University of Louisville
Journal of Respiratory Infections

REVIEW ARTICLE

Most Common Statistical Methodologies in Recent Clinical Studies of Community-Acquired Pneumonia

**Stephen Furmanek[1], Connor English[1], Thomas Chandler[1], Timothy Wiemken[2]*

Abstract

Background: Training new individuals in pneumonia research is imperative to produce a new generation of clinical investigators with the expertise necessary to fill gaps in knowledge. Clinical investigators are often intimidated by their unfamiliarity with statistics. The objective of this study is to define the most common statistical methodologies in recent clinical studies of CAP to inform teaching approaches in the field.
Methods: Articles met inclusion criteria if they were clinical research with an emphasis on incidence, epidemiology, or patient outcomes, searchable via PubMed or Google Scholar, published within the timeframe of January 1st 2012 to August 1st 2017, and contained Medical Subject Headings (MeSH) keywords of "pneumonia" and one of the following: "epidemiologic studies", "health services research", or "comparative effectiveness research" or search keywords of community-acquired pneumonia" and one of the following: "cohort study", "observational study", "prospective study", "retrospective study", "clinical trial", "controlled trial", or "clinical study". Descriptive statistics for the most common statistical methods were reported.
Results: Thirty articles were included in the analysis. Descriptive statistics most commonly contained within articles were frequency (n=30 [100%]) and percent (n=30 [100%]), along with medians (n=22 [73%]) and interquartile ranges (n=19 [63%]). Most commonly performed analytical statistics were the Chi-squared test (n=20 [67%]), logistic regression (n=18 [60%]), Fisher's exact test (n=17 [57%]), Wilcoxon rank sum test (n=16 [53%]), T-test (n=13 [43%]), and Cox proportional hazards regression (n=10 [33%]).
Conclusions: We identified the most common clinical research tests performed in studies of hospitalized patients with CAP. Junior investigators should become very familiar with these tests early in their research careers.

DOI: 10.18297/jri/vol1/iss4/9
Received Date: August 7, 2017
Accepted Date: September 12, 2017
Website: https://ir.library.louisville.edu/jri
Affiliations:
[1]University of Louisville Division of Infectious Diseases
[2] University of Louisville School of Public Health and Information Sciences, Department of Epidemiology and Population Health

Introduction

Community-acquired pneumonia (CAP) is the leading cause of death from infectious disease in the United States. [1] Guidelines for the management of hospitalized patients with CAP have been structured around the results of clinical research activities. Although there is a wealth of data in this area, there are still important knowledge gaps for the management of hospitalized patients with CAP that will need to be resolved with further clinical research. Training new individuals is imperative to produce a new generation of clinical investigators with the expertise necessary to fill these gaps in knowledge. Within the University of Louisville Division of Infectious Diseases, we have implemented novel clinical research training programs with this goal in mind.

In these courses, several challenges were discovered when teaching biostatistics. First, although a comprehensive understanding of all statistical methodologies used in a particular field is ideal, in practice this may be unnecessary. By attempting to teach with a goal of comprehensive understanding, practical ideas are lost in a sea of formulas. Furthermore, clinical investigators are often intimidated by their unfamiliarity with statistics. In an attempt to resolve some of these challenges, we consider that it is important to limit statistical teaching to the most pertinent tests and topics in clinical research. Although this knowledge does not replace the importance of enlisting the expertise of a biostatistician for all stages of a research project, it does facilitate a broad understanding of the basic concepts of biostatistics and assists in critical analysis of study results. This knowledge leads to higher quality studies and a better understanding of how results of other studies can be incorporated into clinical practice.

The primary objective of this study is to define the most common statistical methodologies in recent clinical studies of CAP to inform teaching approaches in the field. Secondary objectives

*Correspondence To: Stephen Furmanek
Work: Address: 501 East Broadway, Suite 120B Louisville, KY 40202
Work Email: stephen.furmanek@louisville.edu

were to: 1) define the most common study designs, 2) define statistical tests to be included in a curriculum to educate clinical investigators interested in clinical research of CAP, 3) generate a glossary defining the most common statistical tests.

