Long-term survival after invasive pneumococcal disease: a matched cohort study using electronic health records in England

Anne Suffel, Fariyo Abdullahi, Eleanor Barry, Jemma Walker, Nick Andrews, Zahin Amin-Chowdhury, Shamez N Ladhani, Daniel Grint, Helen I McDonald, Ian Douglas, Kathryn E Mansfield, Edward P K Parker**

Summary

Background Invasive pneumococcal disease (IPD) is associated with increased long-term mortality, but it is unclear if this is explained by pre-existing comorbidities. We aimed to estimate the long-term survival following IPD in comparison with the general population, adjusting for potential confounders such as underlying comorbidities.

Methods We conducted a matched cohort study comparing long-term survival (>120 days after infection) in individuals with IPD and comparators without IPD. Cases were individuals aged 65 years or older with laboratory-confirmed IPD (2012–19) identified through enhanced national surveillance. Comparators matched on age, sex, and calendar date of laboratory-confirmed diagnosis were drawn from primary care electronic health records in Clinical Practice Research Datalink GOLD. We used Cox regression, stratified by matched set, to compare mortality in people with and without IPD, adjusting for relevant comorbidities, socioeconomic deprivation, and ethnicity.

Findings We included 13 401 IPD cases and 67 005 comparators without IPD. There were 5038 (53·5%) female and 4380 (46·5%) male IPD cases and 19 927 (53·5%) female and 17 351 (46·5%) male comparators without IPD. After adjusting for comorbidities, socioeconomic deprivation, and ethnicity, we found increased all-cause mortality in IPD cases compared with comparators without IPD (hazard ratio 3·74 [95% CI 3·50–3·99]). The predicted median survival was 4·7 years (IQR 2·9–7·4) for IPD cases and more than 11·9 years (IQR 8·7 to >11·9) for comparators without IPD. This increased mortality was consistent across subgroups defined by age, vaccination history, and comorbidity status (including diabetes, chronic respiratory disease, and chronic heart disease).

Interpretation IPD was associated with increased mortality at least 5 years after infection. These findings emphasise the value of IPD prevention and the need for more research into the clinical management of people who have had IPD. Long-term mortality should be incorporated in cost-effectiveness analyses for pneumococcal vaccines.

Funding National Institute for Health and Care Research (NIHR) Health Protection Research Unit in Vaccines and Immunisation (NIHR200929).

Lancet Healthy Longev 2025; 6: 100775

Published Online November 21, 2025
https://doi.org/10.1016/j.lanhl.2025.100775

See Comment https://doi.org/10.1016/j.lanhl.2025.100785

*Joint senior authors

Department for Infectious Disease Epidemiology and International Health, London School of Hygiene & Tropical Medicine, London, UK (A Suffel PhD, E Barry PhD, D Grint PhD, H I McDonald PhD, E P K Parker PhD); **NIHR Health Protection Research Unit in Immunisation, London School of Hygiene & Tropical Medicine, London, UK** (A Suffel PhD, E Barry PhD, Prof I Douglas PhD, E P K Parker PhD); **UK Health Security Agency, London, UK** (F Abdullahi MSc, J Walker PhD, Prof N Andrews PhD, Z Amin-Chowdhury MSc, Prof S N Ladhani PhD); **Centre for Neonatal and Paediatric Infection, St George's University of London, London, UK** (Prof S N Ladhani PhD); **Department of Life Sciences, University of Bath, Bath, UK** (H I McDonald PhD); **Department of Non-communicable Disease Epidemiology, London School of Hygiene & Tropical Medicine, London, UK** (Prof I Douglas PhD); **School of Health and Care Sciences, University of Lincoln, Lincoln, UK** (Prof K E Mansfield PhD)

Correspondence to: Dr Anne Suffel, Department for Infectious Disease Epidemiology and International Health, London School of Hygiene & Tropical Medicine, WC1E 7HT London, UK
anne.suffel@lshtm.ac.uk

Introduction

Invasive pneumococcal disease (IPD) is a severe systemic infection caused by *Streptococcus pneumoniae*, presenting as pneumonia, meningitis, sepsis, and bacteraemia.[1] Annually, in Europe and the USA, there are an estimated 6–10 cases per 100 000 individuals.[2,3] The 30-day case fatality ratio for IPD ranges from 11% to 30%,[1] with higher risk among infants, older adults, and people with underlying health conditions.[4] IPD is also associated with increased long-term mortality.[5–8] This association might reflect a causal effect of IPD on long-term survival; for example, through progression of cardiovascular or chronic kidney disease.[9,10] Alternatively, IPD could act as an indicator of underlying frailty and comorbidities.[11] The extent to which the poor long-term prognosis after IPD is explained by underlying comorbidities remains unclear, as previous studies have often compared IPD deaths to standardised population mortality rates.[6]

To protect against pneumococcal disease, several, predominantly high-income, countries offer the 23-valent pneumococcal polysaccharide vaccine (PPV23) or pneumococcal conjugate vaccines (PCV13, PCV15, or PCV20) to older adults and individuals with underlying comorbidities. PPV23 is currently recommended for adults aged 65 years or older in the UK.[12] Policy decisions regarding the optimal vaccination strategy rely on cost-effectiveness analyses that typically do not account for long-term mortality.[13–15] Accurate estimates of the effect of IPD on long-term mortality (accounting for the effect of comorbidities) could support improved cost-effectiveness estimates and promote wider implementation and higher uptake of pneumococcal vaccination.

Research in context

Evidence before this study

We searched PubMed from database inception up to Jan 27, 2025, with no restrictions in language with the following search terms: (invasive pneumococcal disease) AND (long-term) AND (mortality). We identified four studies examining mortality from invasive pneumococcal disease (IPD) investigating mortality for a time span between 90 days and 10 years after infection. Previous studies were based on small sample sizes and were often restricted to smaller regions, specific hospitals, or specific health insurances. All studies found an increased long-term mortality in comparison with the general population and some studies showed higher mortality rates for individuals with underlying health conditions; however, it remains unclear to what extent these findings can be explained by pre-existing comorbidities.

Added value of this study

Using a combination of a national surveillance dataset and real-world evidence from population-representative English primary care data, we compared individuals who survived IPD for more than 120 days (n=13 401) to comparators without IPD (n=67 005), accounting for various health conditions and risk factors for increased mortality. We found an overall 3·74-times increased mortality in IPD survivors compared with the comparator group over a median follow-up of 657·5 days (IQR 297–4351) in the comparator group and 1505 days (613–2205) in IPD cases. Our findings of increased mortality were consistent across different subgroups by age, underlying health conditions, and vaccination status. A difference in mortality was particularly pronounced in individuals with no underlying health conditions, younger age groups, and individuals who had not received a pneumococcal vaccine ever or influenza vaccine in the past year.

Implications of all the available evidence

IPD presents a substantial risk for mortality in older adults, even years after the initial infection. Clinical care and surveillance for IPD survivors must be extended beyond the initial treatment of the infection. As IPD is vaccine preventable, long-term mortality should be considered in cost-effectiveness analyses and high vaccination coverage is desirable.

This study aimed to estimate the long-term survival after IPD and compared individuals with IPD with matched comparators in the general population, adjusting for potential confounders, such as underlying comorbidities.

Methods

Study design and population

We conducted a cohort study, comparing IPD cases identified from national surveillance data with individuals without IPD (comparators) using routinely collected primary care electronic records in England. Comparators were matched to IPD cases based on age, sex, and calendar date of laboratory-confirmed diagnosis.

We received data governance approval from Clinical Practice Research Datalink (CPRD; protocol number 19_232) and ethical approval from the London School of Hygiene & Tropical Medicine's research ethics committee (reference number 19198). Informed consent was not required, and patients could choose to opt out if they did not want their confidential patient information to be shared.

