected p-value 1.37E-05), was seen for the association between viral encephalitis and AD. To place this in context, we see in FinnGen, 24 out of 406 viral encephalitis cases went on to develop AD (5.9%); this is higher than the general prevalence of AD in the same population at less than 3%. Dementia had the most replicated associations after multiple test correction, with six viral groupings showing significant results: viral encephalitis, other viral diseases, viral warts, all influenza, influenza and pneumonia, and viral pneumonia. No viruses were associated with a protective effect in our study; all were associated with an increased risk of neurodegenerative disease. The overwhelming majority of replicated associations include viruses commonly considered neurotrophic (81%), which means they can invade the central nervous system through peripheral nerves or by crossing the blood-brain barrier. This suggests that these viruses may increase NDD risk by lowering cognitive reserve (resilience to neurodegeneration and the ability to carry out complex mental tasks) by contributing to inflammation in the brain.

Influenza and pneumonia were significantly associated with five of the six neurodegenerative diseases (AD, ALS, dementia, PD, and vascular dementia) in FinnGen; all five were replicated using the cross-sectional data from UKB. As an additional note, these should be considered severe cases of influenza and pneumonia, as all viral codes in FinnGen and UKB are sourced from the electronic medical records (EMRs) of these participants that are collected when they receive medical intervention or hospital-based treatment. Viral encephalitis and ‘viral and other specified intestinal infections’ were significant and replicated for both AD and dementia; the varicella-zoster virus was significant and replicated for both MS and vascular dementia. (As a note, “varicella-zoster virus” is referred to as “Zoster [herpes zoster]” in the FinnGen and UKB coding system.) Please refer to Table 2 for further details about replicated associations. Supplementary Table 2 contains all tested associations found in FinnGen or UKB.

## **THE ASSOCIATION BETWEEN THE EPSTEIN-BARR VIRUS AND MS AS A POSITIVE CONTROL**

As a positive control, we replicated the previously published findings in FinnGen (HR 26.5; CI 3.7 to 191.6;  $P = 0.001$ )<sup>1</sup> showing an association between a preceding Epstein-Barr virus (EBV) exposure and increased risk of MS. Across all follow-up durations in FinnGen, EBV exposure was associated with MS risk at an HR of 3.92 (HR CI 2.57 – 6.00, uncorrected

p-value 2.82E-10 with 43 overlapping cases). This association changed as follow-up time to MS diagnosis increased; at 5–15 years before a diagnosis of MS, the pairing had an HR of 2.08 (HR CI 1.25 – 3.47, uncorrected p-value 5.03E-03 with 21 overlapping cases). The EBV and MS association was significant in FinnGen however, it did not replicate in UKB. We suspect this is because of the way UKB utilizes hospital diagnostic codes. EBV, while a very common exposure, is not usually reported as a cause of hospital admission. This suggests that minor differences in the design of the two biobanks and the standards of reporting certain illnesses may have some impact on the sparsity of these data and related results.

## **INCREASED RISK OF NEURODEGENERATION PERSISTS YEARS AFTER INITIAL VIRUS INFECTION**

For the first part of our study, we looked at FinnGen data for viral exposure across all available years preceding NDD diagnosis (referred to as HR lag 0 in the Risteys database query). When available, we also examined FinnGen data for viral exposure at < 1 year, between 1 and 5 years, and between 5 and 15 years prior to neurodegenerative disease onset (referred to as HR lag 1, 5, and 15 respectively in the Risteys database query).

Of the 22 significant and replicated pairings, lag data was available for 16 pairs. Six of those pairing remained significant 5–15 years before NDD diagnosis, with hazard ratios for associations ranging from 1.49 (dementia and influenza and pneumonia) to 4.98 (AD and meningitis – note: this grouping also included bacterial ICD10 codes for meningitis). Looking at all FinnGen virus-NDD pairings for which data was available, 17 pairs remained significant at 5–15 years between exposure and diagnosis; dementia and infectious mononucleosis (EBV) had the highest hazard ratio (9.00). No pairings were significant at lag 15 for ALS or PD.

At 1–5 years preceding NDD diagnosis, 15 replicated pairings had significant associations, with hazard ratios ranging from 2.24 (PD and influenza and pneumonia) to 24.14 (dementia and viral encephalitis).

Generally, hazard ratios tended to be highest less than 1 year before the diagnosis of a neurodegenerative disease as summarized in Figure 1. Eleven replicated pairings had significant associations at < 1 year preceding NDD diagnosis. The hazard ratio for these pairings ranged from 6.2 (PD and influenza and pneumonia) to 83.03 (dementia and viral encephalitis).

See Table 3 for a summary of all replicated associations from the discovery phase with detailed temporal follow-up data. Supplementary Table 3 contains all associations surveyed for the exposure lag time analysis before NDD diagnosis.

## **VIRAL RISK AFTER DIAGNOSIS OF NEURODEGENERATIVE DISEASE IS MODERATE AS TIME PROGRESSES**

We also looked at the risk of contracting certain viruses after being diagnosed with an NDD. Of our original 73 viral endpoint/NDD pairings, data was available both before and after NDD diagnosis for 33 pairs, including 12 of our significant and replicated pairings. There

was no data available for ALS. All associations surveyed are available in Supplementary Table 4.

At lag 1, eight viral/NDD pairings had sufficient data in both directions (before and after NDD onset). For 100% of these pairings, the hazard ratio for developing an NDD after a viral exposure was higher than the hazard ratio for being infected with a virus after the diagnosis of an NDD. For example, at lag 1, for a patient with influenza, the HR for diagnosis of Alzheimer's disease within the next year was 30.29; at the same lag, the HR of a patient with Alzheimer's contracting influenza within the next year was only 8.84. There was no overlap between the two HR confidence intervals for the influenza/AD pairing, as well as for influenza/dementia and specified intestinal infections/dementia.

At lag 5, there were ten pairings with sufficient data in both directions. For seven of the ten, the HR was higher before NDD diagnosis. The confidence intervals for all of these pairs overlapped, except for Bell's palsy/dementia. For a patient at lag 5 with Bell's palsy, the HR for dementia was 2.49; at the same lag, the HR of a patient with dementia diagnosed with Bell's palsy was only 0.47. We were unable to replicate the significant Bell's palsy/dementia association in UKB, however, we note that all of the cases of Bell's palsy in UKB were in dementia patients.

