

Journal Pre-proof

Accuracy of Penicillin Allergy Diagnostic Tests: A Systematic Review and Meta-analysis

Bernardo Sousa-Pinto, PhD, Isabel Tarrio, MD, Kimberly G. Blumenthal, MD, Luís Araújo, MD, Luís Filipe Azevedo, PhD, Luís Delgado, PhD, João Almeida Fonseca, PhD

PII: S0091-6749(20)30726-0

DOI: <https://doi.org/10.1016/j.jaci.2020.04.058>

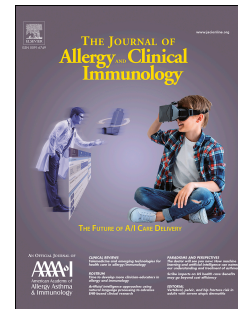
Reference: YMAI 14575

To appear in: *Journal of Allergy and Clinical Immunology*

Received Date: 14 February 2020

Revised Date: 15 April 2020

Accepted Date: 30 April 2020



Please cite this article as: Sousa-Pinto B, Tarrio I, Blumenthal KG, Araújo L, Azevedo LF, Delgado L, Fonseca JA, Accuracy of Penicillin Allergy Diagnostic Tests: A Systematic Review and Meta-analysis, *Journal of Allergy and Clinical Immunology* (2020), doi: <https://doi.org/10.1016/j.jaci.2020.04.058>.

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.

© 2020 Published by Elsevier Inc. on behalf of the American Academy of Allergy, Asthma & Immunology.



Accuracy of penicillin allergy diagnosis tests: A systematic review and meta-analysis

Scopus
PubMed
Web of Science

3952
records

*Selection of studies assessing the
accuracy of testing for penicillin allergy*

105 studies included in the systematic review

Summary
ROC curve

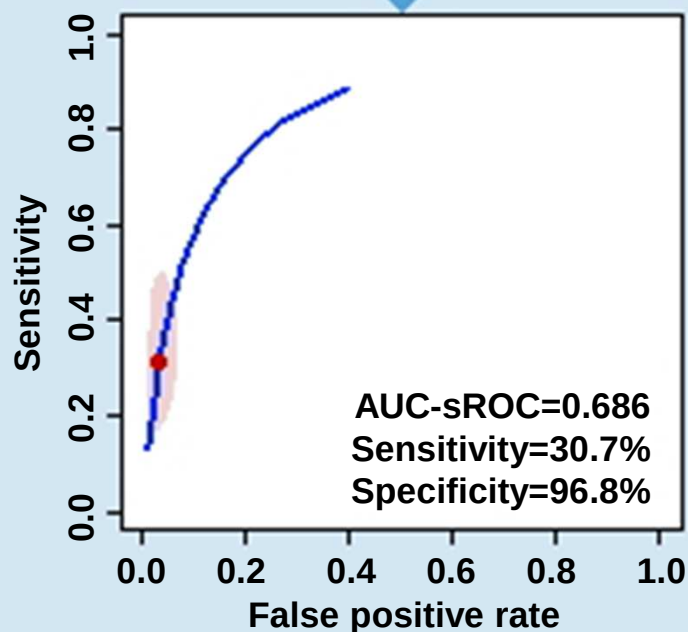
Summary
sensitivity and
specificity

AUC-sROC:
Area under
the summary
ROC curve

Skin tests

20 studies

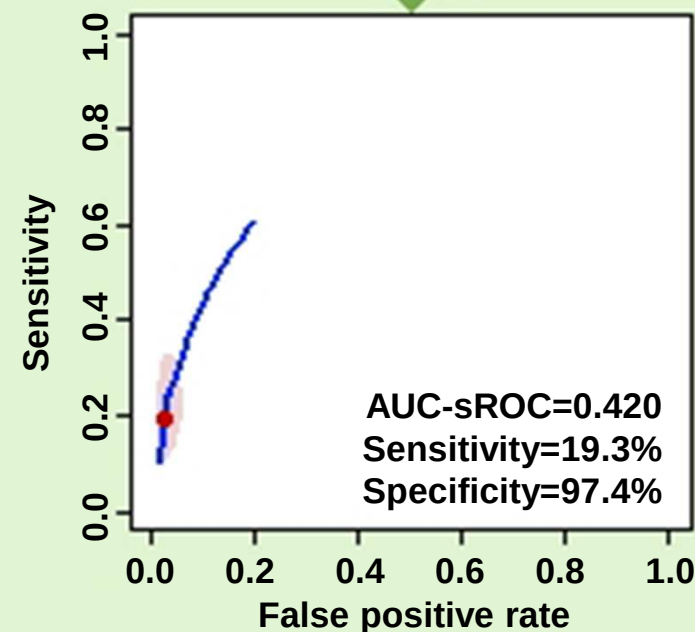
Meta-analysis



Specific IgE
quantification

11 studies

Meta-analysis



Accuracy of Penicillin Allergy Diagnostic Tests: A Systematic Review and Meta-analysis

