Herpes zoster vaccination and new diagnoses of dementia: A quasi-randomized study in Australia

Michael Pomirchy¹, Christian Bommer¹, Fabienne Pradella¹,²,³, Felix Michalik¹,², Ruth Peters⁴,⁵,⁶,⁷, Pascal Geldsetzer¹,⁸,⁹,*

*Corresponding author. Email: pgeldsetzer@stanford.edu

Affiliations:
¹ Division of Primary Care and Population Health, Department of Medicine, Stanford University, Stanford, California, USA
² Heidelberg Institute of Global Health, Heidelberg University, Heidelberg, Germany
³ Gutenberg School of Management and Economics, Mainz University, Mainz, Germany
⁴ Ageing and Neurodegeneration, Neuroscience Research Australia, Sydney, Australia
⁵ School of Psychology, University of New South Wales, Sydney, Australia
⁶ Ageing Futures Institute, University of New South Wales, Sydney, Australia
⁷ Neurology, The George Institute for Global Health, Sydney, Australia
⁸ Department of Epidemiology and Population Health, Stanford University, Stanford, California, USA
⁹ Chan Zuckerberg Biohub – San Francisco, San Francisco, California, USA

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.
Abstract:
Increasing evidence suggests that neurotropic herpesviruses could play a role in the development of dementia, possibly through a neuroinflammatory process. Herpes zoster (HZ) vaccination has been reported to lead to a reduced probability of being diagnosed with dementia in several correlational studies and in a prior analysis by our team in Wales. This present study constitutes the first investigation to use a quasi-randomized study design in an electronic health record dataset from a large and diverse nation (Australia) to aim to determine the effect of HZ vaccination on dementia. In Australia, starting on November 1 2016, live-attenuated HZ vaccination was provided for free to individuals aged 70 to 79 years of age through primary care providers. Thus, those whose 80th birthday was just a few days prior to November 1 2016 never became eligible, whereas those whose 80th birthday was just a few days later were eligible. The key advantage of our approach is that one would not expect that these population groups who differ in their age by only a minute degree would, on average, differ in any of their health characteristics and behaviors. We used detailed primary healthcare records with week-of-birth information from 65 general practices across Australia. We analyzed our data using a regression discontinuity approach. Our sample consisted of 101,219 patients. As expected, patients born just before versus shortly after the date-of-birth eligibility threshold (November 2 1936) for HZ vaccination were well-balanced in their past preventive health services uptake and chronic disease diagnoses. There was an abrupt increase of 15.7 (95% CI: [12.2 – 19.3], p < 0.001) percentage points in the probability of ever receiving HZ vaccination between patients born shortly before versus shortly after the eligibility threshold. The eligibility rules of the HZ vaccination program, thus, created comparison groups just on either side of the date-of-birth eligibility threshold who were similar to each other, except for a large difference in their probability of receiving the intervention (HZ vaccination) of interest. Eligibility for HZ vaccination (i.e., being born shortly before versus shortly after November 2 1936) decreased the probability of receiving a new dementia diagnosis over 7.4 years by 2.0 percentage points (95% CI: [0.3 – 3.7], p = 0.021). Being eligible for HZ vaccination did not affect the probability of taking up other preventive health services (including other vaccinations), nor the probability of being diagnosed with other common chronic conditions than dementia. This study provides important evidence on the potential benefits of HZ vaccination for dementia because its quasi-randomized design allows for conclusions that are more likely to be causal than those of the existing associational evidence.
Introduction

Herpesviruses have long been thought to potentially play a causative role in the development of dementia because they are neurotropic (1), their reactivations in the nervous system become more common with age and can lead to encephalitis (1,2), and they are a near ubiquitous exposure (1). This hypothesis has recently garnered increasing attention (2–5), partly because of the finding that herpesviruses can seed β-amyloid in mice (6, 7) and the evidence for a causal role of the human herpesvirus 4 (the Epstein-Barr virus) in the development of multiple sclerosis (8). Currently, a phase 2 proof-of-concept trial, funded by the US National Institute on Aging, is studying the impact of an antiviral drug against herpesviruses on the cognitive and functional ability of individuals with mild Alzheimer’s dementia (9). An alternative strategy to target herpesviruses is vaccines. Vaccination may be a particularly promising strategy because the immune system likely plays a key role in the development of dementia (10), and there is evidence, especially in the case of live-attenuated vaccines, that vaccines have important off-target health effects induced by broader immune mechanisms (11–13).

Several correlational studies in electronic health record data have investigated the association between herpes zoster (HZ) vaccination and diagnoses of dementia, with all reporting a protective association (14–22). However, these studies are subject to the fundamental concern that those who opt to be vaccinated differ from those who do not in a variety of characteristics that are difficult to measure (23). For instance, detailed information on health behaviors that are likely related to both dementia and vaccination, such as physical activity and diet (24, 25), are virtually never available in electronic health record data. Most recently, using electronic health record data from Wales, our group instead used a quasi-experimental approach to also find that HZ vaccination reduced the incidence of new dementia diagnoses (26).

This is the first study to use a quasi-randomized study design in an electronic health record dataset from a large and diverse nation (Australia) to investigate the effect of HZ vaccination on dementia incidence. Our quasi-randomized design takes advantage of the fact that those aged 70 to 79 years on November 1 2016, when the Australian National Immunisation Programme (NIP) started its HZ vaccination program, were eligible for free live-attenuated HZ vaccination (Zostavax [Merck]) (27-29). Thus, individuals who had their 80th birthday just prior to, or on, November 1 2016 (i.e., born before November 2 1936) were ineligible for HZ vaccination whereas those who had their 80th birthday just after November 1 2016 were eligible. This eligibility rule resulted in an abrupt increase in the probability of ever receiving the HZ vaccine between individuals who differed in their age by merely a week across the date of birth-based eligibility threshold for the vaccination program. The Australian setting, thus, allows for the comparison of dementia incidence between eligible and ineligible groups of individuals who are not expected to differ in their characteristics other than a difference in age of merely a few weeks and a large difference in ever receiving the HZ vaccine. Our analysis approach is, therefore, more likely to reflect a causal relationship than would be the case in more standard epidemiological studies that compare those who received the vaccine with those who did not receive the vaccine whilst assuming that the many differences between these groups (e.g., healthy vaccinee bias (30, 31)) were perfectly measured and modelled in the analysis.