At lag 15, there were two pairings with sufficient data in both directions. For both pairs (influenza/Alzheimer's and Bell's palsy/dementia) the HR was higher before NDD diagnosis. The confidence intervals for Bell's palsy/dementia at this lag also did not overlap.

Because treatments for MS are known to increase the risk of varicella-zoster,<sup>7</sup> we looked to see if the data showed this increase. Data for the varicella-zoster virus/MS pairing were only available at lag 0 for viral exposure before NDD diagnosis, and at lag 0 and lag 15 for viral exposure after NDD diagnosis. The HR for varicella-zoster is higher after MS diagnosis, although the confidence intervals do overlap.

## DISCUSSION

While past studies have looked at individual NDD and viral infection exposures, to our knowledge this is the first systematic investigation of multiple neurodegenerative disease pairings with multiple viruses. The current study found 45 significant associations in longitudinal data from FinnGen between exposure to a viral infection and risk of later developing a neurodegenerative disease, with 22 of these associations replicated in cross-sectional data from the UKB. Influenza (with or without pneumonia), was the most commonly associated viral endpoint, significant in five of the six NDDs studied. Viral encephalitis, intestinal infections, and varicella-zoster virus were also significant and replicated for more than one NDD.

The results described above are supported by recent findings in the literature, which suggest an association between herpes simplex virus (HSV) encephalitis and AD,<sup>8,9</sup> AD and hepatitis,<sup>10,11</sup> genital warts and dementia,<sup>12</sup> EBV and dementia,<sup>13</sup> and MS and HSV.<sup>14</sup> Since the discovery of an association of the 1918 flu pandemic, caused by H1N1 influenza A, with postencephalitic parkinsonism, a link between influenza and PD has been debated.<sup>15</sup>

A recent study using Danish data found an association between influenza and PD with an odds ratio of 1.73 up to 10 years after virus exposure.<sup>16</sup> This is very similar to the hazard ratio for influenza and pneumonia in FinnGen which was 1.72.

While the strongest effect risk pairings are rare, such as AD and viral encephalitis, many relatively common pairings (such as dementia and influenza) are associated with moderately increased risk.

Strikingly, vaccines are currently available for some of these viruses, including influenza, shingles (varicella-zoster), and pneumonia. While vaccines do not prevent all cases of illness, they are known to dramatically reduce hospitalization rates. This evidence suggests that vaccination may mitigate some risk of developing neurodegenerative disease.

Recent research supports this idea. Influenza and pneumonia vaccination has been found to reduce risk for AD and PD.<sup>17 18</sup> Shingles (varicella-zoster) vaccination is associated with a reduced risk of dementia, AD, and PD in both the United States and Wales.<sup>19–21 22</sup> Despite these findings, influenza vaccination coverage in the United States is typically less than 50%.<sup>23</sup> Only about 35% of people over the age of 60 have received a shingles vaccine.<sup>24</sup> Increasing the use of widely available vaccines may give clinicians a way to help their patients reduce their overall risk of neurodegenerative disease later in life.

More research is needed into the role vaccinations might play in the prevention of neurodegenerative disease. Efficacious viral vaccines would limit virus spread, reduce viral load at infection initiation, and prevent aberrant immune reactivity. This may in turn play a role in downstream neurodegenerative disease pathogenesis, as growing and robust evidence points towards the immune system response and neuroinflammation as major disease contributors. Since the viral warts endpoint also includes papillomaviruses, we wonder if increased use of the human papillomavirus (HPV) vaccine in younger populations might lead to a decrease in certain dementias in those groups as they age. The recent advent of messenger RNA vaccine technology holds promise in revolutionizing the prevention of viral infections.

Unlike expensive and invasive tools, such as MRI, PET scans, CSF, and genetic testing, records of common viral illnesses should be an easily accessible part of an individual's medical record. Paying close attention to these potential risk factors may be a quick and easy way to identify those at greater risk of developing a NDD. This information would be valuable to individuals planning for future care and their doctors, as well as researchers who are enrolling patients in clinical trials within a context of precision medicine and predictive models.

These findings also suggest additional avenues to explore for both the treatment and prevention of neurodegenerative diseases. In addition to vaccination, some studies suggest that using antivirals may reduce the risk of dementia in HSV positive patients<sup>25</sup> or in patients with varicella-zoster virus.<sup>26</sup> Antiviral target-specific drugs could also potentially provide ways of ameliorating risk or potentially halting neurodegenerative disease progression, but this remains a goal for future drug development and repurposing efforts. While designing therapies able to completely clear viruses from the body is a daunting

challenge, identifying candidates that can stop virus replication rapidly could be an effective solution for further risk factor reduction.

Clinical trials for neurodegenerative diseases generally require a long follow-up period and are often unsuccessful; one reason may be due to the heterogeneity of disease progression. One potential strategy is to enrich the study cohort with participants that have faster disease progression.<sup>27</sup> In this study, we did not have longitudinal cognitive test scores or motor evaluation results. However, it would be an interesting next step to look deeper into disease-targeted cohorts with more detailed longitudinal evaluations to test if viral infections are associated with differences in disease progression or severity.

This report is not without its limitations. Summary data level access to FinnGen is excellent, although participant-level access could have facilitated additional modeling efforts. Our access to only cross-sectional and prevalent disease data in UKB at this time was also a limitation that we hope to improve in further research. The granularity of viral exposure quantification is of concern in some cases, as there may be different criteria for common or less severe viral exposures in Finland versus the UK. This may introduce some noise into the study data and lead to incompatibility, particularly when comparing hazard estimates to odds ratios. Additionally, both FinnGen and UKB use diagnoses based on medical billing codes and not bloodwork or other assays.

While we are working with neurodegenerative diseases that generally occur in middle-aged to elderly study participants, it is unclear if viral exposures more than 15 years before NDD diagnosis contribute to risk. It would be beneficial to examine early-life exposures in greater detail, however, standardization of electronic records has only occurred in the past two decades, making longer-term retrospective data difficult to access. We were unable to investigate the relationship between viral exposure and age at onset of an NDD as we were limited by what was available in our discovery set, but this would be an important and interesting follow-up to this study.

