CDC Antimicrobial Resistance Threats Report

Drug-resistant *Streptococcus pneumoniae* Data Methods

The total estimate of drug-resistant *Streptococcus pneumoniae* (*S. pneumoniae*) infections was calculated as the sum of the estimated number of drug-resistant invasive and non-invasive *S. pneumoniae* infections. Invasive *S. pneumoniae* infections were defined as detection of *S. pneumoniae* in a normally sterile body site, such as blood, cerebrospinal fluid, or internal body sites like bone. Non-invasive infections included all other illnesses caused by *S. pneumoniae*. The number of invasive drug-resistant *S. pneumoniae* infections was estimated using Active Bacterial Core surveillance (ABCs) data [1]. CDC uses the ABCs platform to conduct surveillance for invasive *S. pneumoniae* infections and antimicrobial resistance [1].

The number of non-invasive drug-resistant *S. pneumoniae* infections was estimated using the following three-step multiplicative model:

1. All-cause burden of syndromes associated with *S. pneumoniae* was estimated using administrative data.
2. Estimates of the proportion of these syndromes attributable to *S. pneumoniae* (etiologic fraction) were determined and multiplied by the all-cause burden to estimate the burden of *S. pneumoniae* disease.
3. Estimates of the *S. pneumoniae* burden of disease were multiplied by the estimated percent resistant.

Estimates from both invasive and non-invasive syndromes were summed to find the total estimate of drug-resistant *S. pneumoniae*.

Definition of drug-resistant *S. pneumoniae*

An isolate (pure sample of *S. pneumoniae*) was eligible for inclusion in the drug-resistant *S. pneumoniae* case definition if it was tested for susceptibility to each of these antimicrobial agents: one penicillin drug (penicillin, amoxicillin) and one macrolide drug (erythromycin) and one cephalosporin (cefotaxime, ceftriaxone) and one tetracycline (tetracycline or doxycycline) and one fluroquinolone (levofloxacin) and trimethoprim-sulfamethoxazole. Any isolate that was nonsusceptible (resistant or intermediate) to at least one of the above antimicrobial agents was considered drug-resistant.

Invasive syndromes

The burden of invasive drug-resistant *S. pneumoniae* infections was estimated using data from ABCs [1]. An invasive case reported to ABCs is categorized as drug-resistant if the corresponding isolate was predicted to have intermediate susceptibility or resistance through whole genome sequencing. Missing demographic information (e.g., age group) was imputed using multiple imputation (MI) combined with Random Forest models. Projected national invasive infection counts were estimated by scaling ABCs catchment-area populations to the total U.S. population within age and race strata. Uncertainty in the estimates encompasses uncertainty from the MI algorithm, the variation from Random Forest models, and uncertainty from bootstrapping.

Non-invasive all-cause syndromes

To establish a list of the major clinical syndromes (encompassing both invasive and non-invasive disease) associated with *S. pneumoniae* (Table 1), a clinical work group comprised of pediatric, adult, and infectious disease physicians developed a comprehensive ICD-10-CM code list for each syndrome identifying medical visits of interest.

The Premier Healthcare database (PHD) was used to project national estimates of the number of hospitalized cases (inpatient visits) for each all-cause clinical syndrome. A validated model was used to develop extrapolated estimates of incident case counts [2]. Using iterative proportional fitting (raking), weights were developed to adjust for hospital characteristics in participating PHD hospitals compared to hospitals in the American Hospital Association annual survey data [2,3]. Events of interest were identified using the primary diagnosis code for each discharge from syndrome-specific ICD-10-CM code lists. A 30-day washout period was applied following each event before a new event could occur for that patient. A weighted-means survey procedure was applied to calculate national estimates of hospitalized cases for each syndrome for all age groups in each calendar year.

For people aged <65 years, outpatient rates of non-hospitalized syndromes were estimated using the MarketScan database [4]. Events of interest were identified from both primary and secondary diagnoses using the ICD-10-CM code list. Outpatient episodes that escalated to inpatient illness within the same ICD-10-CM code category using primary diagnosis codes during the 30 days preceding or 30 days following the outpatient visit were excluded.