Bernardo Sousa-Pinto, PhD (bernardo@med.up.pt)*^{1,2,3};

Isabel Tarrio, MD (mimed1301275@med.up.pt)*¹;

Kimberly G Blumenthal, MD (kblumenthal@mgh.harvard.edu)^{4,5};

Luís Araújo, MD (laraujo@med.up.pt)^{2,3};

Luís Filipe Azevedo, PhD (lazevedo@med.up.pt)^{1,2};

Luís Delgado, PhD (ldelgado@med.up.pt)^{2,3};

João Almeida Fonseca, PhD (jfonseca@med.up.pt)^{1,2}

¹ MEDCIDS - Department of Community Medicine, Information and Health Decision Sciences, Faculty of Medicine, University of Porto, Porto, Portugal

² CINTESIS – Center for Health Technology and Services Research, Porto, Portugal

³ Basic and Clinical Immunology Unit, Department of Pathology, Faculty of Medicine, University of Porto, Porto, Portugal

⁴ Division of Rheumatology, Allergy, and Immunology, Department of Medicine, Massachusetts General Hospital, Boston, Massachusetts, United States of America

⁵ Harvard Medical School, Boston, Massachusetts, United States of America

* These authors contributed equally to this work.

Correspondence to: Bernardo Sousa-Pinto; CINTESIS – Center for Health Technology and Services Research, Rua Dr. Plácido da Costa, Porto, Portugal; Telephone number: +351912362153; Fax number: +351225513623; E-mail address: bernardo@med.up.pt

25

26 **Word count (manuscript text):** 3500 words

27 **Tables and figures count:** 4 Tables, 3 Figures; Online Repository: 5 Tables, 1 Figure, 1 App

28

29 **Role of the funding sources:** This work was supported by the Portuguese Ministry for
30 Science, Technology and Higher Education (*Ministério da Ciência, Tecnologia e Ensino*
31 *Superior*), through the Portuguese Foundation for Science and Technology (*Fundação para a*
32 *Ciência e Tecnologia*) [grant number PD/BD/129836/2017]. Dr. Blumenthal receives career
33 development support from the National Institutes of Health (grant NIH K01AI125631), the
34 American Academy of Allergy Asthma and Immunology (AAAAI) Foundation, and the
35 MGH Claflin Distinguished Scholar Award.

36

37 **Conflicts of interests:** Dr. Blumenthal has a clinical decision support tool for inpatient beta-
38 lactam allergy evaluation used at Partners HealthCare Systems and licensed to Persistent
39 Systems, and receives career development support from the NIH K01AI125631, the
40 American Academy of Allergy Asthma and Immunology (AAAAI) Foundation, and the
41 MGH Claflin Distinguished Scholar Award. The content is solely the responsibility of the
42 authors and does not represent the official views of the National Institutes of Health, AAAAI
43 Foundation, nor the MGH.

44 The rest of the authors declare that they have no relevant conflicts of interest.

Abstract

Background: Having a penicillin allergy label associates with higher risk for antibiotic resistance and increased healthcare use.

Objective: To assess the accuracy of skin tests and specific IgE quantification in the diagnostic evaluation of patients reporting a penicillin/ β -lactam allergy.

Methods: We performed a systematic review and diagnostic accuracy meta-analysis, searching on MEDLINE, Scopus and Web of Science. We included studies conducted in patients reporting a penicillin allergy and in whom skin tests and/or specific IgE quantification were performed and compared with drug challenge results. We quantitatively assessed the accuracy of diagnostic tests with bivariate random-effects meta-analyses. Meta-regression and subgroup analyses were performed to explore causes of heterogeneity. Studies' quality was evaluated using QUADAS-2 criteria.

Results: We included 105 primary studies, assessing 31,761 participants. Twenty-seven studies were assessed by bivariate meta-analysis. Skin tests had a summary sensitivity of 30.7% (95%CI=18.9-45.9%) and specificity of 96.8% (95%CI=94.2-98.3%), with a partial area under the summary ROC curve of 0.686 ($I^2=38.2\%$). Similar results were observed for subanalyses restricted to patients reporting non-immediate maculopapular exanthema or urticaria/angioedema. Specific IgE had a summary sensitivity of 19.3% (95%CI=12.0-29.4%) and specificity of 97.4% (95%CI=95.2-98.6%), with a partial area under the summary ROC curve of 0.420 ($I^2=8.5\%$). Projected predictive values mainly reflect the low frequency of true penicillin allergy.