Using electronic health record data from 65 primary care practices across Australia, we first show that the eligibility rules for the HZ vaccination program created a large difference in the probability of ever receiving the HZ vaccine between individuals born just before versus just after the date-of-birth eligibility threshold (November 2 1936). We then show that, over a 7.4-year follow-up period, those who were born on or just after November 2 1936 (and thus eligible for free vaccination) compared to those born just before (and thus ineligible for life) had a lower probability of being newly diagnosed with dementia. Lastly, we provide evidence that these
findings are unlikely to be driven by confounding factors, such as by demonstrating that there were no differences in health characteristics between individuals just on either side of the November 2 1936 eligibility threshold, the effects of the vaccine were specific to dementia (as opposed to a broader set of chronic conditions), and HZ vaccination did not lead to increased uptake of other preventive health measures. We also provide evidence against the possibility that a different intervention or policy used the identical date-of-birth eligibility cutoff (November 2 1936) as the HZ vaccination program.

Methods
The herpes zoster vaccine rollout in Australia:
Australia's National Immunisation Programme (NIP), first introduced in 1997, is a collaborative program between the Australian State and Territory governments that provides free vaccines to eligible individuals with the goal of preventing diseases (32). The NIP schedule, which is non-compulsory, sets out a list of government-funded vaccines and immunizations that patients may take at various stages of life, starting at birth. In addition, state and territory health departments may fund vaccines not covered by the NIP and implement their own immunization schedules (33).

The NIP for HZ vaccination started on November 1 2016 (27). As of this date, the live-attenuated single-dose HZ vaccine (Zostavax [Merck]) was provided free of charge nationwide in Australian primary care practices for certain eligible age groups. These eligible age groups were those aged 70 years and, through a catch-up program that lasted until October 31 2023, individuals aged 71-79 years. Thus, individuals born on or after November 2 1936 (i.e., those who had their 80th birthday after November 1 2016) were eligible for free HZ vaccination, whereas those born before November 2 1936 (i.e., those who had their 80th birthday before or on November 1 2016) were ineligible and remained ineligible for life. The rollout of HZ vaccination in Australia, therefore, resulted in groups of individuals close to either side of the date-of-birth eligibility threshold for HZ vaccination who are expected to be similar to each other in observed and unobserved characteristics, except for a small difference in age and their eligibility status for HZ vaccination. Further information on the HZ vaccination rollout in Australia is available elsewhere (27-29).

Data source:
We used data from PenCS (34), an Australian-owned health informatics company, which provides detailed primary care electronic health records to researchers. The data included diagnoses, immunizations and other healthcare procedures, as well as prescribed medications from 65 general practitioner (GP) practices across each of Australia’s six states and the Australian Capital Territory. These were GP practices that voluntarily consented to share their electronic health record data for research. 27 GP practices were located in New South Wales, 17 in Queensland, one in South Australia, one in Tasmania, 14 in Victoria, four in Western Australia, and one in the Australian Capital Territory. As classified by the Modified Monash Model (35), 14 practices were in a metropolitan area, 40 in regional centers, 10 in small rural towns, and one in a remote community. The practices ranged in size from approximately 200 to 4,000 patients, with the average practice having 1,672 patients.

General practitioners function as gatekeepers in the Australian healthcare system such that patients generally only qualify for the Medicare Benefits Schedule for specialist care after a referral from a GP (36). The dataset contained data for each patient who visited one of the 65 GP practices between February 15 1993 and March 27 2024, whereby unique patient identifiers allow for tracking of patients across GP practices within the PenCS network. For the purposes
of our analysis, PenCS provided us with patients’ dates of birth in weeks. As is customary in Australia’s primary care records, diagnoses were coded by PenCS using open-ended text fields provided by the GP. The text fields used to define each diagnosis in our analysis are listed in Table S1. PenCS does not link any of its primary care records to hospital records or mortality registers. The data contained date of death as recorded by the GP practice.

This study did not use the MedicineInsight database because it does not provide date of birth at a more granular level than years and is not currently available for research (37, 38).

**Outcome and exposure definitions:**

Our follow-up period began on the start date (November 1 2016) of the HZ vaccination program. Our dataset ended in March 27 2024, which marked the end of the follow-up period.

Our outcome of interest was new diagnoses of dementia made during the follow-up period. If more than one diagnosis for dementia was recorded for an individual patient, we used the date of the first diagnosis. This approach of using the date of the first diagnosis was also used for defining the date of all other diagnoses in our analyses. Given the neuropathological overlap between dementia types and the difficulty in distinguishing dementia types clinically (39-41), as well as our reduced statistical power when studying less common outcomes, we defined dementia as dementia of any type or cause. The codes used to define dementia (as well as all other diagnoses used in our analyses) are listed in Table S1.

Our exposure was eligibility for free HZ vaccination as determined by an individual’s date of birth. Week of birth in our data was coded such that each week started on a Monday. Because November 2 1936 was also a Monday, we were thus able to determine the eligibility status of each patient in the data.

**Statistical analysis**

*Regression discontinuity:* Our analysis is based on the rationale that individuals born very close to either side of the November 2 1936 threshold are expected to be similar to each other in observed and unobserved characteristics except for their eligibility status for HZ vaccination. We tested for differences in our outcomes at the November 2 1936 date-of-birth eligibility threshold for HZ vaccination using regression discontinuity (RD) analysis, which is a well-established statistical technique for causal effect estimation (42). Regression discontinuity provides an unbiased effect estimate as long as any confounding variables do not abruptly change at exactly the November 2 1936 threshold (43, 44). This assumption was unlikely to be violated in this study because there was, to our knowledge, no other relevant policy or intervention that used the identical date of birth threshold as its eligibility criterion as the HZ vaccination program. As described below, we conducted a series of tests to further substantiate that this assumption was met.

