**Supplementary Data 1: Serotypes contained within each pneumococcal vaccination and vaccination group.**

|  |  |
| --- | --- |
| **Vaccination** | **Serotypes** |
| Pneumococcal polysaccharide vaccine, 23 valent (PPV-23, PneumoVax®) | 1, 2, 3, 4, 5, 6B, 7F, 8, 9N, 9V, 10A, 11A, 12F, 14, 15B, 17F, 18C, 19F, 19A, 20, 22F, 23F, 33F. |
| Pneumococcal conjugate vaccine, 7-valent (PCV-7, Prevenar®) | 4, 6B, 9V, 14, 18C, 19F and 23F |
| Pneumococcal conjugate vaccine, 13-valent (PCV-13, Prevenar13®) | 1, 3, 4, 5, 6A, 6B, 7F, 9V, 14, 18C, 19A, 19F, and 23F |
| PCV13-7 serotypes | 1, 3, 5, 6A, 7F, 19A |

**Supplementary Data 2: Characteristics of patients admitted with pneumococcal parapneumonic effusions**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Characteristic | All parapneumonic effusions  (N=559) | SPE  (N=459) | Pleural infection  (N=100) |  |
| Gender - N (%) | | | |  |
| Male | 282 (50) | 223 (49) | 56 (56) | 0.205 |
| Female | 277 (50) | 236 (51) | 44 (44) | 0.205 |
| Age at Presentation – yr | | | |  |
| Age – median (IQR) | 68 (50-78) | 69 (57-78) | 55 (45-74) | <0.001 |
| Smoking Status - N (%) † | | | |  |
| Non-smoker | 127 (23) | 106 (23) | 21 (21) | 0.665 |
| Current smoker | 196 (35) | 148 (32) | 48 (48) | 0.002 |
| Ex-smoker | 238 (43) | 205 (45) | 33 (33) | 0.028 |
| Index of Deprivation - N (%) ‡ | | | |  |
| Unknown/No fixed abode | 2 (0) | 1 (0) | 1 (1) | - |
| 1-2 | 88 (16) | 73 (16) | 15 (15) | 0.804 |
| 3-4 | 115 (21) | 91 (20) | 24 (24) | 0.372 |
| 5-6 | 103 (18) | 80 (17) | 23 (23) | 0.158 |
| 7-8 | 119 (21) | 93 (20) | 26 (26) | 0.183 |
| 9-10 | 134 (24) | 121 (26) | 13 (13) | 0.006 |
| Vaccination Status | | | |  |
| Pneumococcal vaccination status – N/total N (%) # | | | |  |
| < 6 months of admission | 27 (5) | 19 (4) | 8 (8) | 0.088 |
| ≥ 6 months of admission | 231 (41) | 202 (44) | 29 (29) | 0.006 |
| Seasonal influenza vaccination – N/total N (%) # | | | |  |
| < 6 months of admission | 231 (41) | 193 (42) | 38 (38) | 0.462 |
| ≥ 6 months of admission | 53 (9) | 42 (9) | 11 (11) | 0.534 |
| Comorbidities – N (%) | | | |  |
| Chronic Lung Disease | 245 (44) | 201 (43) | 44 (44) | 0.856 |
| - COPD | 164 (29) | 140 (31) | 24 (24) | 0.166 |
| - Asthma | 56 (10) | 43 (9) | 13 (13) | 0.221 |
| - Other Respiratory Disease \* | 71 (13) | 59 (13) | 12 (12) | 0.787 |
| Cardiac Disease | 304 (54) | 261 (57) | 43 (43) | 0.011 |
| - Atrial Fibrillation | 94 (17) | 83 (18) | 11 (11) | 0.090 |
| - Hypertension | 221 (40) | 210 (38) | 36 (36) | 0.709 |
| - IHD | 128 (23) | 108 (24) | 20 (20) | 0.392 |
| - Cardiac Failure | 56 (11) | 53 (12) | 3 (3) | 0.008 |
| - Other Heart Disease \*\* | 23 (4) | 21 (5) | 2 (2) | 0.188 |
| Other Conditions |  |  |  |  |
| Upper GI/Swallow Problem | 214 (38) | 180 (40) | 34 (34) | 0.265 |
| Cerebrovascular Disease | 39 (7) | 29 (6) | 10 (10) | 0.148 |
| Type 2 Diabetes | 95 (17) | 79 (17) | 16 (16) | 0.809 |
| Chronic Liver Failure | 25 (4) | 21 (5) | 4 (4) | 0.672 |
| Intravenous Drug Usage | 24 (4) | 17 (4) | 7 (7) | 0.192 |
| Alcohol excess | 53 (9) | 38 (8) | 15 (15) | 0.028 |
| BMI ≤17 or BMI ≥ 35 | 23 (4) | 16 (3) | 7 (7) | 0.056 |
| Immunosuppression |  |  |  |  |
| Autoimmune Disease | 42 (8) | 33 (7) | 9 (9) | 0.488 |
| Corticosteroids | 22 (4) | 19 (4) | 3 (3) | 0.637 |
| Immunosuppressive Medication | 21 (4) | 19 (4) | 2 (2) | 0.334 |
| Solid Organ Malignancy | 82 (15) | 70 (15) | 12 (12) | 0.44 |
| Haematological Malignancy | 50 (9) | 44 (10) | 6 (6) | 0.212 |
| Chemotherapy | 20 (4) | 18 (4) | 2 (2) | 0.333 |