Conclusion: Skin tests and specific IgE quantification appear to have low sensitivity and high specificity. As current evidence is insufficient for assessing the role of these tests in stratifying patients for delabeling, we identified key requirements needed for future studies.

Clinical Implications

In patients reporting a penicillin allergy, skin tests (at least in non-immediate mild reactions) and specific IgE quantification appear to have high specificity and negative predictive value, but low sensitivity.

Capsule summary

This systematic review suggested that, in patients reporting a penicillin allergy, skin tests (at least in non-immediate mild reactions) and specific IgE quantification have high specificity and negative predictive value, but low sensitivity.

Keywords: Diagnostic accuracy; Drug provocation test; IgE quantification; Penicillin allergy; Skin tests

Abbreviations:

- AUC-sROC: Area under the curve of the summary receiver operating characteristic
- CI: Confidence interval
- DPT: Drug provocation test
- NPV: Negative predictive value
- PPV: Positive predictive value
- sIgE: Specific Immunoglobulin E

Introduction

Beta-lactams are a class of broad-spectrum antibiotics that include penicillins, cephalosporins, carbapenems and monobactams. Due to their high effectiveness, safety and low costs, β -lactams are the most frequently prescribed antibiotics in the world, and those for which hypersensitivity reactions are most often reported^{1,2}. Penicillin allergy reporting is particularly frequent – in Europe and North America, allergy to penicillins is reported by up to 10% of the general population and 15% of hospitalized patients^{1,3-5}. However, several studies have shown that there is an overdiagnosis of penicillin allergy, as only 2-10% of individuals who report penicillin allergy are truly allergic^{3,6}.

Without a proper diagnostic approach, many patients with a penicillin allergy label may unnecessarily receive second-line, broad-spectrum antibiotics, leading to an increased risk of drug failure, hospital-acquired infections and higher length of stay⁷⁻¹⁰. These consequences, along with the use of more expensive antibiotics, explain the higher costs of treating patients with a penicillin allergy label^{7,8}, and why penicillin allergy testing programs have been identified as a vital tool to improve antimicrobial stewardship and curb antibiotic resistance^{3,10}.

Patients reporting a penicillin allergy should have their diagnosis confirmed. The International Consensus on Drug Allergy recommends an approach based on skin tests and, in case of negative results, a confirmatory drug provocation test (DPT; “drug challenge”)¹¹. Two types of skin tests are usually done – (i) skin prick tests, which are the first to be performed, consisting in pricking the skin with small amounts of penicillin reagents; and (ii) intradermal tests, which are performed when skin prick tests are negative, involving placing penicillin reagents just underneath the skin¹². In patients reporting immediate reactions, a blood test for quantification of penicillins specific IgE (sIgE) is indicated by European

113 recommendations¹³. If negative results are obtained, the patient undergoes a DPT, receiving a
114 penicillin drug under clinical observation.

115 The DPT is the reference standard for testing patients reporting a penicillin allergy. However,
116 as it is typically performed after negative results with skin tests and/or sIgE quantification,
117 there is still scarce information on the accuracy of these recommended diagnostic tests. In
118 fact, while skin tests appear to have a high negative predictive value (NPV), there are
119 conflicting reports on their sensitivity and positive predictive value (PPV), with wide ranges
120 of values reported. Therefore, this systematic review aims to evaluate the accuracy of the
121 diagnostic tests – skin tests and sIgE quantification – used in the evaluation of patients
122 reporting a penicillin/ β -lactam allergy, and to identify factors that may explain differences in
123 diagnostic parameter estimates.

Methods

Search strategy and inclusion criteria

This study is a systematic review with meta-analysis of diagnostic test accuracy, complying with the PRISMA-DTA Statement¹⁴.

We included original studies conducted in patients in whom penicillin/ β -lactam allergy was reported, and in which the results of DPT had been compared to those of skin tests (skin prick and/or intradermal tests) and/or sIgE quantification (“index tests”). We excluded studies conducted only in patients with non-penicillin β -lactam allergy, and studies evaluating only patients with specific diseases (e.g., patients with rheumatic fever) or occupations (e.g., nurses).

Three electronic bibliographic databases (MEDLINE/PubMed, Web of Science and Scopus) were searched in July 2018. The queries (presented in Table E1 in the Online Repository) included a database-adaptation of the strategy presented by Astin to search for diagnostic performance studies¹⁵. No restrictions based on primary studies’ language or publication date were applied.