We also cannot rule out potential reverse causality due to potential innate immune dysregulation in NDDs<sup>28</sup>. A significant body of research supports the notion that the neurodegenerative process begins many (10–20) years before diagnosis. It is possible, for example, that a hospitalization for influenza with pneumonia that is recorded as 5 years before an NDD diagnosis could actually be occurring 5 years *after* the degenerative process has already begun. However, if patients are infected with more viruses as their bodies suffer the effects of an NDD before diagnosis, we might expect to see even more cases of a virus after diagnosis.

The lag data counters this idea for some viruses. For example at lag 1, if the prior condition is influenza, the HR for a diagnosis of Alzheimer's disease within the next year is 30.29. However, if the prior condition is Alzheimer's, the HR for an influenza diagnosis is lower, 8.84, with no overlap of the confidence intervals of the two hazard ratios. On the other hand, certain treatments for MS are known to increase the risk of varicella-zoster<sup>7</sup> and we also see this in our lag results. If varicella-zoster virus is the prior condition, the HR for MS is

2.12; however if MS is the prior condition, the HR for varicella-zoster virus is higher: 3.44. Replication in an additional time-sensitive dataset may help clarify this issue.

Additionally, the Eurocentric nature of the resources could be expanded for more globally generalizable results. In future efforts, we plan to address these concerns by incorporating more diverse time series-focused datasets. In addition, we plan to evaluate potential therapeutic and genomic mechanisms connecting these viral exposures with neurodegeneration.

Finally, in light of the current coronavirus pandemic, our results illustrate the need to take seriously the reports of concomitant neurological symptoms accompanying viral exposures and monitor at risk patients to discover if they will be at higher risk of NDDs in the future.<sup>29 30 31 32</sup> The findings reported here cover multiple neurodegenerative diseases from two different biobanks, and previous research supports the described associations between viral infection and NDD risk, suggesting that virus/NDD associations are worthy of further investigation.

## STAR METHODS

### RESOURCE AVAILABILITY

**Lead contact**—Further information and requests should be directed to and will be fulfilled by the lead contact, Mike Nalls (nallsm@nih.gov).

**Materials availability**—Discovery phase data is available in the link here [<https://risteys.finnngen.fi/>] and was accessed in May 2022. The UKB data is available here [<https://www.ukbiobank.ac.uk/>] and last accessed in May 2022 containing data from the 2018 general release. Both the discovery and replication phase dataset are covered by the relevant national IRB groups.

**Data and code availability**—Notebooks containing code used in this analysis can be found in the GitHub link here [[https://github.com/NIH-CARD/NDD\\_virus](https://github.com/NIH-CARD/NDD_virus)]. This paper analyzes existing, publicly available data. In addition, complete summary statistics describing these data/processed datasets derived from these data have been deposited in the supplementary materials connected to this publication and are publicly available as of the date of publication.

### STUDY DESIGN

Our initial study design included a discovery and replication phase, using time-aware data from the FinnGen longitudinal cohort as discovery and cross-sectional data from the UKB as replication.

We downloaded time-aware (stemming from longitudinal cohort data) risk estimates from the FinnGen Risteys portal for six neurodegenerative diseases: Alzheimer's disease, amyotrophic lateral sclerosis, generalized dementia, multiple sclerosis, Parkinson's disease, and vascular dementia. FinnGen endpoints are collections of one or more ICD10 codes, as defined by FinnGen clinical expert groups. For example, the FinnGen endpoint



AB1\_VIRAL\_ENCEPHALITIS is described as “viral encephalitis, not elsewhere classified/unspecified” and contains the ICD-10 codes A85 and A86 (A85.0 Enteroviral encephalitis; A85.1 Adenoviral encephalitis; A85.2 Arthropod-borne viral encephalitis, unspecified; A85.8 Other specified viral encephalitis, Encephalitis lethargica, Von Economo-Cruchet disease; A86 Unspecified viral encephalitis). Diagnoses were based on hospital billing codes and not specific viral assays.

In the neurodegenerative disease data, there were 1385 endpoints available. We searched those endpoints for viral keywords: ‘viral’, ‘virus’, ‘mononucleosis’, ‘epstein’, ‘ebv’, ‘encephalitis’, ‘hepatitis’, ‘meningitis’, ‘warts’, ‘influenza’, ‘bell’s\_palsy’, ‘s\_palsy’, ‘chicken’, ‘shingles’, ‘zoster’, ‘measles’, ‘varicella’, ‘herpes’, ‘myocarditis’, ‘erythema multiforme’, ‘subacute thyroiditis’, and ‘cold’. This left us with 32 viral-related endpoints with available NDD data. In FinnGen all concurrent endpoints include at least 10 cases. In total, we surveyed 73 viral endpoint and neurodegenerative disease pairs in FinnGen.

We attempted to survey all potential associations regardless of prior information on a viral risk factor and NDD pairing. As an additional note, viral codes and descriptions are based on shared criteria between the UKB and FinnGen; please see the supplemental tables for further details on sample counts. As per FinnGen’s portal methods, individuals with multiple viral exposures were not excluded from association tests.

Using the UKB data, we recreated the 32 FinnGen viral endpoints. For example, for the AB1\_VIRAL\_ENCEPHALITIS endpoint discussed above, there was UKB data available for A85.0 (Enteroviral encephalitis), A85.8 (Other specified viral encephalitis), and A86 (Unspecified viral encephalitis), so those codes were used to recreate the endpoint. See Supplementary Table 1 for the exact code groupings for all endpoints. We included any virus-neurodegenerative disease pairs with at least 3 concurrent cases (i.e. individuals with positive diagnoses of both virus and NDD), for a total of 97 pairings in UKB.

For the discovery phase of our analysis, we focused on extracting all precomputed associations centrally generated at FinnGen, that tested viral exposures (i.e. positive diagnosis of a virus) preceding the onset of each of our common NDDs of interest. This is referred to as lag 0 in the Risteys database. While summary statistics from FinnGen (version 9) were accessed and extracted in May of 2022, we concurrently generated compatible cross-sectional results in the UKB to serve as the replication phase, since time-series data in UKB was not available for all viral exposures at that time.