The exposure of interest was IPD, defined as laboratory-confirmed invasive pneumococcal isolate in individuals with a clinical diagnosis of pneumonia, meningitis, or any other invasive infection. We used anonymised data from the UK Health Security Agency (UKHSA) on laboratory-confirmed IPD cases between Jan 1, 2012, and Dec 31, 2019. Hospital laboratories routinely report invasive bacterial infections electronically to UKHSA and submit invasive pneumococcal isolates to the UKHSA Respiratory and Vaccine Preventable Bacteria Reference Unit for confirmation and serotyping. As part of the national surveillance, all confirmed cases are enhanced via a postal questionnaire to general practitioners (GPs) requesting clinical information on comorbidities and vaccination history (table 1).[15]

We included IPD cases aged 65–115 years who survived longer than 120 days after reported infection onset (specimen date). A threshold of 120 days was selected to limit inclusion of deaths resulting from complications of the initial infection. Case data were linked to the Personal Demographics Service (PDS) dataset, which includes date of death in primary care records. Index of multiple deprivation (IMD) quintile was determined based on each individual's postcode.[17] Follow-up started 120 days after IPD infection (index date) and ended at the earliest of death or PDS linkage date (April 11, 2024). If IPD cases had more than one eligible IPD episode, we included only the first (interpreting any subsequent episode as a potential mediator on the pathway to increased mortality).

Data on cause of death, pneumococcal serotype, and disease severity (eg, intensive care unit [ICU] admission or length of hospital stay) were not available.

Comparators without IPD were derived from the CPRD GOLD primary care database (June 2024 build).[18] The June 2024 CPRD GOLD release holds information on 21 464 230 research-quality patients;[19] it is broadly representative of the UK population on age, sex, and ethnicity. CPRD GOLD contains information on demographics (ie, sex, year of birth, and region), clinical events (eg, symptoms and diagnoses), primary care prescriptions, vaccination, and lifestyle information (eg, smoking status). We used linked IMD and mortality data from the Office for National Statistics (ONS) where available.[20]

Eligible comparators without IPD included all individuals who were aged at least 65 years during the surveillance period; registered between Jan 1, 2012, and April 29, 2020

	Definition	History assessed (before index date)
Chronic heart disease	Any clinical record of chronic heart disease, including any type of ischaemic heart disease, heart failure, valve insufficiency or stenosis, cardiomyopathy, and major congenital anomalies (such as Fallot's tetralogy)	Any history
Chronic lung disease	Any clinical record of chronic lung disease, including chronic obstructive pulmonary disorder (including chronic bronchitis and emphysema), cystic fibrosis, fibrotic lung disease, respiratory sarcoidosis, occupational lung diseases, or asthma	Any history for chronic lung disease, asthma record only within the past 3 years or aged ≥65 years (definition adapted from Borgan [2018][16])
Chronic liver disease	Any clinical record of chronic liver disease, including liver cirrhosis, chronic hepatitis, biliary atresia or oesophageal varices, or liver transplantation	Any history
Chronic renal disease	Any clinical record indicating stage 3–5 chronic kidney disease or renal replacement therapy	Any history
Asplenia or splenectomy	Any clinical record of asplenia, splenectomy, or dysfunction of the spleen (excluding sickle cell disease)	Any history
Sickle cell disease	Any clinical record of sickle cell disease or homozygous genotype	Any history
Immunosuppression	Any clinical record of a diagnosis indicating immunosuppression or any prescription indicating immunosuppression	Any history of condition indicating long-term immunosuppression; any prescription within 12 months before index date indicating a therapy with immunosuppressive drugs
Malignancy	Any clinical record of a malignancy involving: solid organs; lymphoid, haemopoietic, and related tissues; or skin	Any history
CNS disease	Any clinical record of a chronic disease involving the CNS, including dementia, brain tumours, haemorrhagic or thrombotic infarction, poliomyelitis, and paralysis	Any history
Cochlear implantation	Any clinical record of a cochlear transplant or related surgery	Any history
Diabetes	Any clinical record of diabetes or diabetes-related disease but excluding gestational diabetes and diabetes insipidus	Any history
Coeliac disease	Any clinical record of coeliac disease	Any history
History of IPD	Any clinical record of meningitis, pericarditis, peritonitis, pleurisy, pneumonia, septic arthritis, or bacteraemia explicitly linked to *Streptococcus pneumoniae*	Any history
Influenza vaccination	Most recent clinical record or prescription of an influenza vaccine	Most recent record
PPV23 vaccination	Most recent clinical record or prescription of PPV23 or unspecified polysaccharide pneumococcal vaccine	Most recent record

Definitions aimed to emulate the comorbidities captured for IPD cases via a postal questionnaire sent to GPs as part of the national IPD surveillance programme. A clinical record refers to a morbidity coded in the GP software (ie, no free text was used). GP=general practitioner. IPD=invasive pneumococcal disease. PPV23=23-valent pneumococcal polysaccharide vaccine.

Table 1: **Definition of comorbidities in the comparator population without IPD**

(corresponding to 120 days after Dec 31, 2019—ie, the latest index date of IPD cases); had at least 1 year of registration; and met CPRD quality-control standards at the individual and practice level.[21] Comparators entered the cohort on the index date of their matched IPD case (120 days after infection onset). We used exact matching, on year of birth and sex, to select five comparators, without replacement, for each confirmed IPD case in order of increasing index date. Follow-up ended at the earliest of end of registration with GP practice, death, at the last data collection from practice, at the end of study period (April 11, 2024, aligning with the PDS linkage date of IPD cases).

We did not exclude comparators with previous IPD as the condition is likely to be poorly recorded in primary care, but performed a post-hoc sensitivity analysis in which we excluded comparators with a previous IPD record in primary care (appendix p 1).

Outcome

The primary outcome was date of death recorded in primary care records. As linkage to the ONS mortality dataset is considered the gold standard for assessing mortality but was not available for IPD cases,[22] we assessed potential outcome misclassification by comparing the date of death in ONS versus primary care records among CPRD comparators with ONS mortality linkage.

Covariates

For IPD cases, demographic information (ie, year of birth, sex, region of residence, ethnicity [White, Black, South Asian, mixed, other], and IMD quintile) were included in the UKHSA surveillance dataset. Data on the following comorbidities reflecting clinical risk at the time of infection (recorded as Yes, No, Unknown, or blank) were ascertained from the GP questionnaire: chronic heart disease, chronic lung disease, chronic liver disease, chronic renal disease, asplenia or splenectomy, sickle cell disease, immunosuppression or immunosuppressive drug, malignancy, CNS disease, cochlear implantation, diabetes, and coeliac disease.[12] PPV23 and influenza vaccination status (and most recent corresponding vaccination dates) were also included.

See Online for appendix

For comparators, we obtained demographic information (ie, year of birth, sex, and region) from primary care records. Region was included to account for geographical differences in risk of infection, access to health care, and surveillance. Ethnicity was defined using a validated algorithm.[23] We assessed the same comorbidities obtained for IPD cases using relevant clinical (morbidity coded) and prescription records (table 1). When a comorbidity was not recorded, we assumed the absence of the comorbidity. Vaccination status was defined using prescription and clinical records.[24] Aside from age and sex, baseline covariates for comparators were defined after matching.