† smoking status as determined by recording in patient medical records from relevant admission or other hospital records

\* Other chronic respiratory disease includes pulmonary fibrosis, bronchiectasis, and other rare/orphan lung diseases

\*\* Other chronic cardiac disease includes valvular heart disease, structural cardiac anomalies and other rare cardiac diseases, but excludes arrhythmias

‡ Index of deprivation was calculated using postcode to derive Lower-layer Super Output Area (LSOA) data from The Office of National Statistics (ONS)

# Vaccination status for pneumococcal and seasonal influenza was verified for each patient from GP records.

**Supplementary Data 3: Clinical features and outcomes in pneumococcal parapneumonic effusions**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Characteristic | All parapneumonic effusions  (N=559) | SPE  (N=459) | Pleural infection  (N=100) | P-value |
| Clinical Characteristics |  |  |  |  |
| Respiratory rate, median (IQR) | 30 (24-32) | 28 (24-32) | 30 (24-32) | 0.124 |
| Oxygen saturations, median (IQR) | 87 (84-91) | 88 (84-91) | 86 (83-90) | 0.213 |
| Hypotension, median % (IQR) | 88 (16) | 69 (15) | 19 (19) | <0.001 |
| AMTS ≤7, median % (IQR) | 162 (29) | 129 (28) | 33 (33) | 0.318 |
| CURB65 Score – N (%) |  |  |  |  |
| 0 | 28 (5) | 20 (4) | 8 (8) | 0.088 |
| 1 | 107 (19) | 83 (18) | 24 (24) | 0.167 |
| 2 | 139 (25) | 112 (24) | 27 (27) | 0.528 |
| 3-5 | 28 5 (51) | 244 (53) | 42 (42) | 0.046 |
| Investigation Results – median (IQR) | | | |  |
| White cell count,×109/L | 17.8 (12.1-23.8) | 17.0 (12.0-22.7) | 20.8 (13.9-27.0) | 0.005 |
| Neutrophil count, ×109/L | 14.9 (10.0-20.2) | 14.2 (9.9-20.1) | 18.5 (10.8-24.1) | 0.003 |
| Lymphocyte count, ×109/L | 1.1 (0.6-1.8) | 1.1 (0.7-1.8) | 1.2 (0.6-1.9) | 0.913 |
| C-reactive protein, mg/dL | 215 (133-341) | 200 (129-312) | 331 (213-412) | <0.001 |
| Sodium, mg/dL | 137 (132-140) | 137 (133-140) | 137 (132-140) | 0.666 |
| eGFR, ml/min/1.73m2 | 55 (34-76) | 55 (35-76) | 56 (33-75) | 0.715 |
| Urea, mmol/L | 9.3 (6.4-13.2) | 9.2 (6.5-13.3) | 9.7 (6.07-12.8) | 0.666 |
| Albumin, g/L | 30 (26-36) | 31 (27-36) | 29 (23-34.3) | 0.666 |
| Complications, N (%) |  |  |  |  |
| Acute renal failure | 198 (35) | 170 (37) | 28 (28) | 0.088 |
| Liver dysfunction | 60 (11) | 55 (12) | 5 (5) | 0.041 |
| ITU Care, N (%) |  |  |  |  |
| ITU admission | 110 (20) | 92 (20) | 18 (18) | 0.649 |
| Non-invasive ventilation | 37 (7) | 35 (8) | 2 (2) | 0.032 |
| ETT insertion | 92 (16) | 80 (17) | 12 (12) | 0.218 |
| Tracheostomy | 26 (5) | 20 (4) | 6 (6) | 0.375 |
| Inotropes | 89 (16) | 79 (17) | 10 (10) | 0.082 |
| Discharge Outcome |  |  |  |  |
| Increased care on discharge – N (%) | 198 (35) | 176 (38) | 22 (22) | 0.002 |
| Length of stay (days) – median (IQR) | 11 (6-20) | 10 (5-20) | 14 (10-21) | <0.001 |
| Mortality, N (%) |  |  |  |  |
| Inpatient mortality | 71 (13) | 60 (13) | 11 (11) | 0.586 |
| 30-day mortality | 74 (13) | 65 (14) | 9 (9) | 0.180 |
| 90-day mortality | 106 (19) | 94 (21) | 14 (14) | 0.039 |
| 1-year mortality | 150 (26) | 134 (29) | 16 (16) | 0.008 |