Event-free person-time (PT) was accumulated for enrollees from the beginning of the year or start of enrollment until the event, unenrollment, or end of the year. After a hospitalized or non-hospitalized event, a 30-day washout window began during which the person did not contribute event-free PT and remained enrolled. If another *S. pneumoniae*-relevant ICD-10-CM code occurred during the washout period, the washout period re-started from the new visit date. Each participant was enrolled for the 30-day washout period before contributing to PT. Syndrome specific rates were estimated as the number of events divided by total enrolled PT for each calendar year. Outpatient streptococcal disease incidence rates for people aged ≥65 years were estimated using Medicare claims data (CMS) with the same analysis plan [5]. Estimates included beneficiaries aged ≥65 years enrolled in either Medicare Fee-for-Service (Parts A and B) or Medicare Advantage (Part C) [6]. Medicare enrollment files provided entitlement and demographic information; death was treated as a censoring event for PT calculations.

Outpatient incidence estimates for people aged <65 years were stratified by age group and sex. Incidence estimates from the MarketScan Commercial Claims database (CCAE) were weighted by the proportion of the population with private insurance, and incidence estimates from the MarketScan Medicaid database were weighted by the proportion of the population without private insurance [7]. Weighted MarketScan CCAE and Medicaid incidence estimates were summed and multiplied by U.S. Census population estimates from the American Community Survey 1-Year Estimates within age-sex strata to generate a nationally representative estimate of the total number of cases associated with each outpatient syndrome [8]. Medicare outpatient incidence estimates were multiplied by U.S. Census population estimates for people aged ≥65 years. Total all-cause, outpatient disease burden was the sum of projected estimates from MarketScan and CMS.

For each year, syndrome, age group, and setting, uncertainty in incidence rates or case counts was propagated using appropriate probability distributions (Table 2), primarily Poisson distributions for outpatient rates and normal or negative binomial distributions for hospitalized case counts. Monte Carlo simulations were conducted to generate distributions of projected case counts. National case counts were projected by multiplying each simulated value by the corresponding census population sizes for each age group. PHD estimates were already reported as national case counts; after propagating uncertainty around these estimates, adjustment using census data was not necessary.

Non-invasive etiologic fraction

The etiologic fraction is the proportion of all-cause disease attributable to *S. pneumoniae*. Using data from a literature review, Premier microbiology data, and ABCs invasive case counts, an internal collaborator panel estimated the appropriate etiologic fraction estimate and bounds for each clinical syndrome caused by *S. pneumoniae*. Etiologic fraction estimates and bounds were stratified by age group (<18 and ≥18 years) and clinical syndrome. All-cause burden for each syndrome and age group was multiplied by the etiologic fraction to determine *S. pneumoniae* disease burden.

The Program Evaluation and Review Technique (PERT a, b, c) distribution was applied to propagate uncertainty in etiologic fraction estimates. This approach translates expert opinion regarding minimum, most likely, and maximum values into a smooth parametric distribution, emphasizing the most likely value. For each age group and syndrome, random variables generated from the PERT distribution were utilized as multiplicative factors in the burden estimates, incorporating uncertainty in the etiologic attribution.

Proportion resistant

ABCs data were used to estimate the proportion of resistant isolates. While ABCs monitors invasive disease, drug resistance data are gathered systematically for all isolates and are population based, making the data less biased than other sources of drug resistance data. The definition of drug-resistant *S. pneumoniae* is included above. When ABCs cases are missing information about tested isolates, drug resistance results were imputed based on Random Forest models built on bootstrapped datasets using MI outputs.

The proportion of invasive cases resistant to antibiotics was estimated using multiple imputation results from ABCs data, which served as proxies for non-invasive infections, assuming the same resistant proportion for both. Uncertainty in these proportions was propagated through final burden estimates.

Combining invasive and non-invasive infections

Drug-resistant non-invasive *S. pneumoniae* infections were estimated by multiplying *S. pneumoniae* disease burden by the ABCs estimated proportion of drug-resistant isolates. Estimates for drug-resistant invasive *S. pneumoniae* infections from ABCs were combined with estimates of drug-resistant non-invasive *S. pneumoniae* infections to estimate a total burden for drug-resistant *S. pneumoniae* infections.