After eliminating duplicates, the titles and abstracts of the obtained studies were independently evaluated by two researchers. Subsequently, the full texts of selected studies were independently read by two investigators, who extracted information regarding the year of publication, country, sampling method, study setting (outpatients, inpatients or emergency department), tested drugs/determinants, participants’ number, sex and age distributions, and inclusion criteria regarding initial reactions – initial reactions were assessed based on involved drug classes (whether studies included participants reporting an allergy to any β -lactam or specifically to penicillins), clinical manifestations, and timing. Immediate reactions were defined as initial reactions occurring within the first hour after contact with the

suspected drug, while non-immediate reactions were defined as those occurring after the first hour of drug contact. In addition, we collected information related to each diagnostic index test, including the performed tests, number of participants with positive and negative results and number of participants submitted to DPT. For primary studies assessing patients with any reported β -lactam allergy but presenting discriminated results for those who specifically reported a penicillin allergy, we only retrieved data concerning the latter.

Data were extracted by means of a purposely-built form – we developed a pilot version, which was modified after assessment of the first ten studies. For publications whose full texts were not available, we contacted corresponding authors. Authors were also contacted to provide relevant missing information. Disagreements in study selection or data extraction were solved by consensus.

Study quality was independently assessed by two researchers according to QUADAS-2 judgements on primary studies' risk of bias and applicability¹⁶.

Quantitative Synthesis

For each study, we retrieved the frequency of positive results to each index test. Based on that and on DPT results, we were able to calculate, for each index test, the number of true positives, true negatives, false positives and/or false negatives.

For tests for which we had more than four primary studies simultaneously presenting results for true positives, true negatives, false positives and false negatives, we performed meta-analyses of diagnostic test accuracy using a bivariate random-effects approach¹⁷. Bivariate random-effects meta-analyses allow for the calculation of summary properties of diagnostic tests (namely, sensitivity, specificity and area under the ROC curve), by pooling the results of different diagnostic accuracy studies, accommodating between-study heterogeneity. This method (which is equivalent to the hierarchical summary receiver operating characteristic

model) has the advantage of jointly estimating sensitivity and specificity, accounting for correlation between them¹⁸. Performing bivariate meta-analysis, we calculated areas under the summary curve (AUC-sROC), and summary sensitivities, specificities, likelihood ratios and diagnostic odds ratios. A correlation coefficient (r) >0.6 between sensitivity and false positive rate was considered to represent considerable threshold effect (i.e., negative correlation between sensitivity and specificity possibility accounted by variability in positivity criteria)¹⁹. Heterogeneity (i.e., between-studies variability beyond expected due to chance or random sampling errors alone) was assessed using an I^2 statistic estimate for bivariate meta-analysis²⁰. An $I^2 > 50\%$ was considered to represent substantial heterogeneity.

Based on the obtained estimates of pooled sensitivity and specificity, and considering a prevalence of true penicillin allergy ranging from 0-10%, we estimated the PPV and NPV of index tests. We performed univariate sensitivity analyses, to obtain PPV and NPV estimates for different prevalence values. Probabilistic sensitivity analyses based on Monte Carlo simulations were additionally performed to account for uncertainty in pooled sensitivity and specificity estimates.

Whenever there were sufficient primary studies, sensitivity analyses were performed to address the possibility of selection bias and/or to account for possible clinically relevant sources of heterogeneity. Such sensitivity analyses restricted to (i) studies in which at least 50% of both patients with positive and negative results on index tests received a DPT; (ii) studies assessing a mix of participants with immediate and non-immediate reactions (excluding only those with very severe index reactions), and testing all with index tests and DPT; and (iii) studies assessing consecutive samples of patients specifically reporting a history of penicillin allergy and – in the case of skin tests – testing for the determinants recommended by European Network of Drug Allergy (penicillin polylysine, minor determinant mixture, benzylpenicillin and an aminopenicillin); and (iv) studies assessing

particular clinical presentations or age groups. In addition, we performed bivariate meta-regression and applied the likelihood-ratio test to identify moderator variables (i.e., variables explaining heterogeneity²¹). Meta-regression tested the following covariates: year of publication, proportion of female subjects, participants' age group (children, adults, or adults and children), study region (Europe, North America, Middle East, or Australia and Far East), setting, sampling method (consecutive sample including all participants with reported allergy, consecutive sample including all tested participants, or convenience sample), number of tested drugs/determinants (and whether the index test assessed at least an aminopenicillin, a cephalosporin, and/or the suspected drug), reported allergy (to overall β -lactams, penicillins or aminopenicillins only) and timing of the initial reaction.

Not all studies allowed for obtaining, for each index test, the number of true positives, true negatives, false positives and/or false negatives. Therefore, in addition to bivariate random-effects meta-analysis, we performed random-effects meta-analysis of individual diagnostic parameter estimates (methods fully described in the Online Repository). The latter corresponds to a less robust approach in the context of meta-analysis of diagnostic studies, as it does not account for sensitivity-specificity correlation¹⁷.