Normal Ranges: White cell count 4.0-11.0×109/L, neutrophil count 2.0-7.5×109/L. lymphocyte count 1.5-4.5×109/L, C-reactive protein <10mg/dL, Sodium 135-145mg/dL, eGFR ml/min/1.73m2>60, urea 2.5-7.1mmol/L, albumin 35-50g/L

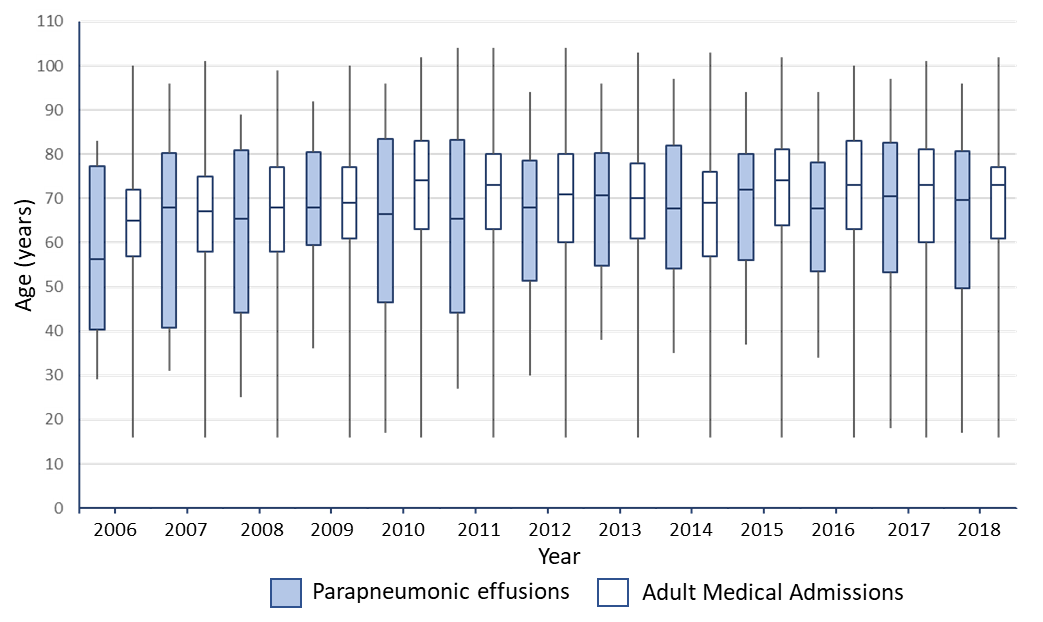
**Supplementary Table 4: Annual population and hospital admissions**