To dig deeper into the span of time between viral exposure and neurodegenerative disease onset, we also accessed FinnGen data for viral exposure at < 1 year, between 1–5 years, and between 5–15 years (referred to as lag 1, lag 5, and lag 15 in the Risteys database) prior to neurodegenerative disease onset. We extracted the same data for viral exposures after NDD diagnosis for any pairs that had this level of data available. See Supplementary Table 4. This provided us with estimates on roughly how long viral exposures may be relevant to NDD risk, while also letting us estimate potential bias of viral exposures that are peri-diagnostic or post-diagnostic with regard to the onset of a neurodegenerative disease.

## STATISTICAL ANALYSES

Summary statistics were downloaded from the FinnGen Ristey's portal for any association identified in our initial study design that met our inclusion criteria for the discovery and replication phases of analysis. All summary statistics were compiled and stratified by neurodegenerative disease outcome. Within each stratum of neurodegenerative disease, we used a Benjamini-Hochberg false discovery rate correction (FDR) with a corrected p-value less than 0.05 to indicate a significant association between that neurodegenerative disease and a particular preceding viral exposure<sup>33</sup>.

For the replication phase of our analysis, all pairs of neurodegenerative disease and virus exposures passing multiple test correction in the discovery efforts in FinnGen were evaluated in compatible cross-sectional data in the UKB. The UKB data lacked the longitudinal components available in FinnGen. To ensure compatibility with FinnGen, only unrelated samples with genetically-confirmed European ancestry were analyzed. While FinnGen employed cox models to estimate longitudinal risk associations, adjusted for age, sex, and socioeconomic status, we used logistic regression to predict the effect of viral exposure on each NDD in the UKB. For our cross-sectional analysis in UKB, we mirrored the covariates in these models to ensure further compatibility with FinnGen and included age at enrollment, genetically determined sex, and Townsend deprivation index.<sup>34</sup> As part of the replication analysis we utilized a stringent FDR cut off for replication, mirroring our discovery analysis. Associations with p-values passing FDR in both the discovery (FinnGen) and replication phases (UKB) were considered replicated and potentially robust.

To attempt to account for additional lifestyle confounders that may affect our results, we also calculated polygenic risk scores (PRS) for diet and physical activity for our UKB participants and reran our analysis, using the two PRS as additional covariates. Their addition did not significantly change our results (with resulting P values consistent across covariate sets at more than three significant digits). PRS were calculated as per Nalls et al 2019<sup>35</sup>, using genome-wide significant loci from the largest available European ancestry genome-wide association studies for physical activity<sup>36</sup> and dietary protein imbalance.<sup>37</sup> See Supplementary Table 5 for results.

Because hospitalization due to viral infection may lead to a diagnosis of an NDD that would otherwise have likely remained undiagnosed for a period of time, and because this would artificially inflate the association of viral infection with NDD risk, we sought to explore whether viral infection could significantly predate NDD diagnosis. To further evaluate the timing of viral exposure to neurodegenerative disease onset, we extracted FinnGen summary statistics for all nominated pairs of outcomes surveyed in the discovery phase that had data available at lag 1, 5, and 15 denoting a maximum time between viral exposure and neurodegenerative disease onset. These were treated to the same disease-stratified FDR correction as described above. All analyses were conducted with Python except the PRS analysis, which was conducted with Plink1.9<sup>38</sup>.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## ACKNOWLEDGEMENTS

### FUNDING

This research was supported in part by the Intramural Research Program of the NIH, National Institute on Aging (NIA), National Institutes of Health, Department of Health and Human Services; project number Z01 AG000535, as well as the National Institute of Neurological Disorders and Stroke. A special thanks to Walter Koroshetz, M.D. for his feedback on an early draft of this paper.

This research has been conducted using the UK Biobank Resource under Application Number 33601.

We want to acknowledge the participants and investigators of the FinnGen study and thank them for their hard work and generosity.