Statistical analysis

Short-term survival of all IPD cases was initially described at 30 days, 60 days, 90 days, and 120 days after infection. We then used Kaplan–Meier curves to estimate long-term survival (from 120 days onwards) for IPD cases and matched comparators without IPD. We used Cox regression models (stratified by matched set) to estimate hazard ratios (HRs) and 95% CIs comparing overall and period-specific (0–1 year, 0–2 year, and 0–5 year) mortality in IPD cases (from 120 days post IPD) versus comparators without IPD. We estimated a sparse model implicitly accounting for matching variables (ie, age, sex, and calendar date of laboratory-confirmed diagnosis; Model 1), then sequentially adjusted for other covariates as follows: region and IMD quintile (Model 2); also adjusting for comorbidities (Model 3; primary analysis); and adjusting for ethnicity (Model 4). All matching variables were considered through stratifying by matched set. If underlying comorbidities or demographic factors represent common risk factors for both IPD and long-term mortality (ie, as confounders), the effect size for the association between IPD and mortality would be attenuated after their inclusion in the model. Model 3 was deemed the primary analysis model given our emphasis on whether long-term mortality after IPD might be explained by underlying comorbidities.

Based on Model 3, which was implemented with adjustment for matching variables to enable prediction without stratification index, we estimated survival for both IPD cases and comparators using the Nelson–Aalen estimator.[16,25] We calculated the median survival time[26] for each individual and the median (IQR) for the predicted survival of the population based on the covariate distribution of the IPD cohort, the comparator cohort, and a standardised IPD cohort (with the same covariate distributions as the comparators). Where the predicted survival probability for an individual was >0·5, the maximum follow-up time was assigned. All analyses were performed in R (version 4.2.2). Code lists and statistical analysis code are available on GitHub.[26]

We included all eligible IPD cases and comparators in Model 1. Each successive model was restricted to individuals remaining in valid matched sets (cases with at least one comparator after exclusion of individuals with missing data). Individuals with missing region, IMD, or ethnicity were excluded from models adjusting for these. When adjusting for comorbidities (ie, in Models 3–4), we included all cases with at least one comorbidity recorded as Yes or No (indicating at least some engagement with the surveillance questionnaire). When other comorbidities were recorded as Unknown or blank, we interpreted this as absence of the comorbidity. For comparators, absence of a medical record for a comorbidity was interpreted as absence of the condition.

We conducted secondary subgroup analyses defined by age group (65–69 years, 70–74 years, 75–79 years, 80–89 years, and ≥90 years); comorbidity status (none stated or unknown, cardiovascular disease, respiratory disease, kidney disease, neurological disease, malignancy, and diabetes); vaccination status (influenza and pneumococcal vaccination history); and clinical manifestation of IPD (meningitis, peritonitis, pericarditis, bacteraemia, pneumonia, arthritis, and sepsis). A series of sensitivity analyses were also performed on Model 3 (appendix p 1).

We calculated hazard ratios (HRs) and 95% CIs to assess differences between the groups.

Role of the funding source

The funder of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report.

Results

13 401 IPD cases and 67 005 matched comparators were included in the study (figure 1). Of these, 9418 cases and 37 278 comparators were eligible for the comorbidity-adjusted analyses (ie, the primary analysis cohort; Model 3). Median follow-up in the primary analysis cohort was 1505·0 days (IQR 613·0–2205·0) for IPD cases and 657·5 days (297·0–4351·0) for comparators. 5038 (53·5%) of IPD cases were female and 4380 (46·5%) were male; 19 927 (53·5%) of comparators were female and 17 351 (46·5%) were male (table 2). IPD cases and comparators differed in region of residency, ethnicity, socioeconomic deprivation, and prevalence of underlying health conditions. Recurrence of IPD during follow-up was observed in 355 (2·6%) of 13 401 cases.

5275 (26·0%) of 20 285 IPD cases aged 65 years and older had died within 30 days of infection and 6884 (33·9%) died within 120 days (appendix p 2). Among IPD cases who survived past 120 days, mortality was higher than matched comparators without IPD throughout follow-up (figure 2). Unadjusted Kaplan–Meier estimates of 1-year survival probability were 83·1% (95% CI 82·3–83·8) in IPD cases and 95·4% (95·2–95·6) in comparators without IPD. Unadjusted estimates of 5-year survival were 45·1% (95% CI 44·1–46·2) in cases and 79·5% (78·9–80·2) in comparators.

The HR for all-cause mortality in IPD cases compared with comparators was 4·11 (95% CI 3·95–4·27) in Model 1, which accounted only for matching variables. In Model 2, which adjusted for region and IMD, the HR was 4·14 (95% CI 3·95–4·35). In Model 3, which further adjusted for pre-existing comorbidities, the HR was 3·74 (95% CI 3·50–3·99), and in Model 4, which also adjusted for

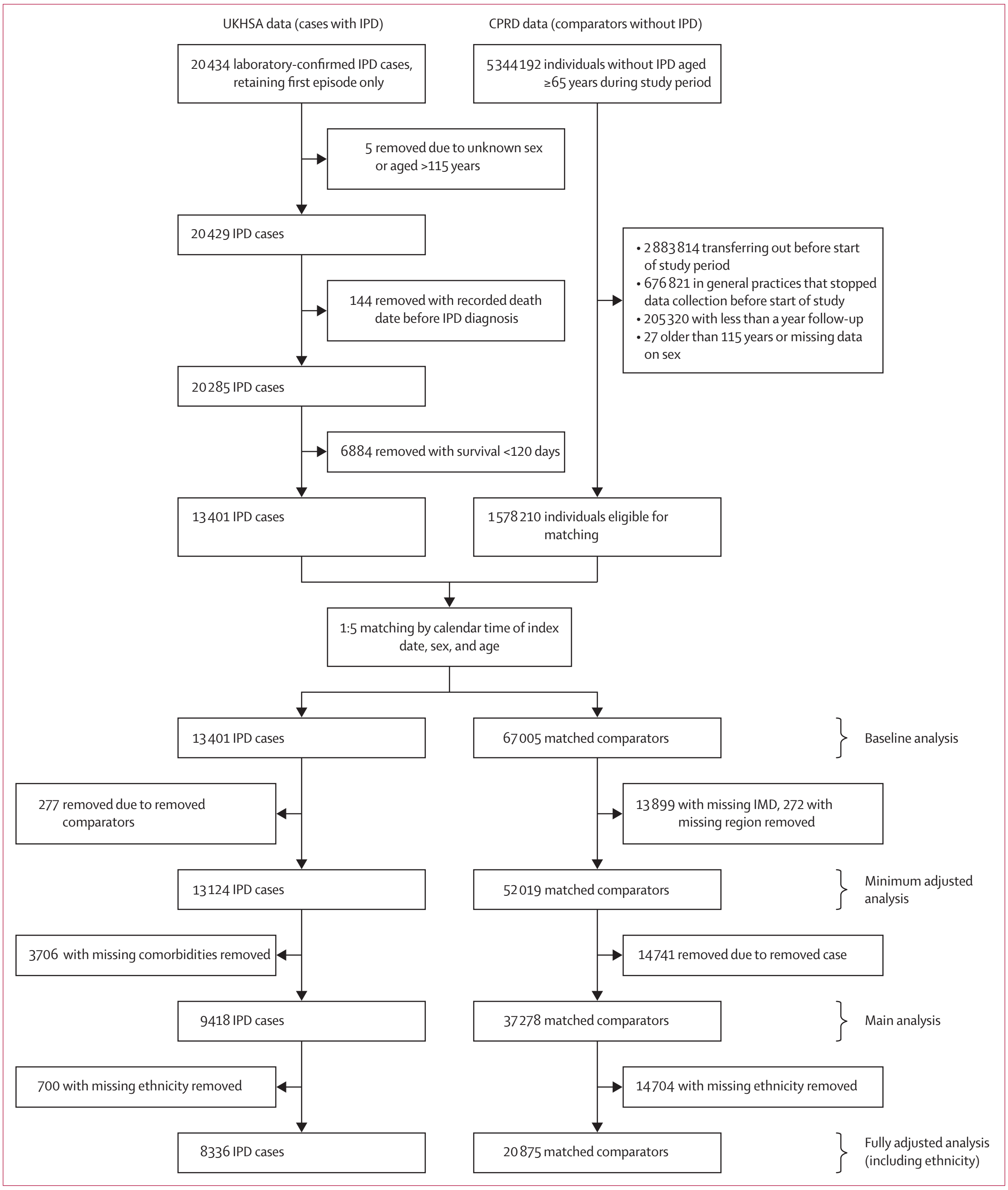

***Figure 1:* Selection of matched study population**

Cases were included if at least one comorbidity was recorded as Yes or No, indicating engagement with the surveillance questionnaire. CPRD=Clinical Practice Research Datalink. IPD=invasive pneumococcal disease. UKHSA=UK Health Security Agency.

ethnicity, the HR was 3·84 (3·52–4·18; appendix [pp 3–7]). Period-specific HRs indicated a slightly higher year-1 excess mortality compared with later periods, although a strong effect of IPD persisted throughout follow-up (figure 2; appendix p 10).