As per recommended practice (42-45), we used local linear regression with triangular kernel weights on observations within a mean squared error (MSE)-optimal bandwidth of the date-of-birth eligibility threshold. In robustness checks, we also implemented our analysis using different bandwidth choices, local quadratic regression, and uniform weights. Local linear regression is the recommended and most robust approach for RD analyses even in situations in which the relationship between the assignment variable (here, date of birth) and the outcome is exponential (45). Triangular kernel weights give more weight to those observations closer to the eligibility threshold and less weight to observations further away (44). The MSE-optimal criterion is used as an objective criterion to balance precision and bias in estimation (44). We calculated the MSE-optimal bandwidth for each analysis separately.
We tested for an effect heterogeneity by gender by i) performing the analysis separately among women and men, and ii) running an interaction model that measured the difference in effects between men and women. The regression equations for all analyses are provided in Text S1.

We did not use a competing risk model for several reasons. First, in the absence of the HZ vaccination program, there is no reason that the competing risk of death should differ across the November 2 1936 date-of-birth eligibility threshold. Second, not adjusting for competing risk of death in our setting is a conservative choice because eligibility for HZ vaccination may reduce (but is very unlikely to increase) all-cause mortality (46, 47). Thus, those eligible for HZ vaccination will, on average, be exposed to a longer time period during which they could become newly diagnosed with dementia. Third, to date, no well-established approach exists for competing risk and survival models in an RD framework (48).

Given its implementation using local linear regression, the effect estimates obtained in RD are absolute effect estimates. We, thus, consistently reported the absolute (in percentage points) rather than the relative effects of HZ vaccination eligibility.

Comparative regression discontinuity:
One common drawback of RD analyses is the often relatively high level of imprecision in estimating effects at the date-of-birth-eligibility threshold (49). To improve precision, we used a variant of the RD design, called comparative RD (CRD), in secondary analyses, by leveraging an additional source of untreated data from older patients in our sample. By adding these data, CRD tends to provide increased statistical power relative to standard RD (49). This improved statistical efficiency can yield point estimates that are more similar to those from randomized controlled trials than those from standard RD (50, 51).

In addition to the main cohort of patients who were born in close proximity to the date of birth-based eligibility threshold, our CRD design analyzed a comparison cohort of patients that was always ineligible for free HZ vaccination. As a result, in our CRD, there were two sets of vaccine-ineligible individuals: i) ineligible patients in the main cohort (i.e., those born before the date of birth-based eligibility threshold); and ii) patients in the comparison cohort. For our comparison cohort, we used the youngest cohort of patients that was older (and thus always ineligible) than the patients born within the MSE-optimal bandwidth around the November 2 1936 eligibility threshold for free HZ vaccination. We defined our comparison cohort using the same MSE-optimal bandwidth size as we used for ineligible patients in our main cohort. To account for the age difference between the main and comparison cohort, we delayed the start of the follow-up period for identifying new dementia diagnoses for the comparison cohort by the same number of weeks as were contained in the MSE-optimal bandwidth for the main cohort. We implemented our CRD analysis using the bandwidth that was considered MSE-optimal (for a given combination of sample and outcome definition) under the standard RD design. We used local linear regression and uniform kernel weights.

Our CRD approach assumed that the relationship between age and dementia incidence was similar for patients in our comparison cohort as for patients ineligible for HZ vaccination in our main cohort. To evaluate this assumption, we compared the trends of age with the incidence of new dementia diagnoses within the MSE-optimal bandwidth around the November 2 1936 eligibility threshold between patients in the comparison cohort and patients ineligible for free HZ vaccination in the main cohort. In addition to visual inspection, we did this by testing whether the age-dementia trends were statistically significantly different from each other between these two
patient groups using regression analysis. Details on these regressions are provided in Text S1. We found no evidence that the assumption for the valid use of CRD was violated (p = 0.387).

**Baseline balance checks:**
The robustness of our study design to confounding rests on the intuition that potential confounding variables are unlikely to change abruptly (i.e., display discontinuities) precisely at the November 2 1936 eligibility threshold. To test the validity of this assumption empirically, we conducted a series of baseline balance checks by testing for differences in outcomes at the November 2 1936 date-of-birth threshold. We used the identical analysis approach as for our main outcome analyses except that we used the incidence of the outcome at any time prior to, rather than after, the start of the HZ vaccination program on November 1 2016. We used three sets of outcomes for our baseline balance checks. The first set of outcomes was the 15 most common clinical diagnoses in our data. The second set was indicators of prior uptake of preventive health services that were available in our data. These were uptake of common vaccinations in older age other than for HZ, like influenza vaccination, the pneumococcal polysaccharide vaccine, and the diphtheria, pertussis, and tetanus (DPT) vaccine, use of antihypertensive medication, use of statin medications, and participation in colorectal or breast cancer screening (defined, as per Australian cancer screening guidelines (52, 53), as uptake of fecal occult blood testing for colorectal cancer screening and mammography for breast cancer screening). The third set of outcomes were dementia diagnoses prior to November 1 2016 and the prevalence of known risk factors for dementia available in our data. These risk factors were obesity, current smoking, hypertension, diabetes, hyperlipidemia, antihypertensive use, and statin use.