Meta-analyses were performed using software R (version 3.5.0), using mada, metafor and glmulti packages. Simulations of predictive values estimates were obtained using TreeAge Pro software. Significance level was defined at $p < 0.05$.

Results

Our search yielded 3952 publications, of which 1387 were duplicates; 2419 studies were excluded based on application of eligibility criteria (Figure 1). We could not obtain the full text of 24 publications despite contacting corresponding authors – 12 of these studies had been written in languages other than English and 16 had been published before 2000. Of the 122 articles meeting the eligibility criteria, 17 were excluded as some participants had also been assessed (receiving the same tests) by studies assessing larger samples. Therefore, a total of 105 studies were included in this systematic review^{13,22-125}. For studies in which some degree of participant overlap with other study(ies) occurred (Table E2)^{36,37,40-42,59,62,69,72,83,106,109}, we presented only unique information from each study (e.g., results from diagnostic tests performed in one study that had not been performed in the other study assessing the same participants), so that no participant has been included more than once.

Of the 105 included primary studies, 27 were included in bivariate random-effects meta-analyses²²⁻⁴⁸ (summary of information in Table I and Table E3). The remainder were only quantitatively assessed by means of individual diagnostic parameter estimates meta-analysis, whose description and results can be consulted in the online Supplement (Tables E2-E4). We developed an interactive app (<https://penaccuracy.med.up.pt>; Online Repository), which also presents results from all analyses, as well as results for other tests whose studies were retrieved by our query.

The meta-analysis assessing the diagnostic accuracy of skin tests comprised 20 studies^{22-29,31,32,34,35,38,40,41,43-46,48} (Figure 2A; Table II). These studies had been mostly performed in Europe^{24,26-29,32,38,40,41,45,48} ($n=11$) and North America^{22,23,25,31,43,44} ($n=6$), and mostly assessed outpatients^{23,24,27-29,34,38,40,41,43-46,48} ($n=14$). Adults and children were exclusively assessed by five studies each^{25,28,29,31,35,38,40,41,44,48}, while nine studies assessed patients of all age groups²²⁻

^{24,26,27,32,34,43,46}. In nine studies^{26,27,32,34,35,38,45,46,48}, a consecutive sample was assessed, while other nine reported a convenience sample^{23,24,28,29,31,40,41,43,44}. We obtained a partial AUC-sROC of 0.686, with a summary sensitivity of 30.7% (95%CI=18.9-45.8%) and a specificity of 96.8% (95%CI=94.2-98.3%). No considerable threshold effect was observed ($r=0.363$). Similar results were obtained in sensitivity analyses addressing selection bias or possible sources of confounding, albeit with less precise estimates (Table II). Heterogeneity was found to be moderate ($I^2=38.2\%$), with the year of publication and performance in Middle East found to be moderators of heterogeneity (Table III). Predictive values estimates are depicted in Figure 3.

For non-immediate reactions (and contrary to immediate reactions), we were able to perform a bivariate random effects meta-analysis on the diagnostic accuracy of skin tests. This analysis was based on seven studies^{26,28,32,35,40,41,46} (Figure 2B; Table II), with almost all participants reporting either maculopapular exanthema or urticaria/angioedema. The partial AUC-sROC was of 0.396 ($r=0.335$), the summary sensitivity was of 44.2% (95%CI=14.2-79.1%), the specificity was of 95.1% (95%CI=88.3-98.0%), and an I^2 of 22.2% was obtained.

For skin prick tests, we were only able to estimate pooled single diagnostic measures (Table E4). Regarding intradermal tests, we performed a bivariate diagnostic random effects meta-analysis (Figure 2C; Table II) based on ten studies^{24,26,28,32,34,38,40,42,46,48} – we observed a partial AUC-sROC of 0.582, a summary sensitivity of 29.9% (95%CI=15.0-50.8%) and a specificity of 95.1% (95%CI=88.1-98.0%). A borderline threshold effect was observed ($r=0.607$) along with severe heterogeneity ($I^2=54.8\%$). Performance in Europe and testing the suspected culprit drug were identified as moderators of heterogeneity (Table III). Among patients reporting non-immediate reactions (largely maculopapular exanthema or urticaria/angioedema)^{26,28,32,40,42,46}, a partial AUC-sROC of 0.404 was obtained (Figure 2D; Table II), with a summary sensitivity of 43.1% (95%CI=12.8-79.5%), and a specificity of

95.1% (95%CI=90.9-97.5%). No relevant threshold effect was observed ($r=-0.194$), and moderate heterogeneity was observed ($I^2=35.3\%$).