## REFERENCES

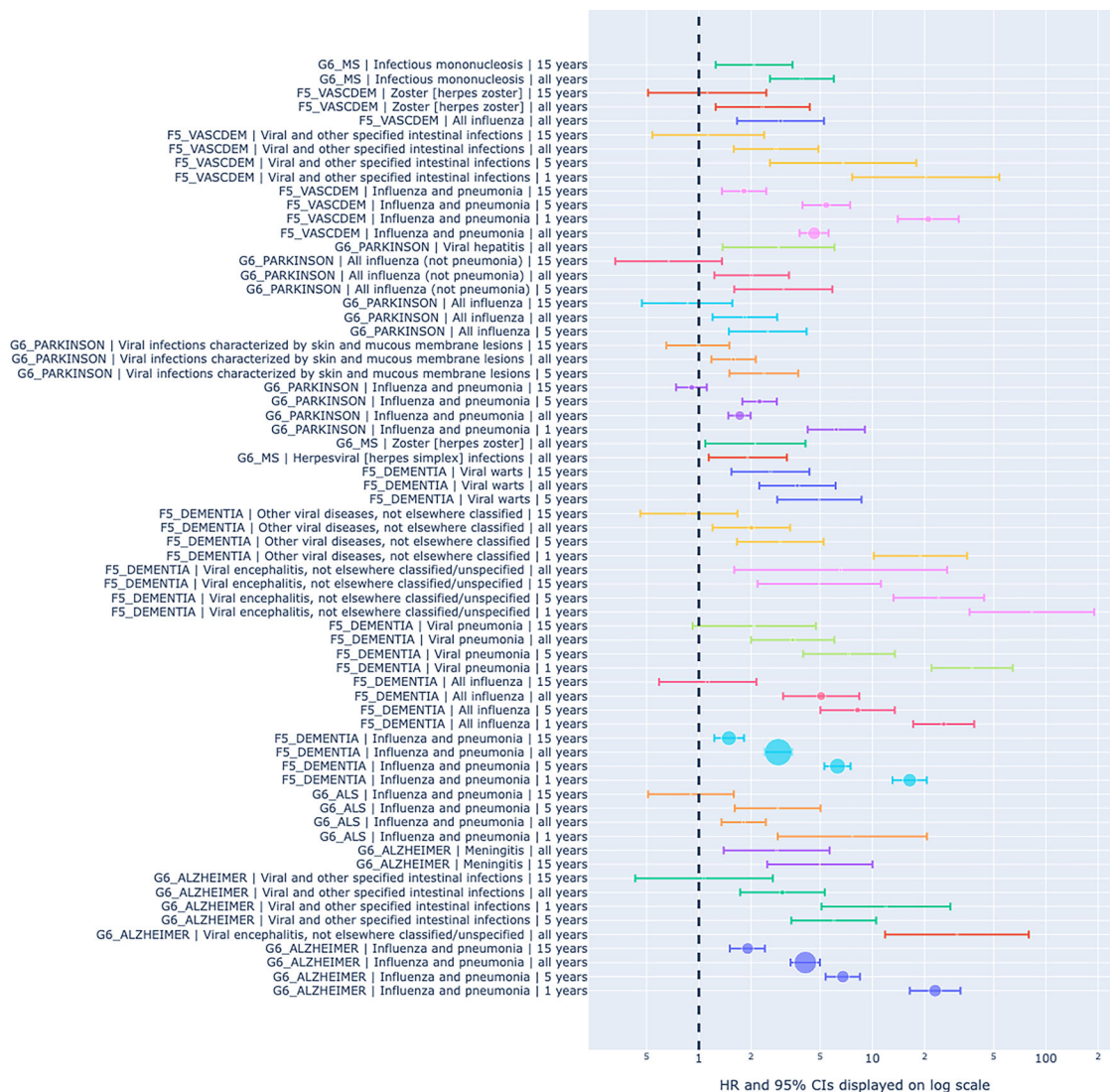
1. Bjornevik K, Cortese M, Healy BC, Kuhle J, Mina MJ, Leng Y, Elledge SJ, Niebuhr DW, Scher AI, Munger KL, et al. (2022). Longitudinal analysis reveals high prevalence of Epstein-Barr virus associated with multiple sclerosis. *Science* 375, 296–301. [PubMed: 35025605]
2. Spudich S, and Nath A (2022). Nervous system consequences of COVID-19. *Science* 375, 267–269. [PubMed: 35050660]
3. Lotz SK, Blackhurst BM, Reagin KL, and Funk KE (2021). Microbial Infections Are a Risk Factor for Neurodegenerative Diseases. *Front. Cell. Neurosci.* 15, 691136. [PubMed: 34305533]
4. Ranjan R, Abhinay A, and Mishra M (2018). Can oral microbial infections be a risk factor for neurodegeneration? A review of the literature. *Neurol. India* 66, 344–351. [PubMed: 29547153]
5. Kurki MI, Karjalainen J, Palta P, Sipilä TP, Kristiansson K, Donner K, Reeve MP, Laivuori H, Aavikko M, Kaunisto MA, et al. (2022). FinnGen: Unique genetic insights from combining isolated population and national health register data. *medRxiv*, 2022.03.03.22271360.
6. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, Downey P, Elliott P, Green J, Landray M, et al. (2015). UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med.* 12, e1001779. [PubMed: 25826379]
7. Arvin AM, Wolinsky JS, Kappos L, Morris MI, Reder AT, Tornatore C, Gershon A, Gershon M, Levin MJ, Bezuidenhout M, et al. (2015). Varicella-zoster virus infections in patients treated with fingolimod: risk assessment and consensus recommendations for management. *JAMA Neurol.* 72, 31–39. [PubMed: 25419615]
8. Itzhaki RF, and Tabet N (2017). Herpes simplex encephalitis and Alzheimer's disease: Is there a link? *J. Neurol. Sci.* 380, 20–21. [PubMed: 28870568]
9. Marcocci ME, Napoletani G, Protto V, Kolesova O, Piacentini R, Li Puma DD, Lomonte P, Grassi C, Palamara AT, and De Chiara G (2020). Herpes Simplex Virus-1 in the Brain: The Dark Side of a Sneaky Infection. *Trends Microbiol.* 28, 808–820. [PubMed: 32386801]
10. Bassendine MF, Taylor-Robinson SD, Fertleman M, Khan M, and Neely D (2020). Is Alzheimer's Disease a Liver Disease of the Brain? *J. Alzheimers. Dis.* 75, 1–14. [PubMed: 32250293]
11. Lin H-C, Xirasagar S, Lee H-C, Huang C-C, and Chen C-H (2017). Association of Alzheimer's disease with hepatitis C among patients with bipolar disorder. *PLoS One* 12, e0179312. [PubMed: 28622343]
12. Lin C-H, Chien W-C, Chung C-H, Chiang C-P, Wang W-M, Chang H-A, Kao Y-C, and Tzeng N-S (2020). Increased risk of dementia in patients with genital warts: A nationwide cohort study in Taiwan. *J. Dermatol.* 47, 503–511. [PubMed: 32189395]
13. Huang S-Y, Yang Y-X, Kuo K, Li H-Q, Shen X-N, Chen S-D, Cui M, Tan L, Dong Q, and Yu J-T (2021). Herpesvirus infections and Alzheimer's disease: a Mendelian randomization study. *Alzheimers. Res. Ther.* 13, 158. [PubMed: 34560893]
14. Xu L, Zhang L-J, Yang L, Yang C-S, Yi M, Zhang S-N, Wang N, Huang C-N, and Liu M-Q (2021). Positive association of herpes simplex virus-IgG with multiple sclerosis: A systematic review and meta-analysis. *Mult. Scler. Relat. Disord.* 47, 102633. [PubMed: 33254090]
15. Hoffman LA, and Vilensky JA (2017). Encephalitis lethargica: 100 years after the epidemic. *Brain* 140, 2246–2251. [PubMed: 28899018]