**Testing for confounding:**
The key advantage of our RD approach is that a potential confounding variable only biases our analysis if it changes abruptly (i.e., displays a discontinuity) at exactly the November 2 1936 date-of-birth eligibility threshold (43, 44). Such a discontinuity could occur if another intervention also used November 2 1936 as its date-of-birth eligibility criterion and had an effect on dementia incidence. We conducted two types of tests to investigate the possible presence of such a competing intervention. First, we implemented the same RD analysis as for new dementia diagnoses (i.e., our primary analysis) for new diagnoses of each of the 15 most common clinical diagnoses during the follow-up period in our data. If another intervention existed that used the identical date-of-birth eligibility threshold as the HZ vaccination program and was not specific to dementia, then we may expect this intervention to also affect health outcomes other than dementia. Second, we reasoned that if November 2 was used as an annual date-of-birth eligibility threshold by another intervention, then we would expect to see effects on dementia of the November 2 threshold not merely for the birth year 1936, but also for other birth years. We, thus, conducted the same RD analysis for new dementia diagnoses as we did for the November 2 1936 date-of-birth eligibility threshold (i.e., our primary analysis) for each of the three years prior to and after 1936 (i.e., date-of-birth eligibility thresholds of November 2 1933, November 2 1934, November 2 1935, November 2 1937, November 2 1938, and November 2 1939). In these tests, given that the maximum follow-up period was shorter for more recent birth years, we restricted the follow-up period for each test to four years such that each test had the same length of follow-up period. However, we additionally conducted these tests using the maximum follow-up period (i.e., until March 27 2024) available for each test.

**Robustness checks:**
We conducted additional robustness checks to those described above. First, we implemented our analysis among “active” patients only, reasoning that delays in the diagnosis of dementia among this patient cohort are likely to be less common. Using the definition of the Royal
Australian College of General Practitioners (RACGP), we considered those patients as “active” who visited their GP at least three times in the two years preceding the end of our dataset (54). This group comprised 61.2% (n=61,903 patients) of our primary analysis cohort. Second, we implemented our analysis both when including and when excluding those patients who had received a diagnosis of dementia prior to the start of the HZ vaccination program on November 1 2016. Third, we verified that our findings were robust to different choices of i) grace periods (i.e., time periods since November 1 2016 after which follow-up time was considered to begin to allow for the time needed for a full immune response to develop after vaccine administration), ii) the bandwidth drawn around the November 2 1936 threshold by using bandwidths ranging from one-half to twice the MSE-optimal bandwidth used in our primary analysis, iii) kernel weights by using both uniform and triangular kernel weights, and iv) functional form assumptions by implementing the analysis using both local linear and quadratic polynomial regression.

This research was approved by the Stanford University Institutional Review Board on June 9 2023 and considered minimal risk (protocol number: 70277).

**Results**

**Characteristics of the study population:**

Our dataset contained data on 101,219 unique patients. Table 1 shows the sociodemographic and clinical characteristics of this sample. We additionally show the characteristics of the 18,402 patients born within the MSE-optimal bandwidth (for our primary analysis on the effect of HZ vaccination on new diagnoses of dementia) of 482 weeks around the November 2 1936 date-of-birth eligibility threshold.
Table 1. Sample characteristics