Population size within the Bristol, North Somerset and South Gloucestershire CCG and Bath & NE Somerset CCG as based on data obtained from the Office for National Statistics (ONS). Number of adult patients seen are listed as a combined total for all 3 NHS Trusts, from data provided by NHS Digital. Data for patients seen in Emergency Departments is derived from Hospital Accident and Emergency Activity. Data for admissions is collated from Monthly Hospital Episode Statistics for Admitted Patient Care, Outpatient and Accident and Emergency data

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Characteristic – N** | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 | 2015 | 2016 | 2017 | 2018 |
| **Population** |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Whole population | 1,031,918 | 1,040,674 | 1,047,210 | 1,053,791 | 1,028,596 | 1,070,120 | 1,080,886 | 1,131,336 | 1,104,211 | 1,118,820 | 1,131,336 | 1,139,791 | 1,140,236 |
| ≥ 16 y | 823,298 | 833,548 | 839,310 | 844,186 | 866,522 | 887,470 | 925,556 | 893,556 | 902,950 | 915,438 | 925,566 | 933,304 | 940,293 |
| 16-34 y | 237,508 | 287,368 | 288,798 | 288,536 | 297,243 | 306,361 | 301,317 | 301,317 | 305,932 | 313,489 | 319,384 | 322,914 | 323,715 |
| 35-49 y | 222,704 | 224,970 | 225,290 | 225,944 | 227,259 | 220,217 | 219,883 | 219,883 | 217,895 | 217,136 | 216,015 | 215,618 | 215,987 |
| 50-64 y | 176,998 | 178,690 | 179,822 | 180,714 | 184,550 | 186,294 | 185,960 | 185,960 | 188,442 | 191,215 | 194,082 | 196,991 | 197,565 |
| 65-79 y | 106,488 | 107,636 | 109,918 | 112,516 | 118,948 | 129,021 | 140,629 | 133,018 | 136,428 | 138,803 | 140,629 | 141,792 | 142,895 |
| ≥80 y | 34,094 | 34,884 | 35,487 | 36,476 | 38,522 | 45,577 | 55,456 | 53,378 | 54,253 | 54,795 | 55,456 | 55,989 | 60,131 |
| **Adults Admissions** |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Accident & Emergency | 266,411 | 261,328 | 270,012 | 266,714 | 268,253 | 271,948 | 279,953 | 286,314 | 275,960 | 28,3137 | 263,566 | 269,868 | 271,328 |
| Unplanned admissions | 103,590 | 96,665 | 123,004 | 105,774 | 108,151 | 107,704 | 106,753 | 109,235 | 114,415 | 120,960 | 126,315 | 133,640 | 137,521 |
| RR admissions \* | 1.00 | 0.93 | 1.19 | 1.02 | 1.04 | 1.04 | 1.03 | 1.05 | 1.10 | 1.17 | 1.22 | 1.29 | 1.33 |
| **Tests Conducted** |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Blood Cultures | 41,975 | 42,586 | 42,033 | 43,735 | 44,983 | 46,620 | 46,021 | 47,955 | 48,336 | 50,028 | 50,852 | 51,186 | 52,904 |
| Urine Antigen | 52 | 143 | 308 | 960 | 1,591 | 1,909 | 2,141 | 2,251 | 2,645 | 3,517 | 3,535 | 3,706 | 3,405 |
| RR testing \* | 1.00 | 1.02 | 1.01 | 1.06 | 1.11 | 1.15 | 1.15 | 1.19 | 1.21 | 1.27 | 1.29 | 1.31 | 1.34 |

\* The relative ratio of unplanned hospital admissions and microbiological testing (blood culture and urine antigen combined) were calculated using 2006 as the reference year

**Supplementary Data 5: Box plot of age of adults hospitalised with parapneumonic effusions and all medical admissions 2006-2018**