Eleven studies were included in the meta-analysis assessing the accuracy of specific

IgE^{24,28,30,33,36-39,42,47,48}. The studies had been mostly performed in Europe ($n=10$)

^{24,28,30,33,36,38,39,42,47,48}, and among outpatients ($n=9$)^{24,28,33,36-39,42,47,48}. Four studies assessed

adults^{28,37,38,47}, while three assessed children^{36,42,48} and other four patients of all age

groups^{24,30,33,39}. The sampling method was consecutive for six studies^{33,36,38,39,42,48} and of

convenience for five studies^{24,28,30,37,47}. We obtained a partial AUC-sROC of 0.420, but

considerable threshold effect was observed ($r=1.000$) (Figure 2E; Table II). This model

resulted in a summary sensitivity of 19.3% (95%CI=12.0-29.4%) and a specificity of 97.4%

(95%CI=95.2-98.6%). Similar results were obtained in a sensitivity analysis (Table II).

Heterogeneity was found to be low ($I^2=8.5\%$), with participants' age identified as a

moderator of heterogeneity (Table III). Predictive values estimates are depicted in Figure 3.

Among patients reporting immediate reactions^{28,33,42,47}, we obtained a partial AUC-sROC of

0.262, with a summary sensitivity of 26.2% (95%CI=18.7-35.3%) and a summary specificity

of 92.1% (95%CI=64.7-98.7%).

Studies' quality was assessed using QUADAS-2 tool¹⁶ – Figure E1 depicts the risk-of-bias

graph, while Table E5 presents the risk-of-bias summary. The main quality concerns of the

assessed studies are related to patients' flow (in most studies, not all tested patients received a

DPT) and to "patient selection domain" (reflecting concerns on sample representativity). Of

all included primary studies, only two (1.9%) were judged as "low risk" on all bias-related

domains^{42,46}, while 12 (11.4%) were judged as "low concern" on all applicability

domains^{24,27,32,34,39,69,77,102,111,114,120,123}.

Discussion

In this systematic review, we assessed the diagnostic accuracy of skin tests and sIgE quantification for patients reporting a penicillin allergy. In all performed analyses, both skin tests and sIgE assessment were found to have a high specificity but a sensitivity to predict hypersensitivity reactions under 50%. Based on those obtained values, and on a plausible range of penicillin allergy prevalences, these tests were projected to have a high NPV, but a poor PPV.

Importantly, predictive values are influenced by the prevalence of the studied condition, in this case corresponding to the pre-test probability of having a true penicillin allergy. For most primary studies we do not have individual participant-level information on the presentation/severity of the index reaction (nor were defined homogenous inclusion criteria regarding the latter). However, we have information that most participants reported a mild index reaction, reflecting the fact that most included studies consist of a description of a mostly low-risk cohort of patients undergoing a diagnostic workup for penicillin allergy. This may explain why the PPV projected by our models were lower than that estimated by a recent review, that identified a PPV of 81%¹²⁶. However, that latter review included only studies with detailed characterizations of each participant's index reaction (and tested with the culprit drug), who, as such, had a higher pre-test probability of having a true/severe allergic reaction.

Although we were not able to estimate the diagnostic accuracy of skin tests and specific IgE for each clinical presentation or risk group¹²⁷, we were able to estimate the sensitivity, specificity and predictive values for penicillin skin tests in patients reporting a delayed maculopapular exanthema or urticaria/angioedema. Therefore, our results apply largely to the most common patients encountered, particularly in North America, namely low risk patients with rashes and/or swelling.

This study has some limitations, particularly associated with the included primary studies (Table IV lists the limitations of current evidence and key requirements needed for future studies). Firstly, only a minority of them performed DPT in participants with positive index results. As a result, our bivariate diagnostic meta-analysis included fewer primary studies than the corresponding individual diagnostic measures meta-analyses. Nevertheless, we obtained comparable results for the analyses in which the two meta-analytic methods were performed. Another limitation concerns the possibility of selection bias – in several studies, not all participants who were subject to an index test underwent a DPT. This is particularly true for patients with positive index tests, with possible implications for sensitivity estimation. In fact, there were studies which only challenged patients with mild or borderline positive skin tests^{23,43}. To account for this possible bias, we performed sensitivity analyses restricted to (i) studies with at least 50% patients with positive results on the index test performing a DPT, and (ii) studies assessing a mix of participants with immediate and non-immediate reactions (excluding only those with very severe index reactions), and testing all with index tests and DPT. Those sensitivity analyses yielded similar results, although with less precision due to the lower number of primary studies. Nevertheless, even in those studies, most participants reported a clinical history compatible with a mild index reaction. In addition, in several studies, more than one index test was performed, with DPT being done only in those patients who had negative results in all index tests – one particularly common example is the sequential execution of skin prick and intradermal tests prior to DPT. Therefore, our pooled estimates do not strictly measure the “pure” accuracy of each index test. Nevertheless, for individual diagnostic measures meta-analyses, similar results are obtained when analyzing only studies performing DPT in all patients with negative and/or positive results in the corresponding index tests. We were not able to assess the accuracy of skin prick tests by bivariate random-effects meta-analysis. Nevertheless, results of individual

diagnostic measures meta-analyses point to a lower sensitivity of skin prick tests (14.6%) when compared to intradermal tests (29.9%) or with the performance of both skin prick and intradermal tests (30.7%) – the performance of skin prick tests without intradermal tests can result in non-identifying a larger proportion of allergic patients.