<table>
<thead>
<tr>
<th>Sociodemographic characteristics</th>
<th>Total, n (%)</th>
<th>Female, n (%)²</th>
<th>Male, n (%)²</th>
<th>Total, n (%)</th>
<th>Female, n (%)²</th>
<th>Male, n (%)²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Married</td>
<td>31,295 (30.9)</td>
<td>15,197 (28.5)</td>
<td>16,071 (34.8)</td>
<td>5,402 (29.4)</td>
<td>2,245 (22.5)</td>
<td>3,153 (38.6)</td>
</tr>
<tr>
<td>Aboriginal or Torres Strait Islander</td>
<td>1,233 (1.2)</td>
<td>662 (1.2)</td>
<td>571 (1.2)</td>
<td>128 (0.7)</td>
<td>73 (0.7)</td>
<td>55 (0.7)</td>
</tr>
<tr>
<td>Foreign-born</td>
<td>1,983 (2)</td>
<td>1,064 (2)</td>
<td>917 (2)</td>
<td>455 (2.5)</td>
<td>243 (2.4)</td>
<td>211 (2.6)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Clinical diagnoses³</th>
<th>Total, n (%)</th>
<th>Female, n (%)²</th>
<th>Male, n (%)²</th>
<th>Total, n (%)</th>
<th>Female, n (%)²</th>
<th>Male, n (%)²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dementia</td>
<td>922 (0.9)</td>
<td>493 (0.9)</td>
<td>429 (0.9)</td>
<td>620 (3.4)</td>
<td>329 (3.3)</td>
<td>291 (3.6)</td>
</tr>
<tr>
<td>Hypertension</td>
<td>22,704 (24.4)</td>
<td>11,643 (21.8)</td>
<td>11,083 (23.9)</td>
<td>5,464 (29.7)</td>
<td>3,095 (31)</td>
<td>2,367 (29)</td>
</tr>
<tr>
<td>Hyperlipidemia</td>
<td>19,930 (19.7)</td>
<td>10,377 (19.4)</td>
<td>9,553 (20.6)</td>
<td>3,894 (21.2)</td>
<td>2,179 (21.7)</td>
<td>1,715 (21.1)</td>
</tr>
<tr>
<td>Heart disease</td>
<td>11,254 (11.1)</td>
<td>4,831 (9.1)</td>
<td>6,413 (13.9)</td>
<td>3,808 (20.7)</td>
<td>1,852 (18.5)</td>
<td>1,956 (23.9)</td>
</tr>
<tr>
<td>COVID-19</td>
<td>7,300 (7.2)</td>
<td>4,124 (7.7)</td>
<td>3,176 (6.9)</td>
<td>1,781 (9.7)</td>
<td>1,005 (10.1)</td>
<td>775 (9.5)</td>
</tr>
<tr>
<td>Osteoarthritis</td>
<td>11,633 (11.5)</td>
<td>6,839 (12.8)</td>
<td>4,788 (10.4)</td>
<td>3,202 (17.4)</td>
<td>1,888 (18.9)</td>
<td>1,313 (16.1)</td>
</tr>
<tr>
<td>Non-hematological cancers</td>
<td>10,190 (10.1)</td>
<td>5,291 (9.9)</td>
<td>4,891 (10.6)</td>
<td>2,908 (15.8)</td>
<td>1,438 (14.4)</td>
<td>1,468 (18)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Respiratory conditions</th>
<th>Total, n (%)</th>
<th>Female, n (%)²</th>
<th>Male, n (%)²</th>
<th>Total, n (%)</th>
<th>Female, n (%)²</th>
<th>Male, n (%)²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diabetes mellitus</td>
<td>9,658 (9.5)</td>
<td>4,288 (8)</td>
<td>5,354 (11.6)</td>
<td>2,121 (11.5)</td>
<td>983 (9.8)</td>
<td>1,137 (13.9)</td>
</tr>
<tr>
<td>Osteoporosis</td>
<td>7,667 (7.6)</td>
<td>5,919 (11.1)</td>
<td>1,745 (3.8)</td>
<td>3,103 (16.9)</td>
<td>2,266 (22.7)</td>
<td>835 (10.2)</td>
</tr>
<tr>
<td>Gout</td>
<td>3,908 (3.9)</td>
<td>794 (1.5)</td>
<td>3,104 (6.7)</td>
<td>967 (5.3)</td>
<td>268 (2.7)</td>
<td>697 (8.5)</td>
</tr>
<tr>
<td>Stroke or TIA</td>
<td>2,349 (2.3)</td>
<td>1,104 (2.1)</td>
<td>1,242 (2.7)</td>
<td>952 (5.2)</td>
<td>480 (4.8)</td>
<td>472 (5.8)</td>
</tr>
<tr>
<td>Back pain</td>
<td>6,056 (6)</td>
<td>3,238 (6.1)</td>
<td>2,818 (6.1)</td>
<td>1,146 (6.2)</td>
<td>671 (6.7)</td>
<td>472 (5.8)</td>
</tr>
<tr>
<td>Depression</td>
<td>5,856 (5.8)</td>
<td>3,621 (6.8)</td>
<td>2,232 (4.8)</td>
<td>1,087 (5.9)</td>
<td>676 (6.8)</td>
<td>409 (5)</td>
</tr>
<tr>
<td>Hematological conditions</td>
<td>2,722 (2.7)</td>
<td>1,454 (2.7)</td>
<td>1,268 (2.7)</td>
<td>977 (5.3)</td>
<td>478 (4.8)</td>
<td>499 (6.1)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chronic kidney disease</th>
<th>Total, n (%)</th>
<th>Female, n (%)²</th>
<th>Male, n (%)²</th>
<th>Total, n (%)</th>
<th>Female, n (%)²</th>
<th>Male, n (%)²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Statin use</td>
<td>29,150 (28.8)</td>
<td>14,069 (26.4)</td>
<td>15,081 (32.6)</td>
<td>6,601 (35.9)</td>
<td>3,297 (33)</td>
<td>3,298 (40.3)</td>
</tr>
<tr>
<td>Antihypertensive use</td>
<td>20,117 (19.9)</td>
<td>10,252 (19.2)</td>
<td>9,863 (21.3)</td>
<td>5,897 (32)</td>
<td>3,277 (32.8)</td>
<td>2,617 (32)</td>
</tr>
<tr>
<td>PPV</td>
<td>18,153 (17.9)</td>
<td>9,786 (18.3)</td>
<td>8,367 (18.1)</td>
<td>6,702 (36.4)</td>
<td>3,621 (36.2)</td>
<td>3,080 (37.7)</td>
</tr>
<tr>
<td>Influenza vaccination</td>
<td>42,853 (42.3)</td>
<td>23,210 (43.5)</td>
<td>19,558 (42.3)</td>
<td>10,275 (55.8)</td>
<td>5,570 (55.7)</td>
<td>4,691 (57.4)</td>
</tr>
<tr>
<td>DPT vaccination</td>
<td>13,806 (13.6)</td>
<td>8,004 (15)</td>
<td>5,790 (12.5)</td>
<td>2,482 (13.5)</td>
<td>1,434 (14.4)</td>
<td>1,047 (12.8)</td>
</tr>
<tr>
<td>Cancer screening⁵</td>
<td>16,426 (16.2)</td>
<td>9,169 (17.2)</td>
<td>7,252 (15.7)</td>
<td>1,277 (6.9)</td>
<td>740 (7.4)</td>
<td>537 (6.6)</td>
</tr>
</tbody>
</table>

Abbreviations: MSE = Mean Squared Error, TIA = Transient Ischemic Attack; PPV = Pneumococcal Polysaccharide Vaccine; DPT = Diphtheria, Tetanus, and Pertussis

1 This is the MSE-optimal bandwidth used in our primary analysis for the effect of HZ vaccination on new diagnoses of dementia. The MSE-optimal bandwidth was 482 weeks.

2 1,620 (1.6%) patients in the entire sample and 234 (1.3%) patients in the MSE-optimal bandwidth had missing information on gender.

3 The clinical diagnoses shown are dementia and the 15 most common diagnoses ever recorded in our data. The codes used to define each condition are shown in Table S1.

4 The codes used to define each indicator of preventive health services uptake are shown in Table S1. All indicators were defined as being recorded at any point in our dataset.

5 Cancer screening refers to the uptake of colorectal or breast cancer screening. As per Australian cancer screening guidelines, this was defined as uptake of fecal occult blood testing (for colorectal cancer screening) and mammography (for breast cancer screening) (52, 53).