The median predicted survival based on Model 3 was more than 11·9 years (IQR 8·7 to >11·9) for comparators without IPD, 4·7 years (2·9–7·4) for cases with IPD, and 5·3 years (3·0–8·7) for cases with IPD after demographic and comorbidity standardisation (figure 3; appendix p 11).

	Cases (n=9418)	Comparators (n=37 278)
Age (years)		
65–69	1723 (18·3%)	6838 (18·3%)
70–74	1979 (21·0%)	7893 (21·2%)
75–79	1848 (19·6%)	7208 (19·3%)
80–89	2883 (30·6%)	11 477 (30·8%)
90+	985 (10·5%)	3862 (10·4%)
Year of infection (reporting date)		
2012	467 (5·0%)	..
2013	673 (7·1%)	..
2014	964 (10·2%)	..
2015	1202 (12·8%)	..
2016	1559 (16·6%)	..
2017	1507 (16·0%)	..
2018	1648 (17·5%)	..
2019	1398 (14·8%)	..
Sex*		
Female	5038 (53·5%)	19 927 (53·5%)
Male	4380 (46·5%)	17 351 (46·5%)
Region		
East Midlands	928 (9·9%)	63 (0·2%)
East of England	885 (9·4%)	1826 (4·9%)
London	834 (8·9%)	4388 (11·8%)
North East	478 (5·1%)	299 (0·8%)
North West	1398 (14·8%)	7506 (20·1%)
South East	1525 (16·2%)	14 041 (37·7%)
South West	1253 (13·3%)	3630 (9·7%)
West Midlands	978 (10·4%)	4932 (13·2%)
Yorkshire and the Humber	1139 (12·1%)	593 (1·6%)
Ethnicity		
Asian	285 (3·0%)	537 (1·4%)
Black	128 (1·4%)	295 (0·8%)
Mixed	111 (1·2%)	79 (0·2%)
Other	61 (0·6%)	178 (0·5%)
White	8133 (86·4%)	21 485 (57·6%)
Missing	700 (7·4%)	14 704 (39·4%)
Index of multiple deprivation (quintiles)		
1 (most deprived)	1998 (21·2%)	10 189 (27·3%)
2	1805 (19·2%)	8569 (23·0%)
3	1912 (20·3%)	7584 (20·3%)
4	1922 (20·4%)	6486 (17·4%)
5 (least deprived)	1781 (18·9%)	4450 (11·9%)
Asplenia or splenectomy	43 (0·5%)	76 (0·2%)
Sickle cell disease	10 (0·1%)	<5
Coeliac disease	44 (0·5%)	149 (0·4%)
Malignancy	2062 (21·9%)	5464 (14·7%)
Diabetes	2201 (23·4%)	6426 (17·2%)
Cochlear implantation	15 (0·2%)	<5
Chronic respiratory disease	3456 (36·7%)	6817 (18·3%)
Chronic liver disease	272 (2·9%)	174 (0·5%)
Chronic kidney disease	2123 (22·5%)	8589 (23·0%)
Chronic cardiovascular disease	3180 (33·8%)	8634 (23·2%)
CNS disease	262 (2·8%)	5880 (15·8%)

(Table 2 continues in next column)

	Cases (n=9418)	Comparators (n=37 278)
(Continued from previous column)		
Immunosuppression	926 (9·8%)	898 (2·4%)
Influenza vaccine last year	6198 (65·8%)	26 705 (71·6%)
PPV23 vaccine ever	6590 (70·0%)	27 995 (75·1%)

Data are n (%). Cell counts with less than 5 were suppressed for anonymity. IPD=invasive pneumococcal disease. PPV23=23-valent pneumococcal polysaccharide vaccine. *Variable named "gender" in both case and comparator datasets but was interpreted as sex.

***Table 2:* Baseline characteristics of IPD cases who survived to 120 days after infection and their matched comparators**

The higher mortality rate in cases with IPD compared with comparators without IPD was consistent across subgroups defined by age, comorbidity status, and vaccination history (figure 4; appendix pp 14–15). The lowest HR was observed for individuals aged 80–89 years (3·41 [95% CI 3·16–3·68]) and 90 years or older (3·00 [2·68–3·34]). The corresponding discrepancies in median predicted survival were greater in younger age groups compared with older age groups (figure 3).

The HR for IPD cases compared with comparators without IPD that had no comorbidities was higher (5·16 [95% CI 4·17–6·39]) than that for subgroups with diabetes, cardiovascular, respiratory disease, and other common comorbidities (figure 4). The subgroup of individuals with no comorbidities was younger than subgroups with comorbidities (figure 4). There was also a lower HR in individuals with a history of PPV23 vaccine or influenza vaccination in the past year compared with unvaccinated subgroups, although the unvaccinated subgroups had a notably lower median age.

The association between IPD and long-term mortality differed by clinical presentation of IPD and was highest for cases presenting with peritonitis (HR 8·97 [95% CI 3·07–26·23]) and sepsis (4·55 [3·71–5·58]), and lowest for cases presenting with arthritis (2·04 [1·43–2·92]) and pericarditis (1·83 [0·50–6·74]; figure 4; appendix pp 12–13).

Results from the following sensitivity analyses were consistent with those of the main analyses: (1) limiting to IPD cases where comorbidities were recorded as Yes or No (HR 3·70 [95% CI 3·45–3·97]); (2) excluding comparators with a previous primary care record of IPD (3·76 [3·52–4·02]); (3) limiting follow-up to Dec 31, 2019 (pre-COVID-19; 3·74 [3·46–4·04]); and (4) applying different methods to account for matching (appendix pp 14–15).

We observed no indication of selection bias associated with exclusions due to missing IMD, region, or ethnicity (appendix pp 14–15); however, the HR was 4·14 (95% CI 3·95–4·35) before exclusion of cases in which all comorbidities were recorded as Unknown or blank, compared with 3·67 (3·46–3·89) after limiting to cases in which at least one comorbidity was recorded as Yes or No.