16. Cocoros NM, Svensson E, Szépligeti SK, Vestergaard SV, Szentkúti P, Thomsen RW, Borghammer P, Sørensen HT, and Henderson VW (2021). Long-term Risk of Parkinson Disease Following Influenza and Other Infections. *JAMA Neurol.* 78, 1461–1470. [PubMed: 34694344]
17. Bukhbinder AS, Ling Y, Hasan O, Jiang X, Kim Y, Phelps KN, Schmandt RE, Amran A, Coburn R, Ramesh S, et al. (2022). Risk of Alzheimer's Disease Following Influenza Vaccination: A Claims-Based Cohort Study Using Propensity Score Matching. *J. Alzheimers. Dis.* 88, 1061–1074. [PubMed: 35723106]
18. Lehrer S, and Rheinstein PH (2022). Vaccination Reduces Risk of Alzheimer's Disease, Parkinson's Disease and Other Neurodegenerative Disorders. *Discov. Med.* 34, 97–101. [PubMed: 36281030]
19. Schnier C, Janbek J, Lathe R, and Haas J (2022). Reduced dementia incidence after varicella zoster vaccination in Wales 2013–2020. *Alzheimers. Dement.* 8, e12293.
20. Scherrer JF, Salas J, Wiemken TL, Hoft DF, Jacobs C, and Morley JE (2021). Impact of herpes zoster vaccination on incident dementia: A retrospective study in two patient cohorts. *PLoS One* 16, e0257405. [PubMed: 34788293]
21. Lehrer S, and Rheinstein PH (2021). Herpes Zoster Vaccination Reduces Risk of Dementia. *In Vivo* 35, 3271–3275. [PubMed: 34697158]
22. Lophatananon A, Mekli K, Cant R, Burns A, Dobson C, Itzhaki R, and Muir K (2021). Shingles, Zostavax vaccination and risk of developing dementia: a nested case–control study—results from the UK Biobank cohort. *BMJ Open* 11, e045871.
23. Lu P-J, Hung M-C, Srivastav A, Grohskopf LA, Kobayashi M, Harris AM, Dooling KL, Markowitz LE, Rodriguez-Lainz A, and Williams WW (2021). Surveillance of Vaccination Coverage Among Adult Populations -United States, 2018. *MMWR Surveill. Summ.* 70, 1–26.
24. Terlizzi EP, and Black LI (2020). Shingles Vaccination Among Adults Aged 60 and Over: United States, 2018. *NCHS Data Brief*, 1–8.
25. Tzeng N-S, Chung C-H, Lin F-H, Chiang C-P, Yeh C-B, Huang S-Y, Lu R-B, Chang H-A, Kao Y-C, Yeh H-W, et al. (2018). Anti-herpetic Medications and Reduced Risk of Dementia in Patients with Herpes Simplex Virus Infections—a Nationwide, Population-Based Cohort Study in Taiwan. *Neurotherapeutics* 15, 417–429. [PubMed: 29488144]
26. Chen VC-H, Wu S-I, Huang K-Y, Yang Y-H, Kuo T-Y, Liang H-Y, Huang K-L, and Gossop M (2018). Herpes Zoster and Dementia: A Nationwide Population-Based Cohort Study. *J. Clin. Psychiatry* 79. 10.4088/JCP.16m11312.
27. Leonard H, Blauwendraat C, Krohn L, Faghri F, Iwaki H, Ferguson G, Day-Williams AG, Stone DJ, Singleton AB, Nalls MA, et al. (2020). Genetic variability and potential effects on clinical trial outcomes: perspectives in Parkinson's disease. *J. Med. Genet.* 57. 10.1136/jmedgenet-2019-106283.
28. Lutshumba J, Nikolajczyk BS, and Bachstetter AD (2021). Dysregulation of Systemic Immunity in Aging and Dementia. *Front. Cell. Neurosci.* 15. 10.3389/fncel.2021.652111.
29. Piekut T, Hurla M, Banaszek N, Szejn P, Dorszewska J, Kozubski W, and Prendecki M (2022). Infectious agents and Alzheimer's disease. *J. Integr. Neurosci.* 21, 73. [PubMed: 35364661]
30. Eldeeb MA, Hussain FS, and Siddiqi ZA (2020). COVID-19 infection may increase the risk of parkinsonism - Remember the Spanish flu? *Cytokine Growth Factor Rev.* 54, 6–7. [PubMed: 32563553]
31. Badrfam R, and Zandifar A (2020). From encephalitis lethargica to COVID-19: Is there another epidemic ahead? *Clin. Neurol. Neurosurg.* 196, 106065. [PubMed: 32659638]
32. Satheesh NJ, Salloum-Asfar S, and Abdulla SA (2021). The Potential Role of COVID-19 in the Pathogenesis of Multiple Sclerosis-A Preliminary Report. *Viruses* 13. 10.3390/v13102091.
33. Benjamini Y, and Hochberg Y (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)* 57, 289–300. 10.1111/j.2517-6161.1995.tb02031.x.
34. Townsend P, Phillimore P, and Beattie A (1988). *Health and Deprivation: Inequality and the North* (Routledge).
35. Nalls MA, Blauwendraat C, Vallerga CL, Heilbron K, Bandres-Ciga S, Chang D, Tan M, Kia DA, Noyce AJ, Xue A, et al. (2019). Identification of novel risk loci, causal insights, and heritable risk

- for Parkinson's disease: a meta-analysis of genome-wide association studies. *Lancet Neurol.* 18, 1091–1102. [PubMed: 31701892]
36. Doherty A, Smith-Byrne K, Ferreira T, Holmes MV, Holmes C, Pulit SL, and Lindgren CM (2018). GWAS identifies 14 loci for device-measured physical activity and sleep duration. *Nat. Commun.* 9, 5257. [PubMed: 30531941]
37. Meddens SFW, de Vlaming R, Bowers P, Burik CAP, Linnér RK, Lee C, Okbay A, Turley P, Rietveld CA, Fontana MA, et al. (2021). Genomic analysis of diet composition finds novel loci and associations with health and lifestyle. *Mol. Psychiatry* 26, 2056–2069. [PubMed: 32393786]
38. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, Maller J, Sklar P, de Bakker PIW, Daly MJ, et al. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* 81, 559–575. [PubMed: 17701901]

### Manuscript Highlights

1. Identified 45 pairs of viral exposures associated with increased risk of NDDs
2. Replicated 22 of the viral exposures/NDD pairings
3. Replicated the previously reported Epstein-Barr and multiple sclerosis association
4. Follow-up shows significantly elevated risk of NDD years after viral exposure



**Figure 1: Graphical Summary of Hazard Ratio Lag for Replicated Viral/NDD Pairs**  
 Graphical summary of hazard ratio lag for replicated associations between viral exposures and neurodegenerative diseases shows increased risk of neurodegeneration even 15 years prior to disease onset. The size of the circle on each line represents relative sample size overlap for the viral exposure and the associated NDD at that time point. The position of the circle on each line denotes the HRs with whiskers showing the 95% CIs, represented on a log scaled x axis. The association between Epstein-Barr virus and multiple sclerosis was included at the top as a reference for scale.