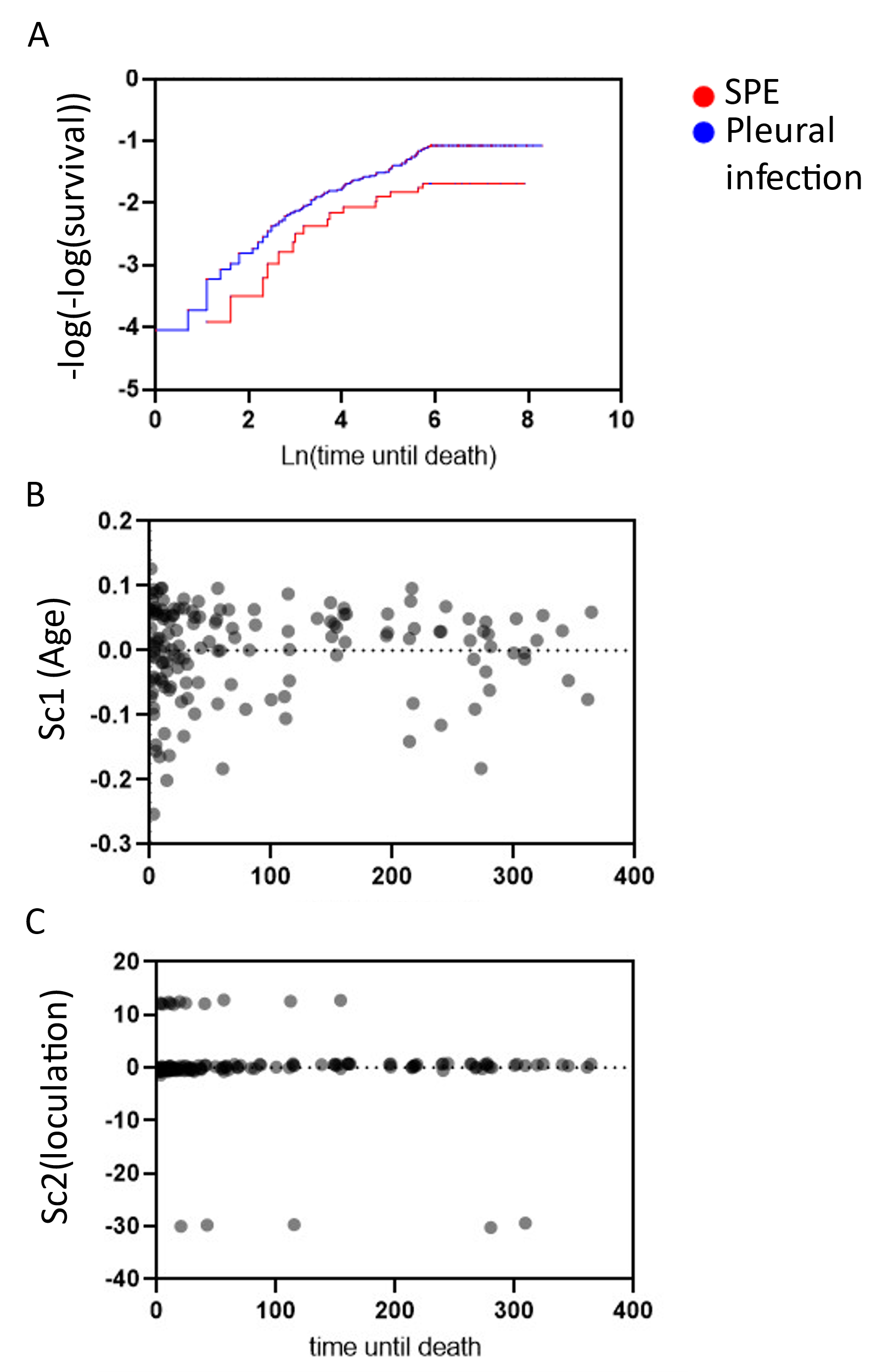
****

**Supplementary Data 6: Characteristics of Patients with pleural infection, by survival following presentation**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Characteristic - N (%) | **All Pleural Infection** | | **SPE** | | **Pleural Infection** | |
| Survived >90 days | Died within 90 days | Survived >90 days | Died within 90 days | Survived >90 days | Died within 90 days |
|  | N=453 | N = 106 | N=365 | N = 94 | N = 86 | N = 14 |
| Gender | | | | | | |
| - Male | 228 (50) | 57 (56) | 181 (49) | 50 (53) | 45 (52) | 7 (50) |
| - Female | 225 (50) | 49 (46) | 184 (51) | 44 (47) | 41 (47) | 7 (50) |
| Age – mean (IQR) | 69  (52-89) | 78  (68-84) | 68  (50-80) | 80  (68-89) | 56  (44-71) | 73  (67-88) |
| Smoking Status | | | | | | |
| - Current smoker | 174 (38) | 28 (26) | 128 (35) | 26 (28) | 44 (51) | 3 (21) |
| - Ex smoker | 180 (40) | 61 (58) | 159 (43) | 52 (55) | 21 (24) | 9 (64) |
| Vaccination Status |  |  |  |  |  |  |
| - PneumoVax (PPV23) | 204 (45) | 56 (58) | 176 (48) | 47 (50) | 27 (32) | 9 (64) |
| - Seasonal Flu | 239 (53) | 39 (40) | 195 (53) | 37 (39) | 43 (50) | 5 (35) |
| Pneumococcal vaccine serotype | | | | | | |
| Unknown serotype | 282/453 (62) | 55/106 (52) | 231/365 (63) | 50/94 (53) | 50/86 (58) | 5/14 (35) |
| PCV13 vaccine serotype | 95/181  (52) | 31/57 (54) | 65/143 (45) | 26/50 (52) | 30/36 (83) | 5/9 (56) |
| Non-vaccine serotype (NVT) | 16/181  (8) | 6/57  (11) | 14/143 (10) | 5/50 (10) | 2/36  (5) | 1/9 (11) |
| CPE/Empyema features | | | | | | |
| - Pus | 59 (13) | 0 (0) | - | - | 59 (66) | 0 (0) |
| - Loculation | 39 (9) | 10 (10) | - | - | 39 (45) | 10 (71) |
| - Fibrinolytic treatment | 27 (6) | 6 (6) | 0 (0) | 0 (0) | 27 (31) | 6 (43) |
| - Surgical Management | 29 (6) | 0 (0) | 0 (0) | 0 (0) | 29 (34) | 0 (0) |
| Co-morbidities | | | | | | |
| Chronic Resp Disease | 166 (37) | 54 (51) | 130 (36) | 49 (52) | 36 (42) | 5 (36) |
| COPD | 123 (27) | 44 (42) | 104 (28) | 40 (42) | 19 (22) | 4 (29) |