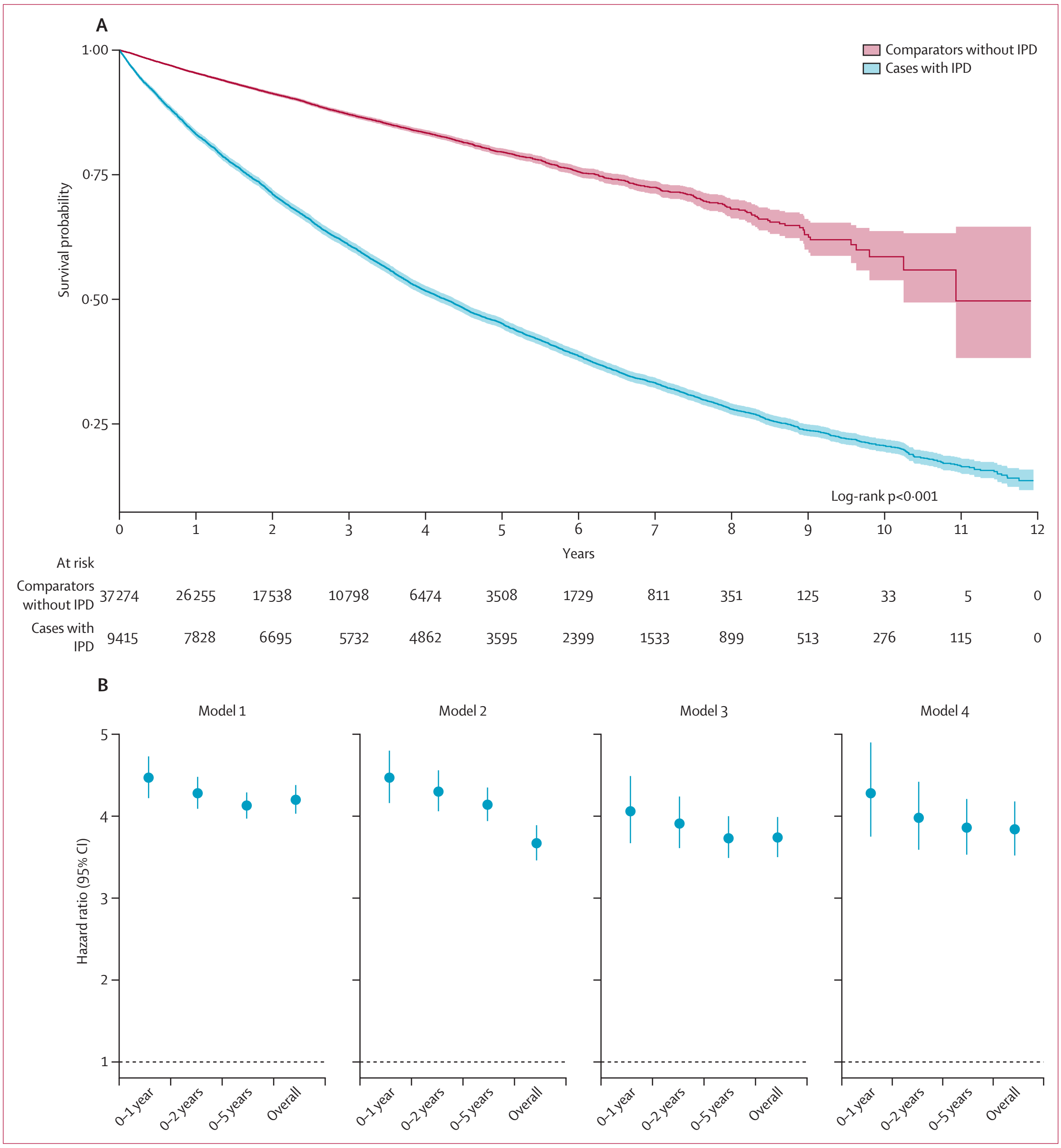

Figure 2: **Association between IPD and long-term survival**
(A) Kaplan–Meier curve comparing the survival of cases with IPD and comparators without IPD matched by age and sex. For IPD cases, day 0 (index date) corresponds to 120 days after the recorded date of infection. Kaplan–Meier steps were delayed until five events had occurred as a precaution against small number disclosures. (B) Hazard ratios (95% CI), estimated using Cox regression stratified by matched set, comparing mortality in IPD cases to comparators without IPD with sequential adjustment for covariates: Model 1: stratified set only, Model 2: additional adjustment for region and IMD; Model 3: additional adjustment for comorbidities; and Model 4: additional adjustment for ethnicity (region, IMD, comorbidities, and ethnicity). Full model outputs (appendix pp 8–9), Kaplan–Meier curves for all cohort definitions (appendix pp 16–17), and age-specific Kaplan–Meier curves (appendix pp 18–19) are shown elsewhere. HR=hazard ratio. IMD=index of multiple deprivation. IPD=invasive pneumococcal disease.

Combining CPRD data with the linked ONS mortality dataset detected only 215 (2·92%) additional deaths in comparison with using primary care data alone (appendix p 15).

Discussion

Among IPD cases aged 65 years or older who lived for 120 days after infection, this study documented substantially higher mortality relative to matched comparators

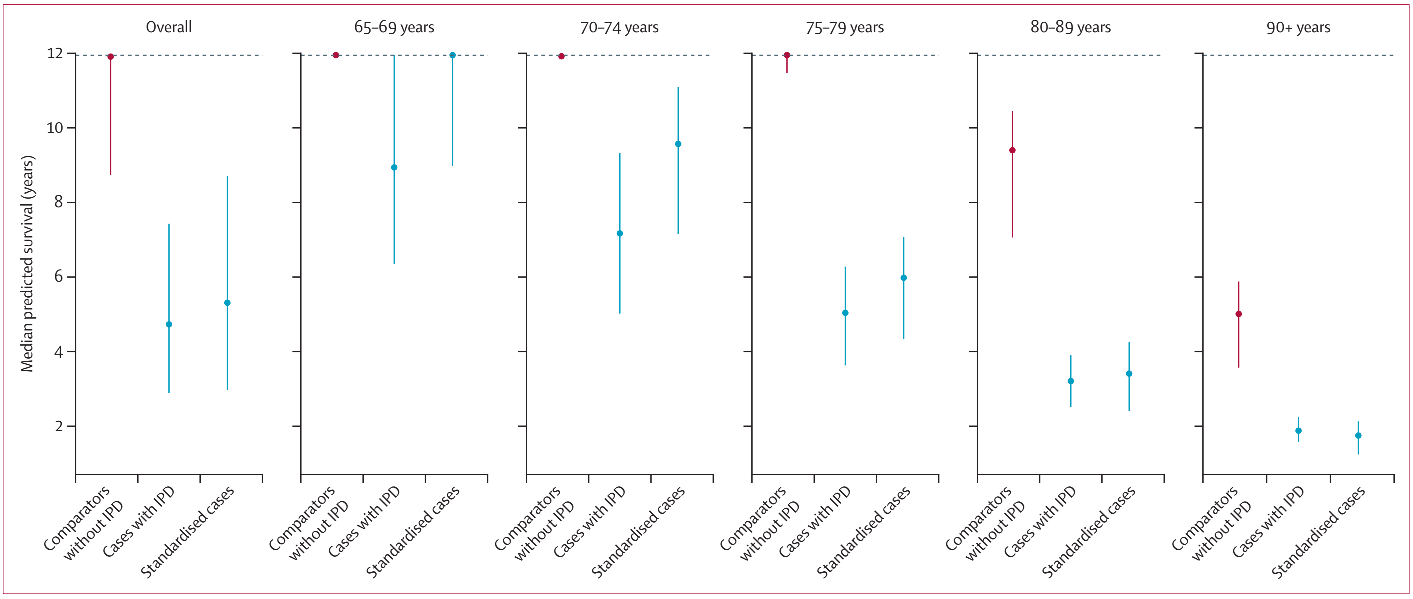

Figure 3: **Association between IPD and predicted survival**

Predicted median survival (IQR) for the overall cohort and stratified by age group. Survival was predicted using the Aalen–Johnson estimator based on models adjusting for region, index of multiple deprivation, underlying comorbidities, and matching variables without stratification. The dashed line indicates the maximum observed follow-up (ie, 12 years). For cases with IPD, we obtained estimates based on the covariate distribution of the IPD cases and a standardised cohort with the same covariate distributions as the comparator population. Full data are presented in the appendix (p 11). IPV=invasive pneumococcal disease.

without IPD over a follow-up period of up to 12 years. This increased mortality was consistent after adjusting for underlying comorbidities, ethnicity, deprivation, and region, and across subgroups defined by age, comorbidities, and vaccination status. Together, our findings suggest that the increased risk of long-term mortality after IPD cannot be explained by the confounding effect of underlying comorbidities.