Methods

This was a literature review including recent publications in the field of CAP. Articles were eligible for inclusion in this study if they: 1) were clinical research with an emphasis on incidence, epidemiology, or patient outcomes, 2) were searchable via PubMed or Google Scholar, 3) were published within the timeframe of January 1st 2012 to August 1st 2017, 4) contained Medical Subject Headings (MeSH) keywords of “pneumonia” and one of the following: “epidemiologic studies”, “health services research”, or “comparative effectiveness research” or search keywords of “community-acquired pneumonia” and one of the following: “cohort study”, “observational study”, “prospective study”, “retrospective study”, “clinical trial”, “controlled trial”, or “clinical study”. A Delphi panel of CAP clinical investigators decided by vote on which 30 articles to include in this review based on their clinical relevance and applicability to clinical practice.

Only statistics mentioned in the body of the article were included in our evaluation; information in supplementary appendices were not evaluated.

Statistical methodologies were divided into five categories: 1) descriptive statistics, 2) inferential statistics and procedures, 3) graphics and figures, 4) study design, and 5) statistical software.

Descriptive statistics were performed to summarize the tests and procedures performed, represented by frequency and percent. The most frequent statistical methodologies were reported. R software version 3.3.2 was used for all analysis.

A test was considered for curriculum if it was present in more than 30% of the articles reviewed. Based on this criterion, we identified statistical tests to be included in a curriculum to educate individuals interested in clinical research of CAP. Tests were grouped into sessions based on their application in clinical research.

Using this curriculum, we developed a glossary for the most common statistical tests for clinical investigators in the field of CAP. Definitions were written so that they are meaningful to non-statisticians and did not include mathematical formulae.

Results

Thirty studies were evaluated based on decisions by the Delphi panel [2-31]. Study characteristics are shown in **Table 1**. Every study that was evaluated included a descriptive analysis of patient characteristics. Every study included frequencies for various descriptive analyses, notated by n, and percent. Medians were the most frequently reported measures of central tendency (n=22 [73%]), and interquartile ranges (IQRs; n=19 [63%]) were the most frequently reported measure of variability.

P-values were reported in 27 (90%) studies. In 19 (63%) studies, cutoff values for statistical significance were explicitly stated. Confidence intervals were reported in 21 (70%) studies. Bivariate comparisons were performed most commonly with Wilcoxon rank sum tests (n=16 [53%]) and Chi-squared tests (n=20 [67%]). Multivariable analyses were most commonly logistic regression (n=17 [60%]) and Cox Proportional Hazards regression (n=10 [33%]).

Table 1. Study characteristics

Variable	n (%)
Study Design	
Retrospective	7 (23)
Prospective	23 (77)
Cohort	19 (63)
Prospective	13 (68)
Retrospective	6 (32)
Case-Control	1 (3)
Randomized Controlled Trial	10 (33)
Descriptive Statistics	
n	30 (100)
%	30 (100)
Mean	16 (53)
Standard Deviation	13 (43)
Median	22 (73)
Interquartile Range	19 (63)
Minimum	1 (3)
Maximum	1 (3)
Analyses or procedures	
Type I error rate or significance level (alpha)	14 (47)
Power (concerning sample size)	4 (13)
p values	27 (90)
Confidence Intervals	21 (70)
T-test (or Z-test)	13 (43)
Shapiro-Wilk’s test	1 (3)
Wilcoxon Rank Sum test	16 (53)
Chi-squared test	20 (67)
Fisher’s Exact test	17 (57)
McNemar’s test	1 (3)
One Way Analysis of Variance	4 (13)
Linear Regression	1 (3)
Linear Mixed Model	1 (3)
Logistic Regression	18 (60)
Hosmer-Lemeshow test	5 (17)
Poisson Regression	1 (3)
Log-Rank test	3 (10)
Cox Regression	10 (33)
Positive/Negative Predictive Value	2 (7)
Area under the curve	3 (10)
Sensitivity analysis	8 (27)
Multiple imputation adjustment	2 (7)
Attributable Fractions	1 (3)
Non-Inferiority testing	4 (13)
Graphics or Figures	
Study Flowchart or Diagram	17 (57)
Pie Charts	2 (7)
Bar Charts	8 (27)
Line Charts	2 (7)
Kaplan Meier Curves	6 (20)
ROC Curve	6 (20)

The most common figure was study flow diagram (n=17 [57%]). The most common graphic was a bar chart (n=8 [27%]).

The majority of studies were cohort studies (n=19, [63%]). The majority of cohort studies were prospective (n=13, [68%]).

The most frequent software used was SPSS (n=9, [30%]). Five studies (17%) mentioned multiple software used, and eight studies (27%) did not specify which software was used for analysis.

Six statistical tests occurred in more than 30% of the articles reviewed. These tests were: Chi-squared tests, Fisher’s exact tests, t-tests, Wilcoxon Rank Sum tests, logistic regression, and Cox Proportional Hazards regression. Curriculum session are depicted in **Table 2**.

A glossary defining the most common statistical tests is depicted in **Table 3.**

Table 2. Statistical Test Curriculum for Clinical Investigators

Session	Material Covered
Session 1	T-test and Wilcoxon Rank Sum test 1. Normality of Data 2. Means vs. Medians 3. Standard Deviation vs. IQR 4. Picking the appropriate test 5. Test statistics 6. Interpretation of Results
Session 2	Chi-squared test and Fisher's exact test 1. Contingency Tables 2. Observed and Expected Counts 3. Picking the appropriate test 4. Test statistics 5. Interpretation of Results
Session 3	Logistic Regression 1. Model building and variable selection 2. The 10:1 rule 3. Assessing model fit 4. Interpretation of results
Session 4	Cox Proportional Hazards Regression 1. When to use vs. Logistic Regression 2. Model building and variable selection 3. The proportional hazards assumption 4. Interpretation of results

Table 3. Glossary of terms

Term	Definition
Chi-squared Test	There are many types of chi-squared tests, but most commonly this refers to either a chi-squared test of homogeneity or chi-squared test of independence in the clinical literature. This test is most commonly used to compare two patient populations in terms of categorical data. The hypotheses for these tests are different, but the end result is nearly the same: a significant p-value (typically one <0.05) indicates that the category levels are different between patient samples, either because the category levels are not dependent per group or because the groups are not homogeneous. More colloquially, a significant P-value suggests there is an association between the predictor variable (typically in the rows of the table) and the outcome variable (typically in the columns of the table). One limitation is that the chi-squared test may not be reliable when sample sizes or category levels are small.
Fisher's Exact Test	An alternative test to the chi-squared test to compare two patient populations in terms of categorical data. This test is used primarily when sample sizes or category levels are small and the chi-squared test is not reliable.
T-test	Most commonly this refers to the two-sample t-test—a test most traditionally used in clinical research to compare the means (averages) of continuous data of two patient populations. A t-test uses means and standard deviations to compare two patient groups. A significant p-value indicates that there is a difference between the two patient groups for that data. The t-test is a "parametric" test, meaning it has several assumptions that are required for it to be used appropriately.
Wilcoxon Rank Sum Test	The "non-parametric" equivalent of a t-test, also often referred to as the Mann-Whitney U-test. In clinical research, a Wilcoxon Rank Sum test is used to compare medians instead of means. It is used primarily with data that is skewed, non-normal, or otherwise violating assumptions of the t-test.
Logistic Regression	A regression used to model log-transformed odds of a dichotomous outcome (e.g. event happened vs. did not happen). Logistic regression can involve one predictor or multiple predictors. Generally, odds ratios are reported for the results. An odds ratio represents the change in odds given the change in that variable. For continuous variables, this corresponds to a 1-unit change (e.g. 1 year increase in Age). For categorical variables, this corresponds to a change from the group without the variable (e.g. those without diabetes) to those with the variable (those with diabetes).
Cox Proportional Hazards Regression	Often referred to as Cox regression. A type of time-to-event analysis that may be involve one predictor or multiple predictors. The outcome variable of cox regression is the hazard (i.e. instantaneous risk) of the event happening. Generally, *hazard ratios* are reported, which can be interpreted as the risk ratio of the event occurring. The hazard ratio is interpreted as the increase or decrease in risk of the outcome at any given time for those with the variable vs those without.

Discussion

Our study defined the most frequent statistics and tests used in clinical research of CAP. The most common descriptive statistics for categorical data were frequencies and percentages. The most common descriptive statistics for continuous data were medians and IQRs. The most common statistical tests can be split into two types: comparing study sample characteristics, and comparing study sample outcomes. Concerning study sample characteristics, the most common tests for categorical data were chi-squared tests, or Fisher's exact tests. The most common test for comparing continuous data between two study groups was the Wilcoxon rank sum test.

Tests comparing outcomes were led primarily by study design. In cohort studies, logistic regression was the most common test for outcome analysis. In randomized controlled trials, analyses were much more varied, with the most common test for outcome analysis being Cox Proportional Hazards Regression analysis.

Sessions of curriculum were chosen based on the application of tests in clinical research. Our data suggests a minimum requirement of four sessions for a basic core statistical curriculum. The t-test and Wilcoxon rank sum test are both used to compare continuous data between two study groups, so these were combined in the curriculum. Likewise, the chi-squared test and Fisher's exact test both compare categorical data between study groups, so these were also combined. While the curriculum will contain more detail on each analytical approach, the glossary provied in **Table 3** gives a fact sheet for clinical investigators on the kinds of tests they should expect to see within clinical research in the field of CAP.

An important limitation is that we concentrated only on clinical research, which is a single topic within the large field of CAP. In order to develop a curriculum for generalized clinical research, a more comprehensive review of clinical studies should be performed. Furthermore, we did not evaluate whether or not each statistical methodology or study design was appropriate for each hypothesis tested. It is possible that other methodologies may have been better suited to answer each of the questions posed by the investigators.

Finally, as articles were chosen by vote from a Delphi panel, it is possible that there may be bias in our results, or that our results may not be replicated by other researchers.

In conclusion, we identified the most common clinical research tests performed in studies of hospitalized patients with CAP. Junior investigators should become very familiar with these tests early in their research careers.

References

1. Xu J, Murphy SL, Kochanek KD, Bastian BA. Deaths: Final Data for 2013. National vital statistics reports : from the Centers for Disease Control and Prevention, National Center for Health Statistics. National Vital Statistics System. 2016;64(2):1–119.
2. Aliberti S, Brambilla AM, Chalmers JD, Cilloniz C, Ramirez J, Bignamini A et al. Phenotyping community-acquired pneumonia according to the presence of acute respiratory failure and severe sepsis. Respir Res. 2014 Mar;15(1):27.
3. Aliberti S, Ramirez J, Giuliani F, Wiemken T, Sotgiu G, Tedeschi S et al. Individualizing duration of antibiotic therapy in community-acquired pneumonia. Pulm Pharmacol Ther. 2017 Aug;45:191–201.
4. Blum CA, Nigro N, Briel M, Schuetz P, Ullmer E, Suter-Widmer I et al. Adjunct prednisone therapy for patients with community-acquired pneumonia: a multicentre, double-blind, randomised, placebo-controlled trial. Lancet. 2015 Apr;385(9977):1511–8.
5. Bonten MJ, Huijts SM, Bolkenbaas M, Webber C, Patterson S, Gault S et al. Polysaccharide conjugate vaccine against

pneumococcal pneumonia in adults. N Engl J Med. 2015 Mar;372(12):1114–25.
6. Bramley AM, Reed C, Finelli L, Self WH, Ampofo K, Arnold SR et al.; Centers for Disease Control and Prevention Etiology of Pneumonia in the Community (EPIC) Study Team. Relationship Between Body Mass Index and Outcomes Among Hospitalized Patients With Community-Acquired Pneumonia. J Infect Dis. 2017 Jun;215(12):1873–82.
7. Burgos J, Luján M, Larrosa MN, Fontanals D, Bermudo G, Planes AM et al. Risk factors for respiratory failure in pneumococcal pneumonia: the importance of pneumococcal serotypes. Eur Respir J. 2014 Feb;43(2):545–53.
8. Carugati M, Franzetti F, Wiemken T, Kelly R, Peyrani P, Blasi F, Ramirez J, Aliberti S. De-escalation therapy among bacteraemic patients with community-acquired pneumonia. Clinical Microbiology and Infection. 2015 Oct 31;21(10):936-e11.
9. Cataudella E, Giraffa CM, Di Marca S, Pulvirenti A, Alaimo S, Pisano M et al. Neutrophil-To-Lymphocyte Ratio: An Emerging Marker Predicting Prognosis in Elderly Adults with Community-Acquired Pneumonia. J Am Geriatr Soc. 2017 Aug;65(8):1796–801.
10. Chalmers JD, Akram AR, Singanayagam A, Wilcox MH, Hill AT. Risk factors for Clostridium difficile infection in hospitalized patients with community-acquired pneumonia. J Infect. 2016 Jul;73(1):45–53.
11. Chen YX, Li CS. Lactate on emergency department arrival as a predictor of mortality and site-of-care in pneumonia patients: a cohort study. Thorax. 2015 May;70(5):404–10.
12. Corrales-Medina VF, Taljaard M, Yende S, Kronmal R, Dwivedi G, Newman AB et al. Intermediate and long-term risk of new-onset heart failure after hospitalization for pneumonia in elderly adults. Am Heart J. 2015 Aug;170(2):306–12.
13. Dharmarajan K, Strait KM, Tinetti ME, Lagu T, Lindenauer PK, Lynn J et al. Treatment for Multiple Acute Cardiopulmonary Conditions in Older Adults Hospitalized with Pneumonia, Chronic Obstructive Pulmonary Disease, or Heart Failure. J Am Geriatr Soc. 2016 Aug;64(8):1574–82.
14. File TM Jr, Rewerska B, Vucinic-Mihailovic V, Gonong JR, Das AF, Keedy K et al. SOLITAIRE-IV: A Randomized, Double-Blind, Multicenter Study Comparing the Efficacy and Safety of Intravenous-to-Oral Solithromycin to Intravenous-to-Oral Moxifloxacin for Treatment of Community-Acquired Bacterial Pneumonia. Clin Infect Dis. 2016 Oct;63(8):1007–16.
15. Gadsby NJ, Russell CD, McHugh MP, Mark H, Conway Morris A, Laurenson IF et al. Comprehensive Molecular Testing for Respiratory Pathogens in Community-Acquired Pneumonia. Clin Infect Dis. 2016 Apr;62(7):817–23.
16. Garin N, Genné D, Carballo S, Chuard C, Eich G, Hugli O et al. β-Lactam monotherapy vs β-lactam-macrolide combination treatment in moderately severe community-acquired pneumonia: a randomized noninferiority trial. JAMA Intern Med. 2014 Dec;174(12):1894–901.
17. Jain S, Self WH, Wunderink RG, Fakhran S, Balk R, Bramley AM et al.; CDC EPIC Study Team. Community-Acquired Pneumonia Requiring Hospitalization among U.S. Adults. N Engl J Med. 2015 Jul;373(5):415–27.
18. Marti C, John G, Genné D, Prendki V, Rutschmann OT, Stirnemann J et al. Time to antibiotics administration and outcome in community-acquired pneumonia: secondary analysis of a randomized controlled trial. Eur J Intern Med. 2017 Sep;43:58–61.
19. McCluskey SM, Schuetz P, Abers MS, Bearnot B, Morales ME, Hoffman D et al. Serial Procalcitonin as a Predictor of Bacteremia and Need for Intensive Care Unit Care in Adults With Pneumonia, Including Those With Highest Severity: A Prospective Cohort Study. Open Forum Infect Dis. 2017 Jan;4(1):ofw238.
20. Mirouse A, Vignon P, Piron P, Robert R, Papazian L, Géri G et al. Severe varicella-zoster virus pneumonia: a multicenter cohort study. Crit Care. 2017 Jun;21(1):137.
21. Peyrani P, Wiemken TL, Metersky ML, Arnold FW, Mattingly WA, Feldman C et al. The order of administration of macrolides and beta-lactams may impact the outcomes of hospitalized patients with community-acquired pneumonia: results from the community-acquired pneumonia organization. Infect Dis (Lond). 2017 Jul;•••:1–8.
22. Postma DF, van Werkhoven CH, van Elden LJ, Thijsen SF, Hoepelman AI, Kluytmans JA et al.; CAP-START Study Group. Antibiotic treatment strategies for community-acquired pneumonia in adults. N Engl J Med. 2015 Apr;372(14):1312–23.
23. Self WH, Williams DJ, Zhu Y, Ampofo K, Pavia AT, Chappell JD et al. Respiratory Viral Detection in Children and Adults: Comparing Asymptomatic Controls and Patients With Community-Acquired Pneumonia. J Infect Dis. 2016 Feb;213(4):584–91.
24. Tagami T, Matsui H, Fushimi K, Yasunaga H. Intravenous immunoglobulin and mortality in pneumonia patients with septic shock: an observational nationwide study. Clin Infect Dis. 2015 Aug;61(3):385–92.
25. Torres A, Sibila O, Ferrer M, Polverino E, Menendez R, Mensa J et al. Effect of corticosteroids on treatment failure among hospitalized patients with severe community-acquired pneumonia and high inflammatory response: a randomized clinical trial. JAMA. 2015 Feb;313(7):677–86.
26. Uranga A, España PP, Bilbao A, Quintana JM, Arriaga I, Intxausti M et al. Duration of Antibiotic Treatment in Community-Acquired Pneumonia: A Multicenter Randomized Clinical Trial. JAMA Intern Med. 2016 Sep;176(9):1257–65.
27. van Deursen AM, van Houten MA, Webber C, Patton M, Scott DA, Patterson S et al. Immunogenicity of the 13-Valent Pneumococcal Conjugate Vaccine in Older Adults With and Without Comorbidities in the Community-Acquired Pneumonia Immunization Trial in Adults (CAPiTA). Clin Infect Dis. 2017;65(5):787–95.
28. Viasus D, Garcia-Vidal C, Manresa F, Dorca J, Gudiol F, Carratalà J. Risk stratification and prognosis of acute cardiac events in hospitalized adults with community-acquired pneumonia. J Infect. 2013 Jan;66(1):27–33.
29. Wesemann T, Nüllmann H, Pflug MA, Heppner HJ, Pientka L, Thiem U. Pneumonia severity, comorbidity and 1-year mortality in predominantly older adults with community-acquired pneumonia: a cohort study. BMC Infect Dis. 2015 Jan;15(1):2.
30. West DM, McCauley LM, Sorensen JS, Jephson AR, Dean NC. Pneumococcal urinary antigen test use in diagnosis and treatment of pneumonia in seven Utah hospitals. ERJ open research. 2016 Oct 1;2(4):00011-2016.
31. Wiemken T, Peyrani P, Bryant K, Kelley RR, Summersgill J, Arnold F, Carrico R, McKinney WP, Jonsson C, Carrico K, Ramirez J. Incidence of respiratory viruses in patients with community-acquired pneumonia admitted to the intensive care unit: results from the Severe Influenza Pneumonia Surveillance (SIPS) project. European journal of clinical microbiology & infectious diseases. 2013 May 1;32(5):705-10.