| Asthma | 48 (11) | 9 (8) | 35 (9) | 9 (10) | 13 (15) | 0 (0) |
| Other Chronic Resp Disease | 57 (10) | 11 (10) | 50 (13) | 9 (10) | 7 (8) | 2 (14) |
| Chronic Cardiac Disease | 226 (50) | 74 (69) | 190 (52) | 68 (72) | 36 (42) | 6 (43) |
| Hypertension | 187 (41) | 57 (54) | 156 (43) | 59 (63) | 31 (36) | 4 (33) |
| Ischaemic Heart Disease | 89 (19) | 38 (36) | 71 (19) | 39 (41) | 18 (21) | 1 (7) |
| AF | 63 (14) | 27 (25) | 59 (16) | 23 (24) | 4 (5) | 4 (29) |
| LVF | 42 (9) | 13 (12) | 39 (11) | 13 (14) | 3 (3) | 0 (0) |
| Other CVS | 16 (4) | 2 (2) | 16 (4) | 5 (5) | 0 (0) | 2 (14) |
| Immunosuppression |  |  |  |  |  |  |
| Autoimmune Disease | 24 (5) | 8 (8) | 16 (4) | 7 (7) | 8 (9) | 1 (7) |
| Oral Corticosteroids | 18 (4) | 4 (4) | 16 (4) | 3 (3) | 2 (2) | 1 (7) |
| Other immunosuppressive med | 17 (4) | 6 (6) | 14 (4) | 5 (5) | 3 (3) | 1 (7) |
| Solid organ cancer | 58 (13) | 24 (23) | 49 (13) | 21 (21) | 9 (10) | 3 (21) |
| Haematological malignancy | 38 (8) | 13 (12) | 34 (9) | 10 (11) | 4 (5) | 3 (21) |
| Recent chemotherapy | 15 (3) | 5 (5) | 14 (4) | 4 (4) | 1 (1) | 1 (7) |
| Other Risk Factors |  |  |  |  |  |  |
| Upper GI/Swallow problem | 168 (37) | 54 (52) | 136 (37) | 48 (51) | 29 (34) | 6 (43) |
| Chronic Liver Disease | 17 (4) | 8 (8) | 14 (4) | 7 (7) | 3 (3) | 1 (7) |
| Diabetes Mellitus | 81 (18) | 16 (15) | 66 (18) | 16 (17) | 15 (17) | 0 (0) |
| CVA disease | 24 (5) | 13 (12) | 19 (5) | 10 (11) | 5 (6) | 3 (21) |
| Alcohol misuse | 41 (9) | 11 (10) | 29 (8) | 9 (10) | 12 (14) | 2 (14) |
| Intravenous drug usage | 16 (3) | 8 (8) | 10 (3) | 7 (7) | 6 (7) | 1 (7) |
| Obesity or BMI<17 | 16 (3) | 4 (4) | 9 (3) | 4 (4) | 7 (8) | 0 (0) |
| Admission Features |  |  |  |  |  |  |
| Length of Stay – median (IQR) | 11 (7-23) | 14 (6-24) | 10 (5-19) | 10 (4-21) | 17 (11-24) | 18 (11-28) |
| ICU/HDU care | 93 (20) | 20 (20) | 75 (20) | 18 (21) | 18 (21) | 2 (14) |
| - Intubation and ventilation | 73 (16) | 20 (20) | 62 (17) | 18 (21) | 11 (13) | 2 (14) |
| - Inotropic support | 71 (15) | 21 (21) | 62 (17) | 19 (22) | 9 (10) | 2 (14) |