Our findings are consistent with previous reports of increased long-term mortality after severe pneumococcal infections. A study of US veterans, with 392 cases, indicated increased mortality for up to 10 years following pneumococcal pneumonia compared with the general population, with particularly high mortality following severe infection.[5] A Norwegian cohort study, with 372 cases, found increased mortality 10 years after hospital admission for IPD, with higher mortality in people with underlying health conditions.[6] A Dutch cohort study, with 228 cases, reported increased mortality up to 3 years after infection for both invasive and non-invasive pneumococcal disease. In contrast to our study, the excess mortality was not apparent among individuals without underlying comorbidities.[7] Notably, previous studies have either looked at IPD cases alone or relative to standardised population mortality rates, making it difficult to assess the potentially confounding effect of underlying health conditions.

The findings of this study support a potential effect of IPD on the long-term risk of mortality. This finding is consistent with a growing literature highlighting the role of severe infections in advancing trajectories of frailty, including reduced quality of life, functional decline, cognitive impairment, and increased mortality.[27,28] Excess mortality was higher in individuals with peritonitis and sepsis compared with other IPD presentations, which could be linked to differences in the host or serotype.[29] Notably, the precision of subgroup-specific estimates was low due to small numbers. We did not have data for the cause of death in IPD cases, which could have offered insight into the causal pathway towards increased mortality. However, previous studies have highlighted an increased risk of cardiovascular disease following infection,[9,30] potentially linked with persistence of an inflammatory and prothrombotic state.[31] Pneumococcal pneumonia has also been linked with increased renal complications including end-stage renal disease.[32] Future studies assessing long-term trends in cause-specific mortality after IPD would be a valuable extension to this work.

The excess mortality among IPD cases was apparent across all subgroups considered. We found a modestly lower effect size in older age groups compared with younger age groups. This finding likely reflects the accrual of competing risks of mortality with increasing age, reducing the excess mortality associated with IPD. We also found the excess mortality to be attenuated in subgroups with PPV23 or recent influenza vaccination. This finding could reflect a protective effect of vaccination on outcomes after IPD (eg, via a direct effect of PPV23 or a reduction in influenza infections during or after IPD). However, the differences should be interpreted with caution given the older age of vaccinated compared with unvaccinated individuals and the potential for confounding by indication (eg, higher vaccine uptake in individuals with underlying health conditions that present a

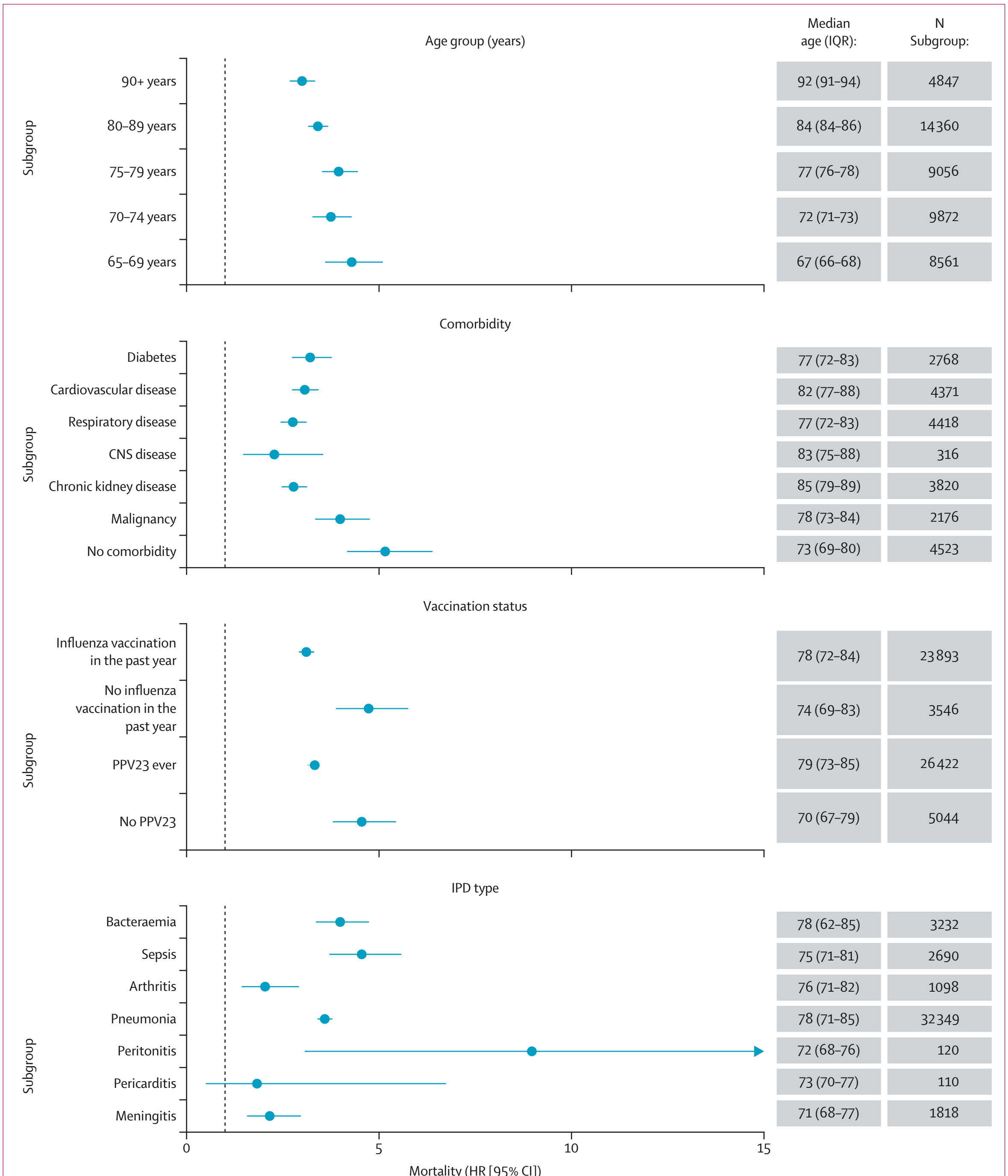

Figure 4: **Association between IPD and long-term survival by subgroup**
HR (95% CI) by age group, underlying comorbidities, and vaccination status. Age of the subpopulations is presented as median (IQR) age in years . Full data are presented in the appendix (pp 12–13). HR=hazard ratio. IPD=invasive pneumococcal disease. PPV23=23-valent pneumococcal polysaccharide vaccine.

competing mortality risk).[33] We did not explore vaccination history as a covariate in adjusted models as IPD is a mediator on the causal pathway between vaccination and mortality, undermining the interpretation of HRs related to vaccination status.