Other important limitations concern the low quality of most included primary studies and – for some studies – lack of discriminated information regarding the timing of initial reactions (on the other hand, patient-reported initial reaction timing may be subject to recall bias). Finally, while DPT was defined as the reference standard, its accuracy is not 100%, and protocol variability might affect the primary studies results.

This study has also several strengths. In particular, we combined two different methods – bivariate diagnostic and individual measure meta-analyses – to estimate summary measures, obtaining comparable results. To minimize the impact of publication bias, we searched three different electronic bibliographic databases without applying exclusion criteria based on the date or language of publication. In addition, we performed meta-regression and sensitivity and subgroup analyses to explore sources of heterogeneity and hint at variables possibly associated with differences in penicillin allergy tests diagnostic accuracy. Heterogeneity was low/moderate in most bivariate random-effects meta-analyses, with substantial heterogeneity observed in only two analyses (in the remaining ten, heterogeneity was found to be low or moderate). The year of publication, the region, and whether the suspected culprit drug was tested were identified as the main variables explaining heterogeneity. A more recent year was associated with a decreased summary sensitivity of skin tests – this may indicate an improvement of DPT procedures, with more recent studies following/contacting patients for a longer period of time after performing a DPT, and not simply a worsening accuracy of skin tests¹²⁸. Regarding the region, recent surveys found important regional differences in the diagnostic workup of penicillin allergy^{129,130}. The included primary studies presented several

other differences in their diagnostic approaches (e.g., regarding testing for minor determinant mixture, whose individual role we were not able to adequately assess). However, these differences were frequently not identified as moderators of heterogeneity, and similar results were obtained with the sensitivity analysis restricted to studies with a “more homogeneous approach” (which included testing for penicillin polylysine, minor determinant mixture, benzylpenicillin and an aminopenicillin).

The topic of this systematic review is particularly relevant and timely since penicillin allergy delabeling is now encouraged as a tool of antibiotic stewardship and quality improvement by several organizations – including the American Academy of Allergy, Asthma & Immunology, the European Academy of Allergy and Clinical Immunology, the Centers for Disease Control and Prevention, and the Infectious Diseases Society of America. Overall, our results suggest that, at least in patients reporting a delayed mild index reaction, skin tests may have a high specificity and negative predictive value, but a low positive predictive value and sensitivity in identifying patients who will develop a hypersensitivity reaction when exposed to penicillins. If that trend is confirmed by subsequent studies, direct performance of DPT may be a feasible and safe option at least in some of those patients – in fact, recent studies have shown that patients whose clinical history is non-severe and imprecise or not compatible with a true reaction may possibly safely undergo DPT without prior skin tests, particularly if the reaction did not occur within one hour of penicillin ingestion^{131,132}. In addition, our systematic review points to the need for further studies minimizing the risk of selection bias and providing, if not individual-level participant data, at least results by type of index reaction clinical presentation. Such evidence would help in patient stratification, so that penicillin allergy delabeling could be safely performed in the most accurate way for each patient.

387 **Author contributions:** BSP and IT participated in search, selection, quality assessment and
388 data extraction from primary studies, as well as in data analysis and manuscript writing.
389 KGB, LA, LFA, LD and JAF participated in data analysis, manuscript writing and critical
390 review of the manuscript.

391

Journal Pre-proof

References

1. Macy E, Poon KYT. Self-reported antibiotic allergy incidence and prevalence: age and sex effects. *Am J Med* 2009; 122(8): 778 e1–7.
2. Sousa-Pinto B, Fonseca JA, Gomes ER. Frequency of self-reported drug allergy: A systematic review and meta-analysis with meta-regression. *Ann Allergy Asthma Immunol* 2017; 119(4): 362–73 e2.
3. Blumenthal KG, Peter JG, Trubiano JA, Phillips EJ. Antibiotic allergy. *Lancet* 2019; 393(10167): 183–98.
4. Lee CE, Zembower TR, Fotis MA, Postelnick MJ, Greenberger PA, Peterson LR, et al. The incidence of antimicrobial allergies in hospitalized patients: implications regarding prescribing patterns and emerging bacterial resistance. *Arch Intern Med* 2000; 160(18): 2819–22.
5. Gomes E, Cardoso MF, Praca F, Gomes L, Marino E, Demoly P. Self-reported drug allergy in a general adult Portuguese population. *Clin Exp Allergy* 2004; 34(10): 1597–601.
6. 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): 557–62.
7. Mattingly TJ, 2nd, Fulton A, Lumish RA, Williams AMC, Yoon S, Yuen M, et al. The Cost of Self-Reported Penicillin Allergy: A Systematic Review. *J Allergy Clin Immunol Pract* 2018; 6(5): 1649–54 e4.
8. Sousa-Pinto B, Cardoso-Fernandes A, Araujo L, Fonseca JA, Freitas A, Delgado L. Clinical and economic burden of hospitalizations with registration of penicillin allergy. *Ann Allergy Asthma Immunol* 2018; 120(2): 190–4 e2.