**Supplementary Data 7: Testing of proportional hazards in Cox Regression model**

(A) Log-Log(survival) curve plot of SPE (red line) and pleural infection (blue line) cases in the cohort. Scaled Schoenfeld residuals are shown for (B) age and (C) loculation, demonstrating log hazard-ratio function is constant over time as gradient of each slope is zero.

****

**Supplementary Data 8: Table of Abbreviations and Definitions**

|  |  |
| --- | --- |
| Abbreviation | Explanation |
| AF | Atrial Fibrillation |
| AMTS | Abbreviated Mental Test Score |
| BMI | Body Mass Index |
| CAP | Community Acquired Pneumonia |
| COPD | Chronic Obstructive Pulmonary Disease |
| CPE | Complex Parapneumonic Effusion |
| CVA | Cerebrovascular Accident |
| CVS | Cardiovascular System |
| ETT | EndoTracheal Tube |
| HIV | Human Immunodeficiency Virus |
| HDU | High dependency unit |
| IHD | Ischaemic Heart Disease |
| IPD | Invasive Pneumococcal Disease |
| IQR | Interquartile Range |
| ITU | Intensive Therapy Unit |
| LRTI | Lower Respiratory Tract Infection |
| LSOA | Lower-layer Super Output Area |
| LVF | Left Ventricular Failure |
| MALDI-TOF | Matrix-assisted laser desorption/ionisation/time of flight |
| ONS | The Office of National Statistics |
| PCV | Polysaccharide conjugate vaccine |
| PHE | Public Health England |
| PPV23 | Pneumococcal Polysaccharide Vaccine |
| RAPID | Renal, Age, Purulence, Infection source, & Dietary factors Score |
| SD | Standard Deviation |
| SPE | Simple Parapneumonic Effusion |
| ST | Serotype |