Across comorbidities, the effect of IPD was lower in individuals with CNS disease and higher in individuals with malignancy, partially overlapping with findings from a Norwegian study in which cardiovascular disease and malignancy were associated with increased mortality among IPD cases.[6] The larger effect size among individuals with no comorbidities might be explained in part by the younger age of this subgroup, although it could also reflect differential ascertainment of comorbidity status in IPD cases (based on GP questionnaire) versus comparators without IPD (based on primary care records). Specifically, if comorbidity ascertainment is more sensitive based on primary care coding than GP questionnaire, IPD cases with comorbidities would be more likely than comparators to be misclassified as having no underlying health conditions, resulting in a bias away from the null in the subgroup analysis that was restricted to individuals with no comorbidities. Alternatively, the larger effect size might reflect a distinct inflammatory response among otherwise healthy individuals who develop IPD, with the cytotoxic[34,35] and physiological stress from infection causing organ damage that results in new morbidities and a long-term excess mortality risk.[36] Additionally, some individuals with undiagnosed health conditions or unmeasured frailty might be more susceptible to IPD but are not receiving any clinical care for their undiagnosed health conditions. These untreated conditions might lead to higher mortality in combination with the IPD infections.

Previous cost-effectiveness analyses related to pneumococcal vaccines have focused on short-term mortality.[13–15] Our findings support the inclusion of long-term mortality in cost-effectiveness models and provide relative and absolute effect sizes that could be used to parameterise future evaluations, including those evaluations to new high-valency formulations that incorporate additional serotypes as well as eligibility for booster doses among subgroups prone to rapid waning of immunity. Notably, 70% of IPD cases and 75% of comparators without IPD in this study had been vaccinated with PPV23, despite all individuals in England being eligible for vaccination from the age of 65 years. By highlighting the substantial mortality attributed to IPD, our findings emphasise the long-term harms of severe pneumococcal infection, supporting efforts to promote vaccine uptake. Our study also highlights the need for proactive clinical management of individuals following recovery from IPD to mitigate their higher risk of long-term mortality.

Strengths of our study include the use of nationally representative datasets for IPD cases and comparators without IPD. The substantial size of our cohort allowed us to adjust for rare comorbidities. Recruitment from 2012–19 enabled us to estimate effects for up to a 12-year follow-up period. By using a matched cohort design, it was possible to compare mortality for individuals of the same age, sex, and calendar period, while the detailed investigation of cases via the national IPD surveillance programme enabled us to adjust for relevant comorbidities that represent potential confounders for the association between IPD and long-term mortality. Our findings were robust to different modelling strategies and to varying interpretations of missing data in the GP questionnaire.

A key limitation of this study is the use of different data sources used for the IPD cases and comparators, and the differential ascertainment of comorbidities resulting from this. Given that the comparators could be censored (eg, by leaving a GP practice) and consequently had a shorter follow-up time than IPD cases, their mortality might have been underestimated, biasing estimates away from the null. GP questionnaires included broad disease categories (eg, respiratory disease). Interpretation of categories might have differed across GPs and, although it is likely that GPs consulted patient health records when filling out the questionnaires, this process might not reflect the definitions and look-back periods we applied in comparators without IPD. Nonetheless, our findings were consistent in analyses restricted to subgroups that are likely to be well defined in both the case and comparator data (eg, diabetes, for which primary care coding is incentivised by the Quality and Outcomes framework).[37] Other factors such as alcohol use disorder, smoking status, and receipt of some medications (eg, immunosuppressants) could also predispose individuals to IPD and higher mortality, but these could not be accounted for in this study.[38] Therefore, we cannot exclude the contribution of residual confounding as an explanation for higher mortality observed in IPD cases, although it is unlikely to fully account for the strong effect size observed across models and subgroups. In addition, our study did not consider the effect of the severity of the initial IPD episode (eg, general ward *vs* ICU admissions) or pneumococcal serotype (eg, vaccine *vs* non-vaccine), which has previously been shown to influence mortality after IPD[4] and is also a key consideration for vaccine policy decisions. These would represent valuable extensions to the present study. Finally, the predicted median survival among comparators exceeded the maximum follow-up time, limiting precise comparison with predicted survival estimates for cases.

In conclusion, IPD affects health far beyond the initial episode, highlighting the importance of close surveillance and long-term support of individuals following IPD. More research is needed to optimise strategies for long-term monitoring and treatment of IPD survivors. Our findings emphasise the importance of efforts to prevent IPD through encouraging higher vaccination uptake and highlight the relevance of long-term mortality during cost-effectiveness analyses and policy decisions related to pneumococcal vaccination.

Contributors

EPKP, JW, NA, HIMcD, FA, ZA-C, SNL, DG, and KEM conceptualised the study and contributed to the study design. FA processed the IPD case data set at the UKHSA. AS and KEM developed the code lists or adapted existing code lists from the Electronic Health Records group at the London School of Hygiene & Tropical Medicine. AS cleaned, analysed, and interpreted the data with input from EPKP, NA, ID, JW, EB, and KEM. AS and EPKP had full access to the data and verified the confidential data on a secure server. AMS wrote the first draft of the manuscript. The manuscript was critically revised by all authors. All authors had final responsibility for the decision to submit for publication.

Declaration of interests
AS is also funded by OpenPhilanthropy which had no involvement in this manuscript. While finalising this manuscript, EB began a role that is partially funded by educational grants from external companies, including Pfizer and Takeda, to support bespoke training in pharmacoepidemiology and real-world evidence. These funds and companies had no involvement in the manuscript. ID holds other research grants from GSK and AstraZeneca, and holds GSK shares. These companies had no involvement in the manuscript. The Immunisation and Vaccine Preventable Diseases Division at UKHSA has provided vaccine manufacturers with post-marketing surveillance reports on meningococcal and pneumococcal infections, which the companies are required to submit to the UK Licensing Authority in compliance with their risk management strategy. A cost recovery charge is made for these reports. EPKP has received funding from UK Research and Innovation/Medical Research Council, NIHR, and the Bill & Melinda Gates Foundation. EPKP has also received funding as a consultant for Tulane University as part of the Safe in Pregnancy and Childhood Study, supported by the Safety Platform for Emergency Vaccines, Task Force for Global Health, and The Coalition for Epidemic Preparedness Innovations. EPKP received reimbursements for providing input on an expert report for the UK COVID-19 Inquiry. These funds and institutions had no involvement in the manuscript. All other authors declare no competing interests.

Data sharing
The study uses data from the CPRD. CPRD does not allow the sharing of patient-level data. The data specification for the CPRD data set is available at: https://www.cprd.com/doi/cprd-gold-june-2024-dataset. Analysis code and code lists are shared in our online repository under the following link: https://github.com/ehr-lshtm/IPD_mortality/

Acknowledgments
This study uses data provided by patients and collected by the UK National Health Service as part of their care and support (usemydata.org). This study is based in part on data from the CPRD obtained under licence from the UK Medicines & Healthcare products Regulatory Agency. The interpretation and conclusions contained in this study are those of the authors alone. This study is funded by the NIHR Health Protection Research Unit in Vaccines and Immunisation (NIHR200929), a partnership between UKHSA and the London School of Hygiene & Tropical Medicine. The views expressed are those of the author(s) and not necessarily those of the NIHR, UKHSA or the Department of Health and Social Care. The funder of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report.