A one-week difference in age led to a large difference in HZ vaccination receipt:

We first provide evidence that eligibility for free HZ vaccination under the NIP led to an abrupt difference in HZ vaccination receipt at the November 2 1936 date-of-birth eligibility threshold. We found that adults born one week after the November 2 1936 date-of-birth eligibility cutoff had a 15.7 percentage point (95% CI: 12.2 – 19.3, p<0.001) higher probability of ever receiving the HZ vaccine than those born just one week earlier (Figure 1). Measured in the MSE-optimal
bandwidth of 255 weeks around the November 2 1936 threshold, the mean HZ vaccination probability was 6.5% (95% CI: 5.6 – 7.3) versus 30.2% (95% CI: 29.0 – 31.4) among those ineligible versus those eligible for the vaccine, respectively. There was no significant difference in the uptake of other preventive health services across the November 2 1936 date-of-birth threshold, including for other vaccines (influenza vaccination, PPV, and DPT vaccination), colorectal or breast cancer screening, statin use, and use of antihypertensive medications (Figure 1).
Fig. 1: The date-of-birth eligibility cutoff led to a large discontinuity in HZ vaccination receipt but there is no such jump at the cutoff in the uptake of other preventive interventions. 1,2,3

Abbreviations: HZ = Herpes Zoster; PPV = Pneumococcal Polysaccharide Vaccine; DPT = Diphtheria, Tetanus, and Pertussis

1 Cancer screening referred to the uptake of colorectal or breast cancer screening. As per Australian cancer screening guidelines, this was defined as uptake of fecal occult blood testing (for colorectal cancer screening) and mammography (for breast cancer screening) (52; 53).

2 Uptake of preventive health services were measured during our follow-up period of November 1 2016 (the start date of the HZ vaccination program) to March 27 2024.

3 Grey dots in panel A show the mean value for each 26-week increment in week of birth. In panels B through G, the increments were 31, 30, 47, 28, 21, and 36 weeks, respectively. Increments were derived by dividing the mean squared error-optimal bandwidth for each analysis by 10. The shading of the dots is in proportion to the weight that observations from each increment received in the analysis.

The HZ vaccination program’s eligibility rules, thus, created two comparison groups of individuals just on either side of the November 2 1936 date-of-birth eligibility threshold who were likely similar to each other except for a small difference in age and a large difference in the probability of having received HZ vaccination. To investigate this expectation further, we tested for each of the 15 most common diagnoses in our data whether there were any discontinuities at the November 2 1936 date-of-birth threshold in the probability of having received a given diagnosis prior to the start of the HZ vaccination program on November 1 2016. We also conducted the same tests for indicators of prior uptake of preventive health services, diagnoses of dementia, as well as risk factors for dementia on which we had information in our data (obesity, hyperlipidemia, hypertension, diabetes, current smoking, use of antihypertensive medications, and use of statins). We found no evidence of any discontinuities at the November 2 1936 threshold (Figure 2 and Figure S10). These tests, therefore, support the expectation that individuals just on either side of the November 2 1936 date-of-birth threshold were indeed similar to each other.
Fig. 2: Baseline balance checks using the 15 most common diagnoses in the PenCS data and indicators of preventive health services uptake.\textsuperscript{1,2,3,4}

Abbreviations: HZ = Herpes Zoster; PPV = Pneumococcal Polysaccharide Vaccine; DPT = Diphtheria, Tetanus, and Pertussis, TIA = Transient Ischemic Attack

1 Dots show the point estimate and horizontal bars the 95% confidence interval.
2 The codes used to define each condition are shown in Table S1.
3 Cancer screening refers to the uptake of colorectal or breast cancer screening. As per Australian cancer screening guidelines, this was defined as uptake of fecal occult blood testing (for colorectal cancer screening) and mammography (for breast cancer screening) (52, 53).
4 Note that COVID-19 is omitted from this analysis of clinical diagnoses because there was no incidence of COVID-19 prior to November 1 2016.

The effect of eligibility for HZ vaccination on new diagnoses of dementia:
Using our RD approach of comparing individuals just on either side of the November 2 1936 date-of-birth eligibility threshold, we find that eligibility for free HZ vaccination under the NIP decreased the probability of receiving a new dementia diagnosis over our 7.4-year follow-up period by 2.0 percentage points (95% CI: 0.3 – 3.7, \(p=0.021\); Figure 3). The effect was similar across follow-up periods ranging from four to seven years, and grace periods ranging from zero to 156 weeks (Figure 3). There was no evidence of a significant treatment effect heterogeneity by gender (Figure S1, Figure S2, and Text S1).
Fig. 3: The effect of being eligible for free HZ vaccination on new diagnoses of dementia. 1,2,3,4,5

Abbreviations: HZ = Herpes Zoster
1 With “grace periods” we refer to time periods since November 1 2016 after which follow-up time is considered to begin to allow for the time needed for a full immune response to develop after vaccine administration.
2 Red (as opposed to white) fillings denote statistical significance (p<0.05).
3 Grey vertical bars depict 95% confidence intervals.
4 Grey dots show the mean value for each 48-week increment in week of birth. The shading of the dots is in proportion to the weight that observations from this 48-week increment received in the analysis.
5 In Panel A, the sample size in the mean squared error-optimal bandwidth is 18,402 adults.

Robustness checks:
In addition to robustness across follow-up and grace periods shown in Figure 2, we detail in the Supplemental Materials that our results were robust to a series of additional checks. First, the effect estimates remained similar in magnitude when using i) uniform kernel weights instead of triangular kernel weights (Figure S3, Panel B), ii) local quadratic instead of local linear regression (Figure S3, Panel C), and iii) bandwidths between one-half and two times the MSE-optimal bandwidth (Figure S4). We conducted the same robustness checks (shown in Figures S5 and S6) for the effect of HZ vaccination eligibility on HZ vaccine uptake. Second, our secondary analysis approach (CRD) confirmed the findings from our primary approach. Specifically, using CRD, we estimated that HZ vaccination eligibility reduced the probability of a new diagnosis of dementia by 1.5 percentage points (95% CI: 0.2 – 2.7, p=0.020; Figure S7) over our 7.4-year follow-up period. Third, we also found a significant reduction in new diagnoses of dementia from HZ vaccination eligibility when restricting our study cohort to the 61,903 frequent primary care visitors (“active” patients) in our data (Figure S8). In fact, the protective effect for new dementia diagnoses was somewhat larger in this cohort (-3.5 [95% CI: -6.5 – -0.5] percentage points, p=0.024) than in our primary study cohort. Fourth, our results remained similar when excluding patients with a diagnosis of dementia recorded prior to the start of the HZ vaccination program (Figure S9).