**Table 1:**

## Summary of Participants in the Study

NDD	FinnGen (discovery phase)			UKB (replication phase)		
	All	Female	Male	All	Female	Male
AD	9301	4079	5222	2342	1199	1143
ALS	483	181	302	357	209	148
dementia	16,499	6984	9515	2200	981	1219
MS	2182	1648	534	1349	985	364
PD	4235	1658	2577	2998	1106	1892
vascular	2335	778	1557	430	152	278
controls	309,154	173,746	135,408	96,390	48,919	47,471

Summary of participants analyzed in this study that includes over 400,000 samples across two national biobanks. In FinnGen, the control numbers listed below are the general population count (total); in the UKB these include exclusions described in the methods section such as concurrent neurodegenerative disease diagnoses.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript



**Table 2:**

Discovery (FinnGen) and Replication (UKB) Results for Replicated Viral/NDD Pairs

NDD	Virus_Description	Cohort	hr/or	ci_min	ci_max	P	N	P, FDR
AD	Viral encephalitis, not elsewhere classified/ unspecified*	FinnGen	30.72	11.84	79.68	1.89E-12	24	1.26E-11
AD	Viral encephalitis, not elsewhere classified/ unspecified*	UKB	22.06	5.47	88.94	1.37E-05	3	5.47E-05
AD	Viral and other specified intestinal infections	FinnGen	3.03	1.73	5.32	1.08E-04	122	3.59E-04
AD	Viral and other specified intestinal infections	UKB	3.09	1.93	4.96	2.63E-06	19	1.32E-05
AD	Influenza and pneumonia*	FinnGen	4.11	3.38	4.99	4.01E-46	2141	8.03E-45
AD	Influenza and pneumonia*	UKB	2.60	2.25	2.99	5.47E-40	231	1.09E-38
AD	Meningitis*	FinnGen	2.81	1.39	5.67	3.95E-03	33	7.47E-03
AD	Meningitis*	UKB	62.20	18.35	210.78	3.29E-11	6	2.20E-10
ALS	Influenza and pneumonia*	FinnGen	1.81	1.35	2.43	7.09E-05	53	7.09E-05
ALS	Influenza and pneumonia*	UKB	7.91	6.02	10.40	1.08E-49	76	5.38E-49
DEM	Viral encephalitis, not elsewhere classified/ unspecified*	FinnGen	6.56	1.60	26.97	9.05E-03	48	2.17E-02
DEM	Viral encephalitis, not elsewhere classified/ unspecified*	UKB	40.12	12.09	133.14	1.61E-09	5	5.07E-09
DEM	Other viral diseases, not elsewhere classified*	FinnGen	2.01	1.20	3.36	7.98E-03	144	2.17E-02
DEM	Other viral diseases, not elsewhere classified*	UKB	3.24	2.19	4.81	5.02E-09	27	1.38E-08
DEM	Viral warts	FinnGen	3.70	2.23	6.14	3.83E-07	74	3.07E-06
DEM	Viral warts	UKB	2.38	1.29	4.39	5.31E-03	11	7.30E-03
DEM	All influenza*	FinnGen	5.07	3.06	8.41	3.04E-10	348	3.65E-09
DEM	All influenza*	UKB	4.56	2.56	8.13	2.55E-07	13	5.10E-07
DEM	Influenza and pneumonia*	FinnGen	2.87	2.40	3.43	6.78E-31	3174	1.63E-29
DEM	Influenza and pneumonia*	UKB	6.12	5.48	6.85	1.63E-222	439	3.59E-221
DEM	Viral pneumonia*	FinnGen	3.48	2.00	6.04	1.01E-05	104	6.04E-05
DEM	Viral pneumonia*	UKB	4.44	1.90	10.36	5.59E-04	6	8.78E-04
MS	Herpesviral [herpes simplex] infections*	FinnGen	1.91	1.14	3.22	1.45E-02	19	2.60E-02
MS	Herpesviral [herpes simplex] infections*	UKB	4.95	1.28	19.12	2.03E-02	3	4.05E-02
MS	Varicella-zoster virus (Zoster [herpes zoster])*	FinnGen	2.12	1.09	4.12	2.66E-02	11	3.98E-02
MS	Varicella-zoster virus (Zoster [herpes zoster])*	UKB	3.76	1.18	12.00	2.54E-02	5	4.57E-02
PD	Viral hepatitis*	FinnGen	2.88	1.37	6.05	5.13E-03	14	1.37E-02
PD	Viral hepatitis*	UKB	2.20	1.06	4.55	3.43E-02	8	4.66E-02
PD	Viral infections characterized by skin and mucous membrane lesions	FinnGen	1.58	1.18	2.13	2.44E-03	77	1.37E-02

NDD	Virus_Description	Cohort	hr/or	ci_min	ci_max	P	N	P, FDR
PD	Viral infections characterized by skin and mucous membrane lesions	UKB	2.46	1.70	3.56	1.88E-06	31	5.10E-06
PD	All influenza (not pneumonia) *	FinnGen	2.01	1.23	3.31	5.71E-03	36	1.37E-02
PD	All influenza (not pneumonia) *	UKB	3.69	2.06	6.59	1.05E-05	13	2.49E-05
PD	All influenza *	FinnGen	1.84	1.20	2.83	5.48E-03	51	1.37E-02
PD	All influenza *	UKB	4.31	2.50	7.43	1.55E-07	15	4.91E-07
PD	Influenza and pneumonia *	FinnGen	1.72	1.48	1.99	1.27E-12	418	1.52E-11
PD	Influenza and pneumonia *	UKB	2.98	2.62	3.38	1.34E-63	294	1.27E-62
VAS	Viral and other specified intestinal infections	FinnGen	2.79	1.59	4.89	3.61E-04	45	5.06E-04
VAS	Viral and other specified intestinal infections	UKB	5.26	2.31	11.95	7.37E-05	6	1.37E-04
VAS	Varicella-zoster virus (Zoster [herpes zoster]) *	FinnGen	2.33	1.25	4.36	7.77E-03	33	9.07E-03
VAS	Varicella-zoster virus (Zoster [herpes zoster]) *	UKB	6.22	2.27	17.00	3.71E-04	4	6.03E-04
VAS	All influenza *	FinnGen	2.96	1.66	5.27	2.36E-04	71	4.13E-04
VAS	All influenza *	UKB	4.99	1.57	15.83	6.39E-03	3	6.92E-03
VAS	Influenza and pneumonia *	FinnGen	4.62	3.81	5.59	1.73E-55	708	1.21E-54
VAS	Influenza and pneumonia *	UKB	6.79	5.40	8.53	9.63E-61	103	1.25E-59

Discovery (FinnGen, reporting hazard ratios) and replication (UKB, reporting odds ratios) analyses showing the 22 replicated associations between virus exposures and neurodegenerative diseases. In this table N reflects the overlapping count of samples with both virus and neurodegenerative disease exposures. The final column for each phase denotes the FDR corrected p-value. Meningitis includes both viral and bacterial codes.