References
1 Drijkoningen JJC, Rohde GGU. Pneumococcal infection in adults: burden of disease. *Clin Microbiol Infect* 2014; **20** (suppl 5)**:** 45–51.
2 Center for Disease Control. Chapter 11: Pneumococcal. 2020. https://www.cdc.gov/surv-manual/php/table-of-contents/chapter-11-pneumococcal.html?CDC_AAref_Val=https://www.cdc.gov/vaccines/pubs/surv-manual/chpt11-pneumo.html (accessed Oct 13, 2025).
3 European Centre for Disease Prevention and Control. Invasive pneumococcal disease - Annual Epidemiological Report for 2018. 2018. https://www.ecdc.europa.eu/en/publications-data/invasive-pneumococcal-disease-annual-epidemiological-report-2018 (accessed Oct 13, 2025).
4 Weinberger DM, Harboe ZB, Sanders EAM, et al. Association of serotype with risk of death due to pneumococcal pneumonia: a meta-analysis. *Clin Infect Dis* 2010; **51:** 692–99.
5 Sandvall B, Rueda AM, Musher DM. Long-term survival following pneumococcal pneumonia. *Clin Infect Dis* 2013; **56:** 1145–46.
6 Floeystad HK, Holm A, Sandvik L, Vestrheim DF, Brandsaeter B, Berild D. Increased long-term mortality after survival of invasive pneumococcal disease: a population-based study. *Infect Dis (Lond)* 2017; **49:** 365–72.
7 Wagenvoort GHJ, Sanders EAM, de Melker HE, van der Ende A, Vlaminckx BJ, Knol MJ. Long-term mortality after IPD and bacteremic versus non-bacteremic pneumococcal pneumonia. *Vaccine* 2017; **35:** 1749–57.
8 Versluys KA, Eurich DT, Marrie TJ, Tyrrell GJ. Invasive pneumococcal disease and long-term mortality rates in adults, Alberta, Canada. *Emerg Infect Dis* 2022; **28:** 1615–23.
9 Corrales-Medina VF, Alvarez KN, Weissfeld LA, et al. Association between hospitalization for pneumonia and subsequent risk of cardiovascular disease. *JAMA* 2015; **313:** 264–74.
10 Restrepo MI, Reyes LF. Pneumonia as a cardiovascular disease. *Respirology* 2018; **23:** 250–59.
11 Grijalva CG. Is pneumonia a risk factor or a risk marker for long-term mortality? *Am J Respir Crit Care Med* 2015; **192:** 532–34.
12 UK Health Security Agency. Green Book–Chapter 25: Pneumococcal. 2023. https://assets.publishing.service.gov.uk/media/684b1fbc1c8d5c94e201abae/Green_Book__Chapter_25_pneumococcal_12_6_25.pdf (accessed Oct 13, 2025).
13 Danelian G, Burton L, Bayley T, et al. The impact and cost-effectiveness of pneumococcal immunisation strategies for the elderly in England. *Vaccine* 2024; **42:** 3838–50.
14 Mendes D, Averin A, Atwood M, et al. Cost-effectiveness of using a 20-valent pneumococcal conjugate vaccine to directly protect adults in England at elevated risk of pneumococcal disease. *Expert Rev Pharmacoecon Outcomes Res* 2022; **22:** 1285–95.
15 de Boer PT, van Werkhoven CH, van Hoek AJ, et al. Higher-valency pneumococcal conjugate vaccines in older adults, taking into account indirect effects from childhood vaccination: a cost-effectiveness study for the Netherlands. *BMC Med* 2024; **22:** 69.
16 Borgan Ø. Aalen–Johansen Estimator. In: Wiley StatsRef: Statistics Reference Online. Wiley, 2018: 1–13.
17 Ministry of Housing Communities and Local Government. English indices of deprivation 2019: technical report. 2019. https://www.gov.uk/government/publications/english-indices-of-deprivation-2019-technical-report (accessed Oct 13, 2025).
18 Herrett E, Gallagher AM, Bhaskaran K, et al. Data resource profile: Clinical Practice Research Datalink (CPRD). *Int J Epidemiol* 2015; **44:** 827–36.
19 Clinical Practice Research Datalink. CPRD GOLD June 2024 (Version 2024.06.001) [Data set]. 2024. https://www.cprd.com/doi/cprd-gold-june-2024-dataset (accessed Oct 13, 2025).
20 Medicines & Healthcare products Regulatory Agency (MHRA). CPRD linked data. 2024. https://www.cprd.com/cprd-linked-data (accessed Oct 13, 2025).
21 Lewis JD, Bilker WB, Weinstein RB, Strom BL. The relationship between time since registration and measured incidence rates in the General Practice Research Database. *Pharmacoepidemiol Drug Saf* 2005; **14:** 443–51.
22 Gallagher AM, Dedman D, Padmanabhan S, Leufkens HGM, de Vries F. The accuracy of date of death recording in the Clinical Practice Research Datalink GOLD database in England compared with the Office for National Statistics death registrations. *Pharmacoepidemiol Drug Saf* 2019; **28:** 563–69.
23 Mathur R, Bhaskaran K, Chaturvedi N, et al. Completeness and usability of ethnicity data in UK-based primary care and hospital databases. *J Public Health (Oxf)* 2014; **36:** 684–92.
24 Suffel AM, Walker JL, Campbell CNJ, Carreira H, Warren-Gash C, McDonald HI. Methodological challenges and recommendations for identifying childhood immunisations using routine electronic health records in the United Kingdom. *medRxiv* 2023; published online March 1. https://doi.org/10.1101/2023.02.28.23286573.
25 Kalbfleisch JD, Prentice RL. The statistical analysis of failure time data. Wiley, 2002.
26 Suffel AM. GitHub repository: Long-term mortality from invasive pneumoccoal disease. 2025. https://github.com/ehr-lshtm/IPD_mortality/ (accessed Oct 13, 2025).
27 Iwashyna TJ, Ely EW, Smith DM, Langa KM. Long-term cognitive impairment and functional disability among survivors of severe sepsis. *JAMA* 2010; **304:** 1787–94.
28 Pandharipande PP, Girard TD, Jackson JC, et al, and the BRAIN-ICU Study Investigators. Long-term cognitive impairment after critical illness. *N Engl J Med* 2013; **369:** 1306–16.
29 Browall S, Backhaus E, Naucler P, et al. Clinical manifestations of invasive pneumococcal disease by vaccine and non-vaccine types. *Eur Respir J* 2014; **44:** 1646–57.

30 Corrales-Medina VF, Madjid M, Musher DM. Role of acute infection in triggering acute coronary syndromes. *Lancet Infect Dis* 2010; **10:** 83–92.

31 Yende S, D'Angelo G, Mayr F, et al, and the GenIMS Investigators. Elevated hemostasis markers after pneumonia increases one-year risk of all-cause and cardiovascular deaths. *PLoS One* 2011; **6:** e22847.

32 Huang ST, Lin CL, Chang YJ, et al. Pneumococcal pneumonia infection is associated with end-stage renal disease in adult hospitalized patients. *Kidney Int* 2014; **86:** 1023–30.

33 Graham S, Walker JL, Andrews NJ, et al. Quantifying and adjusting for confounding from health-seeking behaviour and healthcare access in observational research. *MedRxiv* 2024; published online February 28. https://doi.org/10.1101/2024.02.27.24303434.

34 Torres A, Blasi F, Dartois N, Akova M. Which individuals are at increased risk of pneumococcal disease and why? Impact of COPD, asthma, smoking, diabetes, and/or chronic heart disease on community-acquired pneumonia and invasive pneumococcal disease. *Thorax* 2015; **70:** 984–89.

35 Hughes GJ, Wright LB, Chapman KE, Wilson D, Gorton R. Serotype-specific differences in short- and longer-term mortality following invasive pneumococcal disease. *Epidemiol Infect* 2016; **144:** 2654–69.

36 Kruckow KL, Zhao K, Bowdish DME, Orihuela CJ. Acute organ injury and long-term sequelae of severe pneumococcal infections. *Pneumonia* 2023; **15:** 5.

37 National Institute for Health and Care Excellence. Diabetes–type 2: QOF indicators. 2024. https://cks.nice.org.uk/topics/diabetes-type-2/goals-outcome-measures/qof-indicators/ (accessed Oct 13, 2025).

38 Harboe ZB, Thomsen RW, Riis A, et al. Pneumococcal serotypes and mortality following invasive pneumococcal disease: a population-based cohort study. *PLoS Med* 2009; **6:** e1000081.