Mortality

Mortality associated with drug-resistant *S. pneumoniae* infections was estimated by projecting deaths associated with drug-resistant invasive *S. pneumoniae* infections using a multiple imputation model from ABCs data accounting for missing isolates and susceptibility results. To estimate deaths due to non-invasive syndromes, we projected the national

number of deaths among inpatient visits with primary ICD-10-CM codes for pneumonia and acute exacerbation of chronic bronchitis in PHD. The 3-step multiplicative model previously described was applied to estimate the number of deaths associated with non-invasive drug-resistant *S. pneumoniae* infections. The number of deaths for both invasive and non-invasive drug-resistant *S. pneumoniae* infections were then summed.

Uncertainty in the estimates

The methodology estimated uncertainty at three primary steps: all-cause incidence, etiologic fractions, and the proportion of drug-resistant cases. Statistical techniques such as multiple imputation, expert panel input, and Monte Carlo simulations were utilized to propagate and characterize uncertainty. At each step, uncertainty was quantified and propagated using Monte Carlo simulations or bootstrapping. For each data source, age group, and syndrome, probability distributions were assigned to key variables and simulations conducted to generate outcome distributions. These distributions were combined arithmetically to produce total projected counts and associated uncertainty intervals for both outpatient and inpatient populations. The final burden estimates were characterized by point estimates and 95% credible intervals (range of values likely to include the true value), reflecting cumulative uncertainty from all sources.

References

1. Centers for Disease Control and Prevention. Active Bacterial Core surveillance (ABCs). Accessed February 11, 2026. https://www.cdc.gov/abcs/about/index.html
2. Jernigan JA, Hatfield KM, Wolford H, et al. Multidrug-resistant bacterial infections in U.S. hospitalized patients, 2012–2017. *N Engl J Med*. 2020;382(14):1309–1319. doi:10.1056/NEJMoa1914433
3. American Hospital Association. AHA Annual Survey Database. Accessed November 20, 2025. https://www.ahadata.com
4. Merative. Merative MarketScan Research Databases. Accessed November 20, 2025. https://www.merative.com/documents/merative-marketscan-research-databases
5. Centers for Medicare & Medicaid Services. Parts of Medicare. Accessed November 20, 2025. https://www.medicare.gov/basics/get-started-with-medicare/medicare-basics/parts-of-medicare
6. Centers for Medicare & Medicaid Services. Medicare Monthly Enrollment. Accessed December 18, 2024. https://data.cms.gov/summary-statistics-on-beneficiary-enrollment/medicare-and-medicaid-reports/medicare-monthly-enrollment
7. Nelson LM, Wallin MT, Marrie RA, et al. A new way to estimate neurologic disease prevalence in the United States: Illustrated with MS. *Neurology*. 2019;92(10):469-480. doi:10.1212/WNL.0000000000007044
8. U.S. Census Bureau. ACS public use microdata sample (PUMS). Accessed November 20, 2025. https://search.library.wisc.edu/catalog/9910020716102121

Table 1: Syndromes included in analysis

Pulmonary
- Pneumonia
- Empyema
- Acute exacerbation of chronic bronchitis

Invasive not otherwise categorized
- Meningitis
- Intracranial abscess
- Endocarditis
- Bacteremia
- Sepsis
- Septic shock
- Toxic shock syndrome
- Pericarditis and peritonitis

Bone and joint
- Osteomyelitis and septic arthritis
- Prosthetic joint infections
- Diabetic foot infections

Gynecologic and urinary infections
- Endometritis, chorioamnionitis, puerperal sepsis, septic abortion, and pelvic inflammatory disease
- Urinary tract infections

Head, ears, eyes, nose and throat (HEENT)
- Acute otitis media (AOM)
- Mastoiditis
- Sinusitis
- Conjunctivitis
- Pharyngitis, tonsillitis and epiglottitis
- Neck abscess (retropharyngeal and peritonsillar abscess)
- Cervical adenitis

Skin and soft tissue infections (SSTI)
- Cellulitis
- Necrotizing fasciitis
- Skin abscess
- Myositis
- Erysipelas
- Impetigo

Table 2: Assumed Distribution of Each Data Source to Estimate Uncertainty

Data Source	Input Data	Distribution
CMS (65+)	Rate	~Poisson (λ)
MarketScan	Rate	~Poisson (λ)
Premier	Case count, std. dev	~Normal (μ,σ^2)

For more information on drug-resistant *Streptococcus pneumoniae*, visit: https://arthreats.cdc.gov/spneumoniae

Revised October 2026