

Supplementary File A

SUPPLEMENTARY METHODS

Co-morbidities

For ordered categorical variables (ischaemic heart disease, cancer, liver disease, diabetes) patients were classified as having the lower category of disease until the date of their first code at the higher level, when they permanently changed. If patients only had codes for the higher level (e.g. metastatic cancer) they were immediately categorised as having the higher level of disease.

Neurological disease included Parkinson disease and other extra-pyramidal/movement disorders, epilepsy, multiple sclerosis and other demyelinating diseases, systemic atrophies primarily affecting the central nervous system, hereditary and idiopathic neuropathy, cerebral palsy, Creutzfeldt-Jakob disease and atypical viral infections of the central nervous system.

Disorders of the immune mechanism included aplastic anaemia, immunodeficiency with predominantly antibody defects, combined immunodeficiencies, immunodeficiency associated with other major defects, common variable immunodeficiency, functional disorders of polymorphonuclear neutrophils, chronic myeloproliferative disease, and unspecified immunodeficiency. Four episodes of CAP with HIV infection were also included in this category.

Terminal illness was defined using Read and ICD10 codes stating terminal illness, rather than specific conditions. In addition, GP information on referrals to hospices was included.

There was no historical time limit placed on the recording of co-morbidities.

Frailty

To define frailty, Read & ICD10 code lists, and additional recording fields within the data were used, such as the 'enttype', 'constype' and 'nhsspec' fields in CPRD, and the 'admimeth' field in HES.

Place of residence was defined using CPRD records only, as over 85% of records in HES were coded as admitted from 'Usual place of residence'. Within CPRD the most recent record (using Read codes, or the constype field) was used. The categories lives alone, sheltered accommodation and residential care were mutually exclusive, and where more than one category was used on a date the most severe was assigned.

Recently had carer was defined using CPRD records only. If there were no records relating to the patients place of residence, or the code was from over a year before the CAP index date, then the patient was coded as having a carer if the carer code was within the year prior to the CAP.

If there was a residence code within the last year, then the patient was coded as having a carer if there was a carer code in the time since the residence code.

Self care problems included codes for inability to; wash/clean self, maintain personal hygiene, get dressed without assistance, use the toilet without assistance, perform housekeeping activities, buy and/or prepare food. Appropriate enttypes were also used.

Low weight/poor nutrition included codes for weight loss, malnourishment, lack of appetite, anorexia, and prescriptions for nutritional supplementation. Appropriate enttypes were also used.

Bedsore/ulcer included codes for bedsores, pressure sores/ulcers, immobility and being bedbound.

Mobility problems included codes for problems with general mobility, stairs, walking, being house or chair bound, in a wheelchair, use of a zimmer/walking frame and difficulties getting in/out of bed. Appropriate enttypes were also used.

Incontinence included catheter codes, and those for bladder or bowel incontinence. Appropriate enttypes were also used.

Visual impairment used records from both CPRD and HES (using Read, ICD10 codes and enttypes). To aid data management, visual codes were categorised as long-term codes (such as blindness) which could not be treated, and those which were potentially treatable. These were further categorised as cataracts (diagnosis or removal) and other vision codes. Within HES, cataract codes within the primary code of the first episode (the admitting diagnosis) were assumed to be due to cataract removal surgery. All other HES cataract codes were considered cataract diagnoses.

Patients with any long-term/other/cataract diagnosis code (or a combination of all three) at any point before CAP but without a cataract removal code before the CAP were coded as having a vision problem.

If the patient's most recent code was for cataract removal, then earlier records were examined to determine how many other cataract removals the patient had had. If there were two or more removal codes over six weeks apart they were assumed to have had two cataracts removed.

History of falls used information from both CPRD and HES to define whether a patient had fallen within the last year (excluding records on the CAP incident date). Codes for falls and hip fractures were used as it has been estimated that up to 95% of hip fractures are caused by falls.[1]

Medications

Information for all medication variables was obtained from the therapy files in CPRD. We used the time between the CAP incident date and the closest preceding prescription to define exposure to the drug of interest using different time windows depending on the medication.

Inhaled corticosteroids timing of inhaled corticosteroids (ICS) was based upon recent work by Suissa et al.[2] CAP episodes without an ICS prescription in the previous year were classed as 'non-users'. Those who did have a prescription were categorised as latest prescription up to 60 days prior to CAP, 61-180 days prior, and 181-365 days prior.

Immunosuppressants other than steroids. Patients were categorised as being recent/current immunosuppressant users if they had a prescription record within 120 days of their CAP event date (90 day prescription + 30 day washout).