\* denotes potentially neurotropic virus coding derived from EMR. As a note, “influenza and pneumonia” is a super-set of “all influenza”.

**Table 3:**

Results for Lag Time Associations for Replicated Viral/NDD Pairs

NDD	Virus_Description	Years preceding NDD Onset (lag)	HR	95% CI low	95% CI high	P	N	P, FDR
AD	Influenza and pneumonia	15	1.91	1.51	2.40	4.55E-08	591	1.43E-07
AD	Influenza and pneumonia	5	6.76	5.38	8.48	3.28E-61	662	7.71E-60
AD	Influenza and pneumonia	1	23.00	16.42	32.20	1.95E-74	681	9.16E-73
AD	Meningitis	15	4.98	2.48	10.02	6.57E-06	11	1.82E-05
AD	Viral and other specified intestinal infections	5	5.99	3.41	10.52	4.65E-10	30	1.82E-09
AD	Viral and other specified intestinal infections	1	11.97	5.09	28.17	1.31E-08	30	4.38E-08
ALS	Influenza and pneumonia	5	2.85	1.61	5.03	3.15E-04	19	4.21E-04
ALS	Influenza and pneumonia	1	7.67	2.85	20.63	5.47E-05	11	1.42E-04
DEM	All influenza	5	8.22	5.02	13.47	5.79E-17	160	3.71E-16
DEM	All influenza	1	25.79	17.20	38.65	8.04E-56	100	1.71E-54
DEM	Influenza and pneumonia	15	1.49	1.23	1.82	6.25E-05	963	1.74E-04
DEM	Influenza and pneumonia	5	6.30	5.29	7.50	1.95E-94	1096	6.23E-93
DEM	Influenza and pneumonia	1	16.41	13.07	20.61	6.93E-128	800	4.43E-126
DEM	Other viral diseases, not elsewhere classified	5	2.94	1.66	5.24	2.38E-04	32	5.44E-04
DEM	Other viral diseases, not elsewhere classified	1	18.94	10.19	35.20	1.39E-20	27	9.85E-20
DEM	Viral encephalitis, not elsewhere classified/unspecified	15	4.95	2.18	11.21	1.28E-04	12	3.15E-04
DEM	Viral encephalitis, not elsewhere classified/unspecified	5	24.14	13.23	44.06	3.32E-25	18	3.03E-24
DEM	Viral encephalitis, not elsewhere classified/unspecified	1	83.03	36.29	189.93	1.22E-25	16	1.31E-24
DEM	Viral pneumonia	5	7.34	3.99	13.49	1.40E-10	39	6.91E-10
DEM	Viral pneumonia	1	37.57	21.92	64.40	9.69E-40	30	1.55E-38
DEM	Viral warts	15	2.59	1.54	4.34	3.14E-04	35	6.92E-04
DEM	Viral warts	5	4.95	2.83	8.66	2.09E-08	18	8.38E-08
PD	All influenza	5	2.50	1.49	4.18	4.97E-04	18	2.38E-03
PD	All influenza (not pneumonia)	5	3.07	1.60	5.88	7.16E-04	12	2.86E-03
PD	Influenza and pneumonia	5	2.24	1.78	2.82	4.36E-12	126	3.49E-11
PD	Influenza and pneumonia	1	6.20	4.24	9.05	4.35E-21	71	1.04E-19
PD	Viral infections characterized by skin and mucous membrane lesions	5	2.37	1.50	3.74	2.02E-04	24	1.21E-03
VAS	Influenza and pneumonia	15	1.82	1.36	2.45	6.53E-05	180	1.52E-04
VAS	Influenza and pneumonia	5	5.43	3.96	7.44	8.56E-26	241	3.99E-25
VAS	Influenza and pneumonia	1	21.00	14.01	31.48	3.30E-49	233	2.31E-48

NDD	Virus_Description	Years preceding NDD Onset (lag)	HR	95% CI low	95% CI high	P	N	P, FDR
VAS	Viral and other specified intestinal infections	5	6.78	2.57	17.94	1.14E-04	15	1.99E-04
VAS	Viral and other specified intestinal infections	1	20.31	7.65	53.88	1.46E-09	10	4.10E-09

Detailed analysis of significant hazard ratio lags over time for replicated association showing increased risk of neurodegeneration associated with viral exposures. Data for AD (viral encephalitis), MS (herpesviral and zoster), PD (viral hepatitis), and VAS (zoster and all influenza) were only available for all follow-up durations and not at the granular level, so they are not included in this table. In this table N reflects the overlapping count of samples with both virus and neurodegenerative disease exposures at that point in the follow-up duration.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**Key resources table**

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
FinnGen	Ristey's portal	<a href="https://r9.risteys.finnngen.fi/">https://r9.risteys.finnngen.fi/</a>
UK Biobank	UK Biobank	<a href="https://www.ukbiobank.ac.uk/">https://www.ukbiobank.ac.uk/</a> Application Number 33601
PRS – physical activity	Doherty et al. <sup>36</sup>	N/A
PRS – dietary protein imbalance	Meddens et al. <sup>37</sup>	N/A
Code for modeling and analysis	This paper	<a href="https://zenodo.org/badge/latestdoi/499275067">https://zenodo.org/badge/latestdoi/499275067</a>
Software and algorithms		
Python version 3.8	Python Software Foundation	<a href="https://www.python.org/">https://www.python.org/</a>
Plink 1.9	Purcell et al. <sup>38</sup>	<a href="https://www.cog-genomics.org/plink/">https://www.cog-genomics.org/plink/</a>
Benjamini-Hochberg false discovery rate correction (FDR)	Benjamini and Hochberg <sup>33</sup>	N/A

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript