

A Closer Look at Penicillin Allergy History: Systematic Review and Meta-analysis of Tolerance to Drug Challenge

Martha DesBiens MD, MPH , Peter Scalia PhD, MSc ,
Saiganesh Ravikumar MPH , Andrew Glick MD, MPH ,
Helen Newton MPH , Okechukwu Erinne MBBS, MPH ,
Natalie Riblet MD, MPH

PII: S0002-9343(19)30860-5
DOI: <https://doi.org/10.1016/j.amjmed.2019.09.017>
Reference: AJM 15424



To appear in: *The American Journal of Medicine*

Please cite this article as: Martha DesBiens MD, MPH , Peter Scalia PhD, MSc ,
Saiganesh Ravikumar MPH , Andrew Glick MD, MPH , Helen Newton MPH ,
Okechukwu Erinne MBBS, MPH , Natalie Riblet MD, MPH , A Closer Look at Penicillin Allergy
History: Systematic Review and Meta-analysis of Tolerance to Drug Challenge, *The American Journal
of Medicine* (2019), doi: <https://doi.org/10.1016/j.amjmed.2019.09.017>

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

A CLOSER LOOK AT PENICILLIN ALLERGY HISTORY: SYSTEMATIC REVIEW AND META-ANALYSIS OF TOLERANCE TO DRUG CHALLENGE.

Martha DesBiens, MD, MPH^a; Peter Scalia, PhD, MSc^a; Saiganesh Ravikumar, MPH^a; Andrew Glick, MD, MPH^a; Helen Newton, MPH^a; Okechukwu Erinne, MBBS, MPH^a; Natalie Riblet, MD, MPH^a

^aThe Dartmouth Institute for Health Policy and Clinical Practice
One Medical Center Drive
Lebanon, NH 03766
USA

Martha.T.DesBiens.GR@Dartmouth.edu, Peter.Scalia@Dartmouth.edu, Saiganesh.Ravikumar.GR@Dartmouth.edu,
Helen.E.Newton.GR@Dartmouth.edu, Okechukwu.C.Erinne.GR@Dartmouth.edu, Natalie.BV.Riblet@Dartmouth.edu

All authors had access to the data and participated in writing this manuscript. This manuscript has been approved by all named authors and there are no other persons who satisfy the criteria for authorship but are not listed.

Funding: None

Conflicts of Interest Declaration for all authors: None

Key Words: Penicillin allergy, penicillin challenge test, allergy evaluation

Corresponding author:
Martha T. DesBiens, MD, MPH
Department of Infectious Disease
33 Germantown Rd.
Danbury, CT 06810
Martha.T.DesBiens.GR@Dartmouth.edu
Phone: 1 (603) 443-1469

CLINICAL SIGNIFICANCE

- Twenty-three quasi-experimental studies were reviewed, yielding a large sample of over 5000 participants with reported penicillin allergy undergoing drug challenge.
- Over 94% of these participants tolerated penicillin challenge without any significant reaction.
- Participants challenged based on history alone tolerated penicillin more frequently than those undergoing penicillin skin testing prior to drug challenge.
- Participants challenged with oral formulations tolerated penicillin more frequently than those given parenteral formulations.

ABSTRACT

Background: True allergy to penicillin is rare, despite the high frequency with which it is reported. While most patients reporting penicillin allergy are not prone to anaphylaxis, it is not currently known what percentage of these patients will tolerate dose challenges of penicillin-based antibiotics. This review aims to determine the rate of tolerance in patients reporting penicillin allergy when challenged with penicillin-based antibiotics.

Methods: We searched MedLine, Embase, and Cochrane Library for publications with English language translations between the years 2000 and 2017. We included randomized controlled trials, quasi-experimental, and observational studies of participants reporting penicillin allergy who received at least one systemic dose of a penicillin in the form of a drug challenge. At least two independent reviewers extracted data from included studies and assessed the quality of each included study. To generate primary outcome data, we calculated a summary estimate rate of penicillin tolerance from a pooled proportion of participants receiving penicillin with no adverse effects.

Results: Initial literature search yielded 2,945 studies, of which 23 studies were ultimately included in our review. 5,056 study participants with reported history of penicillin allergy were challenged with a penicillin. After weighting for study sample size, a pooled average of 94.4% (95CI 93.7%, 95%) participants tolerated the dose challenge without any adverse reaction.

Conclusion: Misrepresented penicillin allergy drives unnecessary use of alternative antibiotics which may be less effective, more toxic, and more expensive than using penicillin. In addressing the problem of penicillin allergy over-diagnosis, evaluation should go beyond risk for type 1 hypersensitivity. Our data suggest that 94.4% of 5,056 participants with reported penicillin allergy determined to be clinically appropriate for allergy evaluation tolerated repeat administration of penicillin-based antibiotics without any adverse reactions. This review generates meaningful information useful to clinical predictive analytics, in evaluating and managing patients with a reported history of penicillin allergy.

Authors' Declaration Interest: None of the participating authors have any conflict of interest to disclose.

BACKGROUND

Dangerous allergic reactions to penicillin are rare despite how frequently penicillin allergy is reported. About 10-20% of patients seen in the US healthcare system report an allergy to penicillin, although it is estimated only about 1% of all patients have a true type I hypersensitivity response to penicillin.¹⁻⁴ In addition, around 80% of patients with demonstrable IgE-mediated hypersensitivity will lose their allergy altogether within 10 years of avoiding the drug class.^{1,5,6} Other serious immune-mediated drug reactions such as Stevens Johnson Syndrome or Toxic Epidermal Necrolysis are seen even less frequently, developing in approximately two to seven patients per million person years.⁷ Current epidemiology suggests that the far majority of individuals reporting an allergy to penicillin do not have vulnerability to a serious drug reaction.

Patients reporting penicillin allergy are more likely to receive unnecessarily broad spectrum and less clinically appropriate antibiotics, which may be less effective, more toxic, and more expensive than the penicillins.⁸⁻¹¹ Related to this, patients with reported penicillin allergy have higher rates of *Clostridioides difficile* colitis, methicillin resistant *Staphylococcus aureus* and vancomycin resistant *Enterococcus* infections compared to matched controls.^{3,9,12,13} Penicillin allergy labeling is associated with longer hospital stays, higher drug costs, and more frequent healthcare complications including surgical site infections.^{3,9,10,13-16}

Given the rarity of serious reactions, the expectation for significant over-diagnosis of penicillin allergy and the associated implications, formal penicillin allergy evaluation has been recommended by several large clinical societies, including the Centers for Disease Control and Prevention¹⁷, the Infectious Disease Society of America¹⁸, and the American Academy of Allergy, Asthma, and Immunology¹⁹, among others. Widespread adoption of this practice has been curbed to some extent by a limited ability of general medical providers to predict patient tolerance of penicillin challenge, and concerns of exposing patients to avoidable risk. This review aims to evaluate the degree of tolerance of systemic dosing of penicillin in this patient population, with the goal of informing safe and effective evaluation of penicillin allergy.

METHODS

Review Protocol

We conducted our review following standard Cochrane methodology protocol as outlined in the Cochrane Handbook.²⁰ This protocol informed our approach to identify and review selected studies. PRISMA guidelines were followed for reporting findings.²¹ The original protocol and tracked changes are available for review upon request.

Eligibility Criteria

In accordance with our protocol, we conducted a comprehensive literature search for studies including at least one of our predetermined outcomes. We included randomized control trials, clinical trials, observational studies, cross sectional studies, and quasi-experimental studies. We evaluated studies that included adults with reported history of penicillin allergy as our population of interest, defining this as a mean sample age over 18 years. Some included studies additionally assessed pediatric patients along with adults; we excluded studies that focused exclusively on pediatric patients. Our search was limited to studies published after the year 2000 in order to ensure extracted data represented current clinical practice patterns and contemporary antibiotics.²² No limitations on publication country were implemented, although studies were only included if an English translation was available. As our study question assessed tolerance to penicillin in participants reporting prior allergy in a before/after design, trials were not expected to include a control group. Duration of follow-up was short, as adverse allergic reactions are expected to present within days to weeks of exposure.

Outcome measures

Our primary outcome of interest was tolerance of a penicillin-based antibiotic. Tolerance was defined as receipt of a systemic dose of a penicillin without any adverse reaction related to potential hypersensitivity response, with adequate follow up. Assessment of hypersensitivity reaction was left to the discretion of the study investigators. If no formal judgement was made about whether an adverse reaction was related to hypersensitivity, we considered this reaction evidence of intolerance to maintain a more conservative definition.

Secondary outcomes included differences in tolerance rates with or without initial penicillin skin testing, presence or absence of infection at the time of challenge, difference in outcomes in parenteral versus oral challenge testing, and in ambulatory versus inpatient care settings. We also evaluated antibiotic-related adverse reactions in response to drug challenge to provide a safety summary of the evaluations. These secondary outcomes were intended to provide a more comprehensive overview of potential benefits and harms of re-challenging reported penicillin allergy.

Search Methods

Databases, search terms, limits

We conducted a comprehensive literature review, searching MedLine, Cochrane Library, EMBASE, and ClinicalTrials.gov for relevant articles meeting inclusion criteria. We searched for articles published from January 2000 through June 2018 (MedLine through 6/2018; Cochrane Library issue 10 of 12, 10/2017; EMBASE through 10/2017; ClinicalTrials.gov through 10/2017). We used exploded MeSH terms and

key words to generate sets for the following themes: Penicillin, Allergy, and Drug Challenge. We then used the Boolean term “AND” to find their intersection (see **Appendix 1**). To increase the sensitivity of the search we additionally implemented a “snowballing” technique, reviewing the reference lists of all included articles, and searched trial registries for any unpublished studies.

Study Selection

Search results were collated, and five independent reviewers, including one content expert, initially screened titles and abstracts for relevancy. If there was a disagreement or uncertainty, the case was brought to the entire group where discussion reached consensus. Following this initial screen, we thoroughly reviewed full texts to assess for inclusion criteria (see **Appendix 2**). Studies were selected for final review if they contained data pertaining to at least one of our pre-stated inclusion criteria.

Assessment of Methodological Quality

Two independent reviewers assessed each eligible study for methodological quality using the Methodological Index for Non-Randomized Studies (MINORS) criteria.²³ This tool assesses non-randomized, non-comparative studies utilizing an eight-item scale (see **Appendix 3**) to compute a summary score ranging from zero to sixteen. Items were scored as follows: 0 if ‘not reported’, 1 if ‘reported but inadequate’, or 2 if ‘reported and adequate’. A score of 16 is considered ideal for non-comparative studies. Disagreements in scoring were resolved through discussion between reviewers to provide consistent scoring.

Data Collection

Data from the screenings of the full text was abstracted and compiled using a data collection form for final analysis. We collected the following data: (i) author, year, and country of publication; (ii) study setting and study design; (iii) sample demographics (e.g., gender); (iv) number of participants with reported penicillin allergy; (v) identification of a bacterial infection at the time of the challenge; (vi) if an initial skin test was conducted; (vii) the type of challenge; (viii) percentage of patients tolerating the challenge; (ix) any adverse reactions noted.

Statistical Analysis

Primary Outcome Measure

The primary objective of this systematic review was to determine the probability of tolerance to a systemic penicillin challenge among patients reporting history of penicillin allergy. In each included study, we measured the proportion of patients tolerating penicillin challenge among all patients receiving

a penicillin challenge. Two included studies, Borch et al 2006 and Macy et al 2010, measured the proportion of patients tolerating a penicillin challenge in two separate study samples in their trials, so we considered these study populations separate in our analysis. Therefore, we had 25 distinct patient populations included in our meta-analysis.

Meta-Analysis

We developed a meta-proportion using a Bayesian β -binomial model²⁴ to accommodate the high proportion of included studies that measured near-100 percent (97 percent or greater) or 100 percent challenge tolerance (12 of 24 studies) and to allow for greater flexibility in the computation of our summary measures. Our parameter of interest was the mean of our posterior binomial distribution, which represented the average probability of tolerance of a systemic penicillin challenge.

The expectation that 90% of patients reporting penicillin allergy, on average, will tolerate this drug without serious type I hypersensitivity response has been validated in prior studies.¹⁻⁴ However, the true degree of drug tolerance in patients at low risk for serious drug reactions is less well understood. Our systematic review of this population targeted patients in inpatient, outpatient and emergency department settings to be as comprehensive as possible. We therefore assumed a non-informative, or “flat” prior probability of penicillin tolerance.

We calculated the I^2 statistic to examine the variation between included studies. If the I^2 was >50% it indicated statistically significant heterogeneity.

Secondary Outcome Measures

For our secondary outcomes, we used descriptive statistics and qualitative analysis. To determine any significant difference in rates of tolerance between subgroups, odds ratios were calculated along with χ^2 tests for independence. We qualitatively described details of intolerance and toxicities from studies providing these data.

RESULTS

Description of studies

Of the 5,554 articles that met our search criteria, 2,609 duplicate entries were initially removed leaving 2,945 studies to review. Review of the titles and abstracts found 2,669 to be unrelated to our study question, leaving 276 articles for full text review. Evaluation by at least two independent reviewers

yielded the 23 studies ultimately included in our final review (see **figure 1** describing study selection flow diagram, **Appendix 4** listing included studies).

Table 1 presents the characteristics of the 23 included studies. All 23 were quasi-experimental studies published in peer-reviewed journals between 2001 and 2017. Ten of the studies were conducted in the US, four in Denmark, two in Australia, two in Canada, two in Israel, and the remaining three were conducted in Italy, Slovenia, or Spain.

Table 2 presents the methodological quality of included studies assessed by MINORS criteria. The 23 studies had a range in scores of 7 to 15, with a mean of 10.7.

Primary Outcome

A total of 6,113 participants in our meta-analysis had a reported penicillin allergy, of which 5,056 were challenged with a penicillin-based antibiotic. The most common challenge antibiotic class was aminopenicillins, followed by natural penicillins. The pooled weighted average tolerance calculated with random effects was 94.4% (95 Credible Interval 93.7, 95.0%). See **figure 2**.

Secondary Outcomes

Participants challenged based on history alone tolerated the dose challenge more frequently than those undergoing skin testing first (97% vs 86%, $p<0.0001$). A total of 595 participants were challenged without prior skin testing, and of these 579 tolerated the dose challenge (97%). A total of 4,882 participants had skin testing prior to challenge, of which 4,280 (88%) had negative skin tests, and 4,461 underwent dose challenge, with 86% fully tolerant.

Participants challenged with oral formulation of penicillins tolerated the dose challenge more frequently than those given parenteral administration of penicillins (95% vs 91%, $p<0.0001$).

Participants challenged within an inpatient care setting tolerated the dose challenge more frequently than those challenged in an ambulatory clinic (97% vs 86%, $p<0.001$).

Tolerance rates were similar between subgroups with and without bacterial infection at the time of challenge (94% vs 95%, $p=0.11$). Among participants with bacterial infections at the time of dose challenge, 1,073 of 1,148 tolerated the challenge compared to 3,701 of 3,908 of patients without concomitant infection. See **table 3**.

Adverse Reactions

Overall, 282 out of 5,056 (5.6%) participants had some adverse reaction reported following systemic penicillin challenge and were considered not tolerant of the antibiotic. Of these, 228 (4.5%) noted reactions that may be considered to be consistent with hypersensitivity: 85 participants developed isolated urticarial without systemic symptoms, 129 developed mild rash, 4 noted subjective pruritus without any objective findings, 1 developed facial erythema, 1 developed facial and hand swelling, 1 noted shortness of breath, 1 noted a metallic taste with subjective throat tightening but no shortness of breath or wheeze, 5 participants were coded as having “anaphylaxis”, however upon record review all five of these reactions occurred after on hour of dose challenge, and none of these had systemic symptoms or required epinephrine. 1 patient developed interstitial nephritis. Five participants noted reactions not likely to be related to hypersensitivity (2 developed gastrointestinal symptoms, 1 developed migraine, 1 developed linear IgA bullous dermatosis, and 1 noted an isolated vaginal ulcer). 49 reactions were not described.

DISCUSSION

Type I hypersensitivity to penicillin, the more common of the serious drug reactions, may be safely and effectively evaluated in appropriate patients.^{1,2,25-27} In patients determined to be at low risk for serious drug reaction, either through history alone or history combined with negative penicillin skin testing, drug provocation challenge is considered the gold standard for ruling out a penicillin allergy.^{1,28} This review evaluates the outcomes of such drug provocation on a large scale, assessing a large pooled sample of participants reporting penicillin allergy across a wide geography. We show that over the overwhelming majority of these participants tolerate systemic dosing of penicillin, with appropriate clinical oversight.

Summary of the main results

We analyzed 23 studies detailing penicillin challenge to patients who reported prior history of penicillin allergy. Penicillin challenge was fully tolerated in over 94% of 5,056 participants evaluated, indicating that on a population level, the vast majority of adult patients reporting penicillin allergy may safely receive penicillin treatment when indicated, after appropriate evaluation. This adds to previously accepted estimates that over 90% of patients reporting allergy to penicillin will not have serious type 1 hypersensitivity reactions to the drug.¹⁻⁴

It is worth noting that our review used a conservative definition of drug “tolerance”. Among the adverse events noted, many were not likely representative of serious allergic response. In clinical practice, many of these reactions may be safely monitored, or treated symptomatically, without interruption of antibiotic management.

Dose challenge was tolerated more frequently in participants based on allergy history alone, compared to those receiving prior skin testing. While many study protocols determined a priori to skin

test all participants, several left this decision to the discretion of the treating clinicians. The difference in tolerability among these subgroups is likely impacted by higher risk patients being selected to receive skin testing prior to dose challenge. Allergy history alone nonetheless safely identified individuals appropriate for dose challenge without prior skin testing, which supports this practice even in more resource-limited settings without allergy subspecialty support.

We note an interesting discrepancy among tolerability rates of participants challenged in inpatient versus ambulatory settings, where tolerance was greater in the former. We hypothesized that patients challenged with concomitant infection, which is more likely to be the case with admitted patients among these studies, would tolerate the challenge less commonly, owing to confounding infectious signs and symptoms being misrepresented as drug intolerance. However, we found that inpatients tolerated dose challenge more frequently than those evaluated in the ambulatory setting, and presence or absence of infection did not impact tolerance rates. Inference of infection was limited to aggregate analysis in most studies, and this may have impacted our results. Primary outcome data from ambulatory studies relied mostly on patient report, which may also have influenced this comparison.

Adverse reactions were noted in about 5% of participants evaluated. The majority of these reactions were considered mild, although reactions were not well described in 49 participants. There were no serious adverse events noted by study authors.

Overall completeness and applicability of evidence

A total of 5,056 participants received a systemic penicillin challenge. Included studies derived from multiple countries. Study-level reported rates of tolerance ranged from 81% to 100%, although it is worth noting lower tolerance rates were reported in studies with relatively small sample sizes, suggesting that chance is a likely explanation for this deviation. High rates of penicillin tolerance are consistent across multiple clinical settings. As our initial search spanned four databases, the available literature was comprehensively explored. Thus, our summary estimate of 94.4% of patients reporting penicillin allergy ultimately tolerating administration of these antibiotics may be considered reliable.

Limitations

We selected studies conducted from the year 2000 onwards to keep papers consistent with contemporary clinical practice. We excluded studies in which the challenge sample size was less than 10 to ensure reliable and consistent data. Additionally, we excluded studies that focused exclusively on pediatric samples, as immunology and allergy epidemiology tend to differ between pediatric and adult populations. This makes the results more generalizable to the adult population over 18 years of age.

A notable limitation of our study was that all included studies were quasi-experimental. The nature of our research question tended to preclude randomization, as defining proper control groups was unlikely to answer the study question. The quasi-experimental study design represents the best of limited options. Another limitation was that no study was able to achieve the ideal score of 16 on the MINORS criteria for assessment of quality of non-randomized studies. Many studies did not conduct a prospective study size calculation with power analysis, scoring 0 in this category; most studies did not blind the investigators assessing the outcome of penicillin tolerance, leading to scores of '0' in 'unbiased assessment of study endpoint' category.

Regarding our statistical analysis, Bayesian models are significantly influenced by the selection of the prior distribution. Indeed, Bayesian models are meant to build upon existing information and in that way, are particularly well suited for meta-analysis. However, in selecting a non-informative, or flat, prior probability of tolerance in our model, we disregarded potentially informative clinical knowledge. We believe, however, that selection of a flat prior was necessary because prior information was not generalizable to the patient populations included in our meta-analysis.

Agreements/disagreements with other studies or reviews

Harandian et al. in 2015 published a meta-analysis to pool penicillin tolerance when based on immediate reactions to penicillin skin test or challenge in patients with reported penicillin allergies²⁹. The authors found a rate of tolerance, defined as lack of an immediate reaction, of 92.22% (CI 90.96, 93.47%), but they went on to hypothesize that delayed, non-life-threatening reactions may be more commonly experienced. Contrary to this hypothesis, our review demonstrates very high rates of tolerance of repeat drug challenge with regards to all reactions, including delayed type reactions.

CONCLUSIONS

Implications for practice

Patients reporting prior penicillin allergy who need a β -lactam antibiotic pose an important dilemma. The summarized results of this review provide good evidence that most adult patients reporting penicillin allergy, when determined to be low risk for anaphylaxis, tolerate repeat penicillin administration without any adverse reactions. This review supports expanded penicillin allergy evaluation in appropriate patients. Such evaluation may be done safely with appropriate historical screening, and will lead to more appropriate antibiotic use, improved clinical outcomes, and improved regional antibiotic stewardship.⁸

REFERENCES

1. Drug allergy: an updated practice parameter. *Annals of allergy, asthma & immunology : official publication of the American College of Allergy, Asthma, & Immunology*. 2010;105(4):259-273.
2. Sogn DD, Evans R, 3rd, Shepherd GM, et al. Results of the National Institute of Allergy and Infectious Diseases Collaborative Clinical Trial to test the predictive value of skin testing with major and minor penicillin derivatives in hospitalized adults. *Archives of internal medicine*. 1992;152(5):1025-1032.
3. Lee CE, Zembower TR, Fotis MA, et al. The incidence of antimicrobial allergies in hospitalized patients: implications regarding prescribing patterns and emerging bacterial resistance. *Archives of internal medicine*. 2000;160(18):2819-2822.
4. Macy E. Penicillin and beta-lactam allergy: epidemiology and diagnosis. *Current allergy and asthma reports*. 2014;14(11):476.
5. Trubiano JA AN, Phillips EJ. Penicillin Allergy Is Not Necessarily Forever. *JAMA Intern Med*. 2017;318(1):82-83.
6. Blanca M, Torres MJ, Garcia JJ, et al. Natural evolution of skin test sensitivity in patients allergic to beta-lactam antibiotics. *The Journal of allergy and clinical immunology*. 1999;103(5 Pt 1):918-924.
7. Chan HL, Stern RS, Arndt KA, et al. The incidence of erythema multiforme, Stevens-Johnson syndrome, and toxic epidermal necrolysis. A population-based study with particular reference to reactions caused by drugs among outpatients. *Archives of dermatology*. 1990;126(1):43-47.
8. Miller EL. The penicillins: a review and update. *Journal of midwifery & women's health*. 2002;47(6):426-434.
9. Macy E, Contreras R. Health care use and serious infection prevalence associated with penicillin "allergy" in hospitalized patients: A cohort study. *The Journal of allergy and clinical immunology*. 2014;133(3):790-796.
10. Blumenthal KG, Parker RA, Shenoy ES, Walensky RP. Improving Clinical Outcomes in Patients With Methicillin-Sensitive Staphylococcus aureus Bacteremia and Reported Penicillin Allergy. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America*. 2015;61(5):741-749.
11. King EA, Challa S, Curtin P, Bielory L. Penicillin skin testing in hospitalized patients with beta-lactam allergies: Effect on antibiotic selection and cost. *Annals of allergy, asthma & immunology : official publication of the American College of Allergy, Asthma, & Immunology*. 2016;117(1):67-71.
12. Kelly CP, LaMont JT. Clostridium difficile--more difficult than ever. *The New England journal of medicine*. 2008;359(18):1932-1940.
13. Jeffres MN, Narayanan PP, Shuster JE, Schramm GE. Consequences of avoiding beta-lactams in patients with beta-lactam allergies. *The Journal of allergy and clinical immunology*. 2016;137(4):1148-1153.
14. Blumenthal KG, Ryan EE, Li Y, Lee H, Kuhlen JL, Shenoy ES. The Impact of a Reported Penicillin Allergy on Surgical Site Infection Risk. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America*. 2018;66(3):329-336.
15. Gupta K, Strymish J, Abi-Haidar Y, Williams SA, Itani KM. Preoperative nasal methicillin-resistant Staphylococcus aureus status, surgical prophylaxis, and risk-adjusted postoperative outcomes in veterans. *Infection control and hospital epidemiology*. 2011;32(8):791-796.
16. Macy E. Penicillin allergy: optimizing diagnostic protocols, public health implications, and future research needs. *Current opinion in allergy and clinical immunology*. 2015;15(4):308-313.
17. Prevention CfDCa. Is it really a penicillin allergy? Evaluation and Diagnosis of Penicillin Allergy for

- Healthcare Professionals. 2018.
18. Barlam TF, Cosgrove SE, Abbo LM, et al. Implementing an Antibiotic Stewardship Program: Guidelines by the Infectious Diseases Society of America and the Society for Healthcare Epidemiology of America. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America*. 2016;62(10):e51-77.
19. Immunology AAoAAa. Position statement: penicillin allergy.
20. Higgins JP GS, editor(s). Cochrane Handbook for Systematic Reviews of Interventions. *Version 510 [updated March 2011] The Cochrane Collaboration, 2011*. 2008.
21. Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *Journal of clinical epidemiology*. 2009;62(10):e1-34.
22. Macy E, Schatz M, Lin C, Poon KY. The falling rate of positive penicillin skin tests from 1995 to 2007. *The Permanente journal*. 2009;13(2):12-18.
23. Slim K, Nini E, Forestier D, Kwiatkowski F, Panis Y, Chipponi J. Methodological index for non-randomized studies (minors): development and validation of a new instrument. *ANZ journal of surgery*. 2003;73(9):712-716.
24. Karabatsos G, Talbott E, Walker SG. A Bayesian nonparametric meta-analysis model. *Research synthesis methods*. 2015;6(1):28-44.
25. Bernstein IL, Li JT, Bernstein DI, et al. Allergy diagnostic testing: an updated practice parameter. *Annals of allergy, asthma & immunology : official publication of the American College of Allergy, Asthma, & Immunology*. 2008;100(3 Suppl 3):S1-148.
26. Goldberg A, Confino-Cohen R. Skin testing and oral penicillin challenge in patients with a history of remote penicillin allergy. *Annals of allergy, asthma & immunology : official publication of the American College of Allergy, Asthma, & Immunology*. 2008;100(1):37-43.
27. American Academy of Allergy AI. Diagnostic Tests for Allergy Evaluation. *Choosing Wisely: An initiative of the ABIM Foundation*. 2012.
28. Torres MJ, Adkinson NF, Jr., Caubet JC, et al. Controversies in Drug Allergy: Beta-Lactam Hypersensitivity Testing. *The journal of allergy and clinical immunology In practice*. 2019;7(1):40-45.
29. Harandian F, Pham D, Ben-Shoshan M. Positive penicillin allergy testing results: a systematic review and meta-analysis of papers published from 2010 through 2015. *Postgrad Med*. 2016;128(6):557-562.

FIGURES

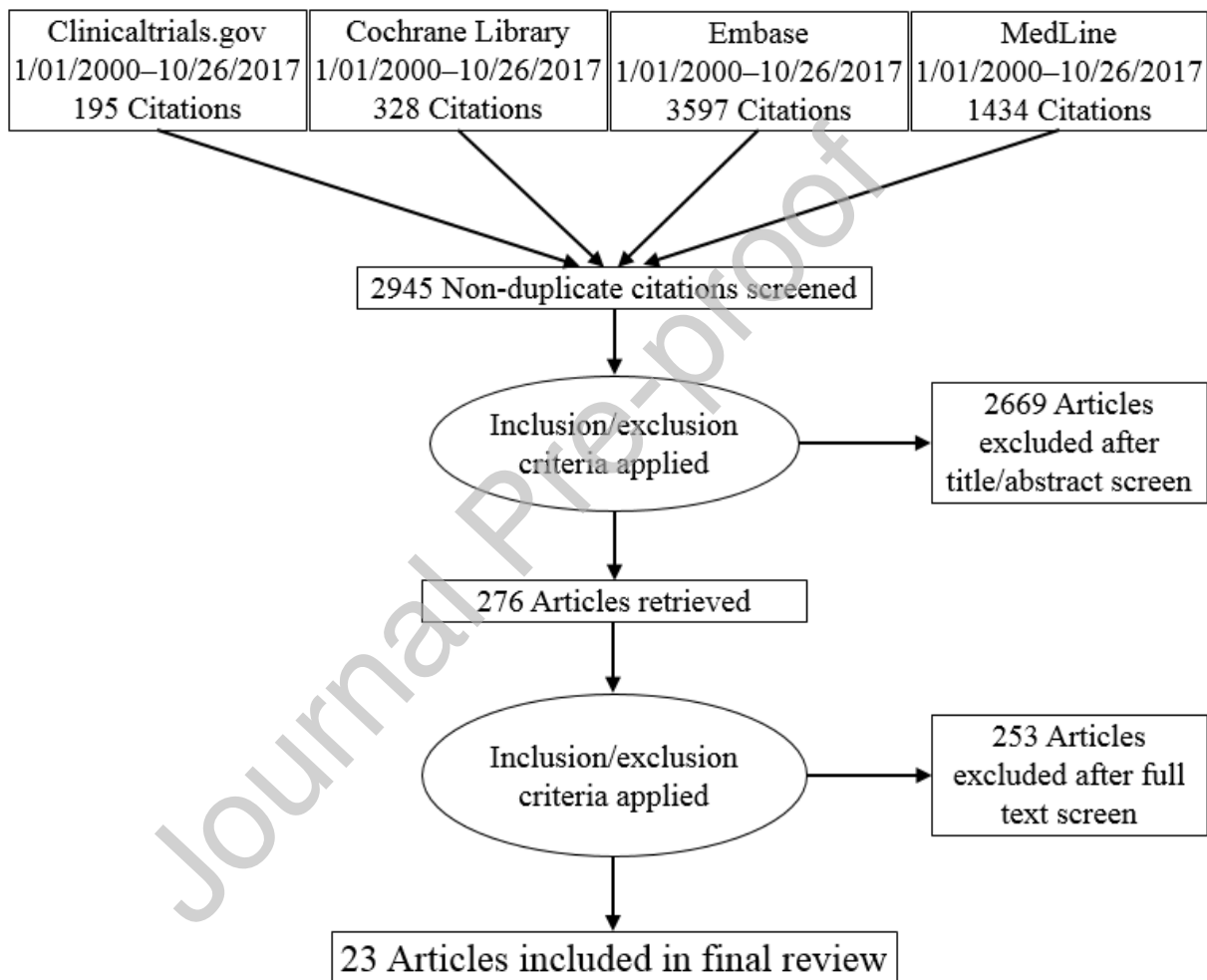
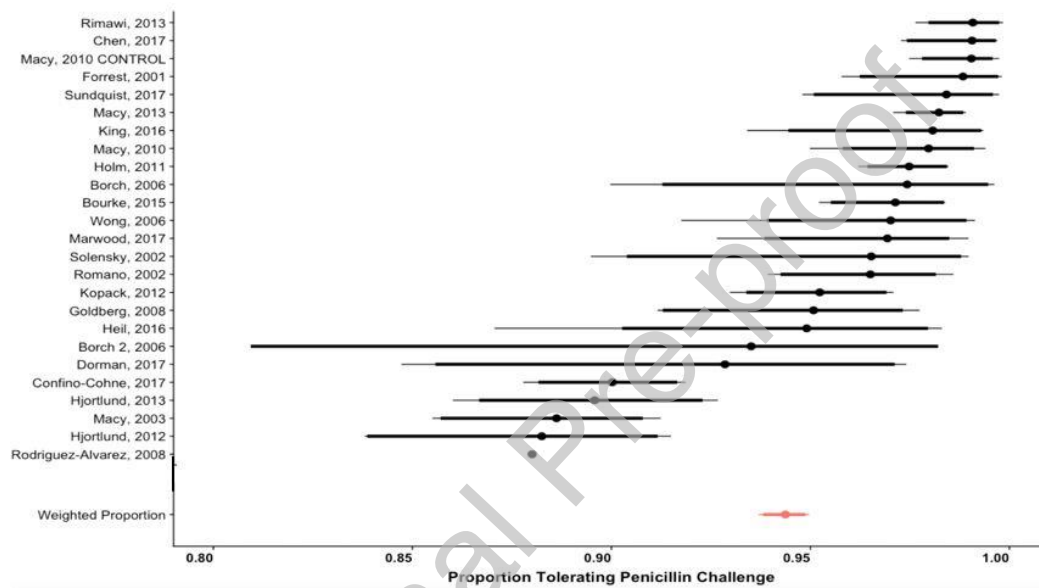
Figure 1: Flow diagram of search outputs and inclusion criteria.

Figure 2: Reported Rates of Penicillin Tolerance**Figure 2:** Pooled weighted average tolerance to penicillin drug challenge calculated with random effects was 94.4% (95 CI 93.7, 95.0%). Lower tolerance rates were reported in studies with relatively smaller sample sizes.

TABLES

Table 1: Qualitative description of included studies.

Author, Yr	Country	Study Design	Participants (n)	Clinical Setting	Gender, % Female	Age yrs, mean (range)	Infection Present?	Initial PST, n (n negative)	Type of PCN challenge	DCT, n	Percent Tolerant
Forrest et al, 2001	Canada	Quasi-exp, retrospective; pre-post	159	Inpatient	NR	NR	Y	72 (68)	FD IV, all classes	61	0.897
Romano et al, 2002	Italy	Quasi-exp, prospective	259	Ambulatory	0.73	37 (13-84)	N	259 (141)	Graded PO challenge followed by 1- week FD (all classes PO & IV)	130	0.962
Solensky et al, 2002	USA	Quasi-exp, prospective	58	Ambulatory	0.62		N	58 (53)	FD PO: 3 subsequent courses of natural penicillins	48	0.958
Macy et al, 2003	USA	Quasi-exp, prospective cohort study	568	Ambulatory	0.688	39.1 (0.2-86)	Y	568 (568)	FD PO: 80% (454) aminopenicillins, 20% (114) natural penicillins	568	0.886
Borch et al, 2006	Denmark	Combined cross sectional & Case control	96	Inpatient	0.63	61 (2-99)	N	13 (12)	Graded PO/IV challenge, natural penicillins	12	1.000
Borch et al, control 2006			43	Inpatient	0.63	61 (2-99)	Y	0	Graded PO/IV challenge, natural penicillins	15	0.867
Wong et al, 2006	Canada	Quasi-exp, retrospective review	91	Ambulatory	0.36	27 (6m-82)	N	91 (75)	Graded 7-day PO challenge, aminopenicillins	72	0.972
Goldberg et al, 2008	Israel	Quasi-exp, prospective trial	169	Ambulatory	0.627	52 (5-81)	N	169 (96)	Single dose PO challenge: 66.5% (62.6) natural penicillins, 33.5% (31.5) aminopenicillins	94	0.947
Rodriguez-Alvarez et al, 2008	Spain	Quasi-exp, retrospective	23	Ambulatory	0.48	23.4 (0.8-70)	N	23 (22)	Graded PO challenge: 83% (19) aminopenicillins, 4% (1) natural penicillins, 4% (1) β -Lactamase resistant penicillins	21	0.810
Macy et al, 2010	USA	Quasi-exp, prospective review	150	Ambulatory	0.694	51.8	N	150 (144)	Single dose PO challenge: natural penicillins & aminopenicillins	144	0.979

Macy et al, control 2010			307	Ambulatory	0.625	43.2	N	307 (304)	Single dose PO challenge: natural penicillins & aminopenicillins	304	0.993
Holm et al, 2011	Denmark	Quasi-exp, retrospective	580	Ambulatory	0.77	39 (13-87)	N	0	Graded 1-day challenge: natural penicillins (IV, PO, or both)	580	0.976
Hjortlund et al, 2012	Denmark	Quasi-exp, prospective	405	Ambulatory	0.686	46	N	405 (340)	Graded 1-day IV challenge followed by 7-day full PO course: natural penicillins	252	0.869
Kopac et al, 2012	Slovenia	Quasi-exp, prospective review	606	Inpatient	0.76	42 (14-85)	N	274 (260)	Graded PO challenge: 53% (227) natural penicillins & 45% (198) aminopenicillins	426	0.951
Hjortlund et al, 2013	Denmark	Quasi-exp, prospective	342	Ambulatory	0.7	(adult)	N	339 (304)	Graded IV challenge followed by 7-day full PO course: natural penicillins	291	0.887
Macy et al, 2013	USA	Quasi-exp, prospective cohort study	500	Ambulatory and Inpatient	0.634	40.7 (1.1-93.4)	N	500 (496)	Single dose PO challenge: aminopenicillins	496	0.984
Rimawi et al, 2013	USA	Quasi-exp, prospective	146	Inpatient	0.6	(adult)	Y	146 (145)	Graded PO challenge (natural penicillins) followed by full dose course (all classes)	145	1.000
Bourke et al, 2015	Australia and USA	Quasi-exp, prospective review	405	Inpatient	0.671	47.4 (15-85)	N	341 (299)	Single dose PO challenge: natural penicillins	301	0.970
Heil et al, 2016	USA	Quasi-exp, prospective & National Provider Survey	90	Inpatient	0.69	(adult)	Y	76 (64)	Single dose PO aminopenicillins, followed by FD course (all classes)	54	0.944
King et al, 2016	USA	Quasi-exp, retrospective	50	Inpatient	0.54	62	Y	49 (49)	Graded PO challenge: aminopenicillins	49	0.980
Chen et al, 2017	USA	Quasi-exp, prospective	252	Inpatient	0.544	49.3 (16-87)	Y	228 (224)	Graded PO challenge: aminopenicillins	224	0.996
Confino-Choen et al, 2017	Israel	Quasi-exp, prospective cohort study	642	Ambulatory	0.51	19.9 (0.1-83)	N	642 (400)	Graded PO challenge: 97% (599) aminopenicillins, 3% (18) natural penicillins	617	0.898

Dorman et al, 2017	USA	Quasi-exp, prospective	32	Ambulatory (4) and Inpatient (28)	0.53	46.9 (24-79)	Y (28)	32 (32)	Single dose PO challenge, or full course IV (all classes)	32	0.906
Marwood et al, 2017	Australia	Quasi-exp, prospective	103	ED	0.54	42 (19-69)	N	103 (87)	Graded 9-day PO challenge: aminopenicillins	84	0.964
Sundquist et al, 2017	USA	Quasi-exp, prospective cohort study	37	Ambulatory	0.51	52	N	37 (37)	Single dose PO challenge: aminopenicillins	36	1.000

Legend: Yr, yrs: year, years; Quasi-exp: quasi-experimental; PST: penicillin skin test; PCN: penicillin; DCT: dose challenge test; FD: full dose; IV: intravenous; PO: oral

Table 2: Methodological quality of included studies according to the MINORS Criteria.

Study	Clearly Stated Aim	Inclusion of consecutive patients	Prospective data collection	Endpoints appropriate to the aim of the study	Unbiased assessment of study endpoint	Follow up period appropriate for study aim	Loss to follow-up < 5%	Prospective study size calculation	Total Score
Bourke, 2015	2	2	2	1	0	2	0	0	9
Confino-Cohen, 2017	2	2	2	1	0	2	1	0	10
Goldberg, 2008	2	2	2	2	2	2	1	0	13
Kopac, 2012 (not 2010)	1	2	2	1	0	1	2	0	9
Rodriguez-Alvarez, 2008	2	2	0	2	2	2	1	0	11
Wong, 2006	2	2	1	2	0	2	2	0	11
Dorman, 2017	2	2	2	2	0	2	2	0	12
Macy, 2003	2	2	2	2	1	2	2	2	15
Macy, 2010	2	2	2	2	1	2	2	2	15
Macy, 2013	2	2	2	2	1	2	2	0	13
Chen, 2017	0	2	2	2	0	1	2	0	9
Heil, 2016	2	1	2	2	0	1	2	0	10
Hjortlund, 2012	0	2	0	2	0	2	0	0	6
Hjortlund, 2013	2	2	1	2	0	2	0	0	9
Holm, 2011	2	2	0	2	0	1	0	1	8
Marwood, 2017	2	2	1	2	0	1	2	2	12
Rimawi, 2013	2	2	2	1	0	1	2	2	12
Borch, 2006	2	1	2	2	0	2	2	0	11
Forrest, 2001	2	2	2	1	0	0	0	0	7
King, 2016	2	1	2	2	0	2	1	0	10
Romano, 2002	2	1	2	2	0	2	2	0	11
Solensky, 2002	2	2	2	2	0	2	1	2	13
Sundquist, 2017	2	2	2	2	0	2	1	0	11

Each item was scored from 0 to 2. A score of 0 indicates the information was not reported, a score 1 indicates that the information was reported but inadequate, while a score of 2 indicates the information was adequately reported. The ideal score is 16 for each study.

Table 3: Subgroup Analysis for Penicillin Tolerance

Subgroup	N studies	Proportion Tolerant	P-value
Ambulatory Evaluation	13	86%	p<0.000
Inpatient/ED Evaluation	11	97%	1
Parenteral Dose Challenge	8	91%	p<0.000
Oral Dose Challenge	16	95%	1
Penicillin Skin Testing Prior to Challenge	23	86%	p<0.000
No Penicillin Skin Testing Prior to Challenge	2	97%	1

Table 3: Participants were more likely to tolerate penicillin challenge: in hospital settings (inpatient and/or in the emergency department) than in ambulatory settings; with oral challenge formulations vs parenteral challenge formulations; and without skin testing prior to drug challenge vs with prior skin testing.

SUPPLEMENTS

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51

APPENDIX 1: Search Strategy

Search	Query	Items found
#4	#1 AND #2 AND #3	1,413
#3	challenge test OR drug challeng* OR incremental challenge OR oral challenge OR drug provocation test OR graded challenge OR dose challenge OR test dose[tiab] OR skin test OR patch test[tiab]	190,856
#2	Drug Hypersensitivity[mesh] OR Cell-mediated hypersensitivity OR Delayed-type hypersensitivity OR penicillin hypersensitivity OR penicillin allerg*[tiab] OR hypersensitivity[mesh] OR allergy and immunology[mesh] OR adverse effects[subheading] OR Drug-Related Side Effects and Adverse Reactions[Mesh] OR drug related side effects OR adverse effects	2,354,837
#1	penicillins[mesh] OR penicillin*	102,582

56 **APPENDIX 2: Inclusion Criteria and Outcomes of Interest**
57

Category	Inclusion Criteria	Justification/Explanation	Code
Study Design	Randomized controlled trials (RCT), non-randomized controlled trials, quasi-experimental studies, cohort, cross-sectional, and case-series where the total number of patients challenged is >10 and were conducted after the year 2000.	RCTs are the gold standard for evaluating interventions and provide the strongest level of evidence. They balance confounding variables present at baseline between the comparison groups. However, it is difficult to design RCTs surrounding penicillin tolerance, as it is difficult to randomize and have a separate control group when examine number of patients who tolerated penicillin challenge and will therefore accept weaker study designs including quasi-experimental. By including studies where more patients than 10 patients were challenged, we are ensuring our final summary estimates reflect better quality data. By including studies after the year 2000, we are incorporating data that represent modern clinical practice protocols.	1
Condition of interest:	Penicillin allergy by report, with receipt of systemic dose of an antibiotic.	Patients with reported penicillin allergy represent the population of concern. These patients must have an indication for receipt of antibiotics, either for treatment or test dose purposes, to facilitate studying our outcome(s).	2
Intervention:	Penicillin challenge: either observed dosing, test dose challenge, graded dose challenge; all with or without skin testing. Penicillin-based antibiotic regimen will include receipt of any of the antibiotics within the “penicillins” drug class.	There are two major ways to challenge penicillin allergy: either with challenge dosing or skin testing. We include these general criteria as “penicillin challenge”, which will include test dose challenge, graded dose challenge, and observed dosing. These interventions may be carried out with or without skin testing. Antibiotics within the “penicillins” drug class include: - Penicillin V - Penicillin G - Procaine penicillin - Amoxicillin +/- clavulanate - Ampicillin +/- sulbactam - Piperacillin - Piperacillin + tazobactam - Carbenicillin - Ticarcillin +/- clavulanate - Oxacillin	3

		- Dicloxacillin - Nafcillin	
Comparison:	May include “pre-intervention” or “pre-treatment” groups	In many common infections, penicillin-based regimens are considered first-line therapy. In patients reporting penicillin allergy, the most common clinical practice is to avoid the penicillin drug class altogether. Therefore, as the nature of our research question precludes a separate control group in most circumstances, we will use the “pre-intervention” stage as the comparator, with the assumption that without the challenge intervention these participants had previously avoided penicillins. This will be compared to the “post-intervention” state following the penicillin challenge, to generate a comparison.	4
Population:	Participants with stated penicillin allergy, who are admitted to the hospital This may include adult or pediatric populations, although our population of interest is adults over the age of 18 Studies that focus exclusively on pediatric participants will be excluded, although mixed populations will be considered.	Considering the natural duration and history of drug allergies and intolerances, and differences in immune constitution and maturation between children and adults, it is most appropriate to study this research question in adults. However, in studies with combined pediatric and adult populations, if the mean age is over 18 years these will be considered for inclusion; as pediatric patients are more likely to have a reported penicillin allergy, this inclusion will generate a more conservative measure of our summary estimate.	5
Outcomes:	≥1 pre-specified outcomes of interest (see below)	We will evaluate studies with at least one of our five pre-specified outcomes. We will consider unanticipated outcomes identified during our search process as well and will plan to amend our protocol as necessary and appropriate, to include post-hoc analysis.	6

APPENDIX 3: Methodological Quality Assessment Tool for Quasi-Experimental Studies

Methodological items for non-randomized studies	Score [†]
1. A clearly stated aim: the question addressed should be precise and relevant in the light of available literature 2. Inclusion of consecutive patients: all patients potentially fit for inclusion (satisfying the criteria for inclusion) have been included in the study during the study period (no exclusion or details about the reasons for exclusion) 3. Prospective collection of data: data were collected according to a protocol established before the beginning of the study 4. Endpoints appropriate to the aim of the study: unambiguous explanation of the criteria used to evaluate the main outcome which should be in accordance with the question addressed by the study. Also, the endpoints should be assessed on an intention-to-treat basis. 5. Unbiased assessment of the study endpoint: blind evaluation of objective endpoints and double-blind evaluation of subjective endpoints. Otherwise the reasons for not blinding should be stated 6. Follow-up period appropriate to the aim of the study: the follow-up should be sufficiently long to allow the assessment of the main endpoint and possible adverse events 7. Loss to follow up less than 5%: all patients should be included in the follow up. Otherwise, the proportion lost to follow up should not exceed the proportion experiencing the major endpoint 8. Prospective calculation of the study size: information of the size of detectable difference of interest with a calculation of 95% confidence interval, according to the expected incidence of the outcome event, and information about the level for statistical significance and estimates of power when comparing the outcomes <i>Additional criteria in the case of comparative study</i> 9. An adequate control group: having a gold standard diagnostic test or therapeutic intervention recognized as the optimal intervention according to the available published data 10. Contemporary groups: control and studied group should be managed during the same time period (no historical comparison) 11. Baseline equivalence of groups: the groups should be similar regarding the criteria other than the studied endpoints. Absence of confounding factors that could bias the interpretation of the results 12. Adequate statistical analyses: whether the statistics were in accordance with the type of study with calculation of confidence intervals or relative risk	

[†]The items are scored 0 (not reported), 1 (reported but inadequate) or 2 (reported and adequate). The global ideal score being 16 for non-comparative studies and 24 for comparative studies.

APPENDIX 4: Included Studies

1. Forrest DM, Schellenberg RR, Thien VV, King S, Anis AH, Dodek PM. Introduction of a practice guideline for penicillin skin testing improves the appropriateness of antibiotic therapy. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America*. 2001;32(12):1685-1690.
2. Romano A, Viola M, Mondino C, et al. Diagnosing nonimmediate reactions to penicillins by in vivo tests. *International archives of allergy and immunology*. 2002;129(2):169-174.
3. Solensky R, Earl HS, Gruchalla RS. Lack of penicillin resensitization in patients with a history of penicillin allergy after receiving repeated penicillin courses. *Archives of internal medicine*. 2002;162(7):822-826.
4. Macy E, Mangat R, Burchette RJ. Penicillin skin testing in advance of need: Multiyear follow-up in 568 test result-negative subjects exposed to oral penicillins. *Journal of and Clinical*. 2003;111(5):1111-1115.
5. Borch JE, Andersen KE, Bindslev-Jensen C. The prevalence of suspected and challenge-verified penicillin allergy in a university hospital population. *Basic and Clinical Pharmacology and Toxicology*. 2006;98(4):357-362.
6. Wong BBL, Keith PK, Wasserman S. Clinical history as a predictor of penicillin skin test outcome. *Annals of , Asthma and*. 2006;97(2):169-174.
7. Goldberg A, Confino-Cohen R. Skin testing and oral penicillin challenge in patients with a history of remote penicillin allergy. *Annals of allergy, asthma & immunology : official publication of the American College of Allergy, Asthma, & Immunology*. 2008;100(1):37-43.
8. Rodriguez-Alvarez M, Santos-Magadan S, Rodriguez-Jimenez B, Reig-Rincon De Arellano I, Vazquez-Cortes S, Martinez-Cocera C. Reproducibility of delayed-type reactions to betalactams. *et Immunopathologia*. 2008;36(4):201-204.
9. Kopac P, Zidarn M, Bajrovic N, Erzen R, Music E. Epidemiology of hypersensitivity reactions to penicillins in Slovenia. *European Annals of and Clinical*. 2010;42(2):45-46.
10. Macy E, Goldberg B, Poon KYT. Use of commercial anti-penicillin IgE fluorometric enzyme immunoassays to diagnose penicillin allergy. *Annals of , Asthma and*. 2010;105(2):136-141.
11. Holm A, Mosbech H. Challenge Test Results in Patients With Suspected Penicillin Allergy, but No Specific IgE. *Allergy, asthma & immunology research*. 2011;3(2):118-122.
12. Hjortlund J, Mortz CG, Skov PS, et al. One-week oral challenge with penicillin in diagnosis of penicillin allergy. *Acta dermato-venereologica*. 2012;92(3):307-312.
13. Hjortlund J, Mortz CG, Skov PS, Bindslev-Jensen C. Diagnosis of penicillin allergy revisited: The value of case history, skin testing, specific IgE and prolonged challenge. : *European Journal of and Clinical*. 2013;68(8):1057-1064.
14. Macy E, Ngor EW. Safely Diagnosing Clinically Significant Penicillin Allergy Using Only Penicilloyl-Poly-Lysine, Penicillin, and Oral Amoxicillin. *Journal of and Clinical : In Practice*. 2013;1(3):258-263.
15. Rimawi RH, Cook PP, Gooch M, et al. The impact of penicillin skin testing on clinical practice and antimicrobial stewardship. *Journal of Hospital Medicine*. 2013;8(6):341-345.
16. Bourke J, Pavlos R, James I, Phillips E. Improving the Effectiveness of Penicillin Allergy De-labeling. *The journal of allergy and clinical immunology In practice*. 2015;3(3):365-334.e361.
17. Heil EL, Bork JT, Schmalzle SA, et al. Implementation of an Infectious Disease Fellow-Managed Penicillin Allergy Skin Testing Service. *Open forum infectious diseases*. 2016;3(3):ofw155.
18. King EA, Challa S, Curtin P, Bielory L. Penicillin skin testing in hospitalized patients with β -lactam allergies: Effect on antibiotic selection and cost. *Annals of , Asthma and*. 2016;117(1):67-71.

- 135 19. Chen JR, Tarver SA, Alvarez KS, Tran T, Khan DA. A Proactive Approach to Penicillin Allergy
136 Testing in Hospitalized Patients. *The journal of allergy and clinical immunology In practice*.
137 2017;5(3):686-693.
- 138 20. Confino-Cohen R, Rosman Y, Meir-Shafir K, et al. Oral Challenge without Skin Testing Safely
139 Excludes Clinically Significant Delayed-Onset Penicillin Hypersensitivity. *The journal of allergy*
140 *and clinical immunology In practice*. 2017;5(3):669-675.
- 141 21. Dorman SM, Seth S, Khan DA. Risk of Allergic Reactions to Recurrent Intravenous Penicillin
142 Administration in Penicillin Skin Test Negative Patients. *The journal of allergy and clinical*
143 *immunology In practice*. 2017.
- 144 22. Marwood J, Aguirrebarrena G, Kerr S, Welch SA, Rimmer J. De-labelling self-reported penicillin
145 allergy within the emergency department through the use of skin tests and oral drug
146 provocation testing. *Emergency medicine Australasia : EMA*. 2017;29(5):509-515.
- 147 23. Sundquist BK, Bowen BJ, Otabor U, Celestin J, Sorum PC. Proactive penicillin allergy testing in
148 primary care patients labeled as allergic: outcomes and barriers. *Postgraduate Medicine*.
149 2017:1-6.