Testing for confounding:
For our effect estimates to be unbiased, the key assumption that needs to be fulfilled is that no confounding variable changed abruptly at precisely the November 2 1936 date-of-birth eligibility threshold (43, 44). Such a discontinuity of a confounding variable at the November 2 1936 date-
of-birth threshold could occur if another intervention or policy used the identical date-of-birth eligibility criterion as the HZ vaccination program. We investigated this possibility in three ways.

First, because another intervention that used a November 2 1936 date-of-birth eligibility criterion and was not specific to dementia would be unlikely to only affect dementia diagnoses without also having an impact on other common diagnoses, we investigated whether being eligible for HZ vaccination based on one’s date of birth had an effect on common disease diagnoses other than dementia. Using the same RD approach as in our primary analysis for dementia, we conducted this test for new diagnoses of each of the 15 most common diagnoses in the PenCS data. Unlike with dementia, being eligible for HZ vaccination based on one’s date of birth had no significant effect on the incidence of any of these 15 conditions over our 7.4-year follow-up period (Figure 4, Panel A).

Second, we conducted the same analysis as for common clinical diagnoses shown in Figure 4 for indicators of preventive health service uptake. The rationale for these analyses was twofold: to investigate whether i) another intervention aimed at improving preventive health service use (e.g., another vaccination program) used a November 2 1936 date-of-birth eligibility criterion, and ii) HZ vaccination itself may have led to increased uptake of other preventive health services. For each of our indicators (influenza vaccination, pneumococcal vaccination, DPT vaccination, statin use, use of antihypertensive medications, and cancer screenings), we found no evidence that HZ vaccination eligibility affected preventive health services uptake (Figure 4, Panel B).

Third, if another intervention used a November 2 date-of-birth eligibility criterion, then we may expect to see differences in the effect of this threshold on new diagnoses of dementia for birth years other than 1936. We, thus, implemented the same analysis as for our primary analysis.
(shown in Figure 2), but shifted the date-of-birth eligibility threshold to each of the three years before and after 1936. We found that the only date-of-birth threshold that resulted in a reduction in new diagnoses of dementia was the threshold used by the HZ vaccination program (i.e., November 2 1936; Figure S11).

Discussion
The finding that HZ vaccination has a beneficial effect on the dementia disease process would be of great significance for both population health (given the availability of HZ vaccination as a simple, one-off, and inexpensive intervention (55–57)) and dementia research. In addition to the use of detailed electronic health record data from a diverse and large country, the fundamental advance of our study is its quasi-randomized study design. Using this approach, we provide evidence that is more likely to be causal than that reported in previous associational studies (14–22). Our study found similar protective effects from HZ vaccination for dementia incidence as has been reported in these associational studies, as well as in our analysis of data from Wales’s population of three million, mostly White (58), individuals (26).

Australia implemented its HZ vaccination program using a specific (maximum) date-of-birth eligibility threshold (27, 59), which created population groups that differed in their age by only a minute degree but had large differences in the probability of receiving the HZ vaccine. The rollout of the HZ vaccine, therefore, created two comparison groups who are likely to be similar to each other on observed and unobserved characteristics except for this difference in their probability of receiving HZ vaccination. The key strength of our study is, thus, that a potentially confounding variable can only bias our findings if it changes abruptly at precisely the date-of-birth eligibility threshold that was used for the HZ vaccination program in Australia (43, 44). Bias could, therefore, arise from the existence of an intervention that used the identical date-of-birth threshold (November 2 1936) as an eligibility criterion as the HZ vaccination program. However, we are not aware of any such intervention. Importantly, we have also provided several lines of evidence in our manuscript that such a competing intervention is unlikely to exist. First, such a competing intervention, like another vaccination program, would be unlikely to affect the incidence of dementia diagnoses without also having some effect on the incidence of other health conditions. We have shown that the date-of-birth eligibility threshold for HZ vaccination only affected the incidence of dementia diagnoses but none of the other 15 most common diagnoses in our data. Second, we have shown that the November 2 threshold that was used for the HZ vaccination program only has a significant protective effect for dementia incidence in the birth year (1936) that was used by the HZ vaccination program, but none of the three years preceding and succeeding 1936. Third, if another intervention that also used November 2 1936 as its eligibility threshold had been implemented prior to the start of the HZ vaccination program, then we may have expected to observe imbalances across the November 2 1936 threshold in existing disease diagnoses or past preventive health services uptake (e.g., vaccination coverage) at the time of the start of the HZ vaccination program. We found no evidence that such baseline imbalances exist in our data.

It is important to note that our conclusions are also unlikely to be affected by ascertainment bias. If attending the primary care provider for HZ vaccination provided an opportunity for the health system to identify previously undetected cases of dementia, our analysis would underestimate, rather than overestimate, the vaccine’s effectiveness in reducing the incidence of new diagnoses of dementia. Additionally, if healthcare visits for herpes zoster episodes were an important way for the health system to identify previously undiagnosed chronic conditions, we would have expected to see effects of HZ vaccination eligibility on a wider range of health outcomes beyond just dementia. We would have also expected a substantially smaller or absent
effect of HZ vaccination on the incidence of dementia diagnoses among patients who frequently visit their primary care provider because one additional healthcare visit is presumably less likely to have an important influence on diagnosing previously undetected dementia in this population. We, however, found no such pattern. Lastly, although it is possible that attending primary care for HZ vaccination led to increased uptake of other preventive health actions, we have shown that there was no effect of HZ vaccination eligibility on indicators of preventive health services uptake in our data, which included uptake of pneumococcal, influenza, and diphtheria-tetanus-pertussis vaccination, as well as the use of statins and antihypertensive medications.