Oral steroids. Patients were categorised as being recent/current steroid users if there was evidence of a prescription for oral steroids within or lasting to within the 90 days pre-CAP.[3] If there was a prescription in the 14 days pre-CAP but not the 90 days before patients were not classified as 'on

steroids'. Due to a lack of information on current dose, we used the 90 day model (rather than the WCD model), as it achieved the best fit among the 'binary' models in the Dixon paper.

Statins use was defined as a prescription within the previous six months (183 days).

Antibiotics prescriptions (excluding those in BNF chapter 5.1.13, for urinary tract infections) were classified as in the 1-7 or, 8-28 days pre CAP.

Vaccines

Vaccination status was determined using GP records alone. It was decided to prioritise records from the CPRD 'immunisation' and 'therapy' files, as they were thought more likely to be recording a current event rather than an historical one. As it takes time for a vaccine to elicit an antibody response, patients were considered 'unvaccinated' for the 14 days post-vaccination date.[4, 5] Patients with only a 'history of vaccination' code were classified as being vaccinated >5 years ago, as there was no way of identifying the correct date. Timings were categorised in similar way to Vindogrova et al.[6]

Influenza vaccine. The categories in this variable needed to reflect that influenza vaccine is updated and given yearly. The vaccination season was defined as 1st September to 31st August the next year (in line with the date used for Public Health England vaccination uptake levels).

Those classed as unvaccinated included CAPs without a prior influenza vaccination record, those who were first vaccinated less than 14 days pre CAP, or first vaccinated post-CAP.

Those vaccinated more than 2 weeks prior to CAP were classified according to their most recent vaccination date as: vaccinated this season pre CAP, vaccinated last season, vaccinated 2-5 years ago and vaccinated >5 years ago.

Pneumococcal vaccine. Those classed as unvaccinated included CAPs without a prior pneumococcal vaccination record, those who were first vaccinated less than 14 days pre CAP, or first vaccinated post-CAP.

Those vaccinated more than 2 weeks prior to CAP were classified according to their most recent vaccination date as: vaccinated more than 14 days, but less than one year ago, vaccinated 1-2 years ago, vaccinated 2-5 years ago and vaccinated >5 years ago.

Health behaviours

Excess alcohol included those with the Read/ICD code for alcoholic, or a code specifying harm due to alcohol. Patients recorded as consuming ≥ 6 units of alcohol a day, or ≥ 42 units a week, and those with prescriptions for anti-alcohol drugs at any time before/on the date of the CAP episode were also included.

Smoking status was derived using Read/ICD10 codes, prescriptions for nicotine replacement therapy and appropriate enttypes. Patients could not be a non-smoker if they had previous current or ex-smoker records, and so were re-coded to ex-smokers. Patients on nicotine replacement therapy were coded as current-smokers. The most recent record prior to the CAP episode was used.

REFERENCES

- 1 Nyberg L, Gustafson Y, Berggren D, et al. Falls leading to femoral neck fractures in lucid older people. *J Am Geriatr Soc*. 1996;**44**:156-60. DOI.
- 2 Suissa S, Patenaude V, Lapi F, et al. Inhaled corticosteroids in COPD and the risk of serious pneumonia. *Thorax*. 2013;**68**:1029-36. DOI: 10.1136/thoraxjnl-2012-202872.
- 3 Dixon WG, Abrahamowicz M, Beauchamp M-E, et al. Immediate and delayed impact of oral glucocorticoid therapy on risk of serious infection in older patients with rheumatoid arthritis: a nested case–control analysis. *Ann Rheum Dis*. 2012. DOI: 10.1136/annrheumdis-2011-200702.
- 4 Public Health England, Department of Health. Influenza (Chapter 19) 2013. In: Immunisation Against Infectious Disease (The Green Book) [Internet]. Available from: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/239268/Green_Book_Chapter_19_v5_2_final.pdf.
- 5 Public Health England, Department of Health. Pneumococcal (Chapter 25). 2013. In: Immunisation Against Infectious Disease (The Green Book) [Internet]. Available from: <https://www.gov.uk/government/publications/pneumococcal-the-green-book-chapter-25>.
- 6 Vinogradova Y, Hippisley-Cox J, Coupland C. Identification of new risk factors for pneumonia: population-based case-control study. *Br J Gen Pract*. 2009;**59**:e329-38. DOI: 10.3399/bjgp09X472629 PubMed PMID: 19843413. Pubmed Central PMCID: 2751937.