We believe that our findings call for investments into further research in this area. Ideally, a clinical trial would be conducted to investigate the effect of HZ vaccination on the dementia disease process. However, we also believe that funding research into the mechanisms through which HZ vaccination could affect dementia is an important investment. Several potential mechanisms for this connection have already been recognized. For example, reactivations of the varicella zoster virus have been linked to long-lasting cognitive impairment through vasculopathy (60, 61), amyloid deposition and aggregation of tau proteins (62), neuroinflammation (63–66), as well as cerebrovascular disease resembling that seen in Alzheimer’s disease, including small to large vessel disease, ischemia, infarction, and hemorrhage (67, 68, 63–66). Additionally, there is a substantial body of evidence suggesting that the herpes simplex virus may contribute to the development of dementia (3, 4), along with suggestive evidence that reactivations of the varicella zoster virus may lead to reactivations of the herpes simplex virus in the brain (69). Lastly, it is conceivable that HZ vaccination affects the dementia disease process through a pathogen-independent immunomodulatory pathway, a hypothesis that has been elaborated recently elsewhere (70).

The estimated effect size in our analysis was large in relative terms. However, it is important to recognize two limitations of our data when interpreting this effect size. First, the 95% confidence intervals around our estimates were comparatively wide, meaning that our data were compatible with considerably smaller effect sizes than our point estimates. The width of our confidence intervals may well also be the reason for which we did not observe the same gender effect heterogeneity as we have observed in our analysis of a Welsh dataset (26). Second, there likely was substantial underdiagnosis of dementia in our data. An estimated 8.4% of all Australians over the age of 65 are living with dementia (71), whereas only about 1.4% of patients in the PenCS data in the same age group in 2023 have been diagnosed with dementia. The underdiagnosis of conditions is a well-recognized limitation of working with primary care records from Australia, and not unique to dementia nor the PenCS data (72, 73). Importantly, the degree of underascertainment of dementia is unlikely to differ between birth cohorts born just before versus just after November 2 1936. Assuming that the absolute degree of underdiagnosis of dementia did not differ across the November 2 1936 threshold, then our absolute effect estimates remained unbiased but dementia underascertainment in our data would have led us to (potentially markedly) overestimate the relative effect size for the effect of HZ vaccination on dementia incidence.

Underreporting in our data was also the reason for which we refrained from scaling our effect estimates to the proportion of eligible patients who took up the vaccine. This would have allowed us to estimate the effect of actually receiving (as opposed to merely being eligible for) HZ vaccination. We reasoned that HZ vaccination is likely substantially underreported in our data because uptake of preventive health services in general appeared to be severely underreported. For instance, pneumococcal and annual influenza vaccination coverage among adults aged 65 years and older in Australia is thought to be approximately 55% and 75%, respectively (74). In our data, however, the corresponding percentages in this age group were
only 27% and 33%, respectively. If the degree of underreporting of vaccinations was similar or larger for HZ vaccination as for influenza and pneumococcal vaccination, then any attempt to estimate the effect of receiving (as opposed to merely being eligible for) HZ vaccination using RD will greatly overestimate the effect of HZ vaccination receipt on dementia incidence. We, therefore, chose to only analyze the effect of being eligible for HZ vaccination.

Our study has several additional limitations. First, since the HZ vaccination program only started on November 1, 2016, we were limited in the amount of follow-up time in which we could measure future dementia diagnoses. The maximum follow-up period in this analysis was 7 years and 5 months. Second, our analysis only provided “local” estimates of the effect of HZ vaccination on the incidence of dementia, i.e., estimates for patients who were approximately 79 and 80 years old at the time of the start of the HZ vaccination program. Our main results do not directly speak to the effect of HZ vaccination on dementia incidence for individuals who fall into other age groups. Third, given that we had data from a non-random sample of primary care practices in Australia, our dataset was unlikely to be representative of all primary care patients in the country. Fourth, because the recombinant subunit zoster vaccine (Shingrix [GSK]) was covered by the NIP starting only on November 1, 2016 (and only became available for private purchase in 2021), our effect estimates apply to the live-attenuated HZ vaccine (Zostavax [Merck]) only. Lastly, our CRD analyses relied on the assumption that the main cohort and comparison cohort exhibited the same trend in untreated outcomes across dates of birth. We tested whether this assumption was likely met by measuring whether the difference in slopes between the two cohorts was statistically distinguishable from zero. In addition, we plotted the data points for the main and comparison cohorts to visually assess whether trends were parallel. Our CRD approach passed both tests. However, most importantly, our main findings were unaffected by whether RD (which does not rely on this parallel trends assumption) or CRD was used for the analysis.

References:
4. R. F. Itzhaki, Overwhelming Evidence for a Major Role for Herpes Simplex Virus Type 1 (HSV1) in Alzheimer’s Disease (AD); Underwhelming Evidence against. Vaccines (Basel) 9, 679 (2021).


**Acknowledgments:**
The authors are grateful for the continuous advice and support from the PenCS team who made this data available for research.

**Funding:**
National Institutes of Health/National Institute of Allergy and Infectious Diseases, DP2AI171011 (PG)
National Institutes of Health/National Institute on Aging, R01AG084535 (PG)
Chan Zuckerberg Biohub investigator award (PG)

**Authors’ contributions:**
M.P. devised the methodology, analyzed and processed the data, created data visualizations, interpreted the results, wrote the original draft, and reviewed and edited the original draft. C.B. devised the methodology, interpreted the results, and reviewed and edited the original draft. F.P. and F.M. interpreted the results, and reviewed and edited the original draft. R.P. interpreted the results, and reviewed and edited the original draft. P.G. conceived the overall project, acquired funding, conceived the study, devised the methodology, was responsible for administration and supervision, interpreted the results, wrote the original draft, and reviewed and edited the original draft.

**Competing interests:**
The authors declare no competing interests.

**Data and materials availability:**
The data that support the findings of this study are available from PenCS (32). Researchers must request access to the data directly from PenCS. The authors have no permission to share the data. All codes to define variables are available in the Supplement. All statistical analysis...
code (in R) will be made available in a publicly accessible GitHub repository upon acceptance of
the manuscript for publication.