9. Blumenthal KG, Lu N, Zhang Y, Li Y, Walensky RP, Choi HK. Risk of meticillin resistant *Staphylococcus aureus* and *Clostridium difficile* in patients with a documented penicillin allergy: population based matched cohort study. *BMJ* 2018; 361: k2400.
10. Shenoy ES, Macy E, Rowe T, Blumenthal KG. Evaluation and Management of Penicillin Allergy: A Review. *JAMA* 2019; 321(2): 188–99.
11. Demoly P, Adkinson NF, Brockow K, Castells M, Chiriac AM, Greenberger PA, et al. International Consensus on drug allergy. *Allergy* 2014; 69(4): 420–37.
12. Blumenthal KG, Shenoy ES. Am I allergic to penicillin? *JAMA* 2019; 321(2):216.
13. Fontaine C, Mayorga C, Bousquet J, Arnoux B, Torres MJ, Blanca M, et al. Relevance of the determination of serum-specific IgE antibodies in the diagnosis of immediate β -lactam allergy. *Allergy* 2006; 62(1): 47-52.
14. McInnes MDF, Moher D, Thombs BD, McGrath TA, Bossuyt PM, PRISMA-DTA Group. Preferred Reporting Items for a Systematic Review and Meta-analysis of Diagnostic Test Accuracy Studies: The PRISMA-DTA Statement. *JAMA* 2018; 319(4): 388–96.
15. Astin MP, Brazzelli MG, Fraser CM, Counsell CE, Needham G, Grimshaw JM. Developing a sensitive search strategy in MEDLINE to retrieve studies on assessment of the diagnostic performance of imaging techniques. *Radiology* 2008; 247(2): 365–73.
16. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 2011; 155(8): 529–36.
17. Reitsma JB, Glas AS, Rutjes AW, Scholten RJ, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *J Clin Epidemiol* 2005; 58(10): 982–90.
18. Harbord RM, Deeks JJ, Egger M, Whiting P, Sterne JA. A unification of models for meta-analysis of diagnostic accuracy studies. *Biostatistics* 2007; 8(2): 239–51.

19. Deville WL, Buntinx F, Bouter LM, Montori VM, de Vet HC, van der Windt DA, et al. Conducting systematic reviews of diagnostic studies: didactic guidelines. *BMC Med Res Methodol* 2002; 2: 9.
20. Zhou Y, Dendukuri N. Statistics for quantifying heterogeneity in univariate and bivariate meta-analyses of binary data: the case of meta-analyses of diagnostic accuracy. *Stat Med* 2014; 33(16): 2701–17.
21. Lipsey MW. Those confounded moderators in meta-analysis: Good, bad, and ugly. *The Annals of the American Academy of Political and Social Science* 2003; 587(1): 69–81.
22. Levine BB, Zolov DM. Prediction of penicillin allergy by immunological tests. *J Allergy* 1969; 43(4): 231–44.
23. Warrington RJ, Simons FER, Ho HW, Gorski BA. Diagnosis of penicillin allergy by skin testing: The manitoba experience. *CMAJ* 1978; 118(7): 787–91.
24. Blanca M, Vega JM, Garcia J, Carmona MJ, Terados S, Avila MJ, et al. Allergy to penicillin with good tolerance to other penicillins; study of the incidence in subjects allergic to betalactams. *Clin Exp Allergy* 1990; 20(5): 475–81.
25. Sogn DD, Evans R, III, Shepherd GM, Casale TB, Condemi J, Greenberger PA, 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. *Arch Intern Med* 1992; 152(5): 1025–32.
26. Romano A, Di Fonso M, Papa G, Pietrantonio F, Federico F, Fabrizi G, et al. Evaluation of adverse cutaneous reactions to aminopenicillins with emphasis on those manifested by maculopapular rashes. *Allergy* 1995; 50(2): 113–8.
27. Sastre J, Quijano LD, Novalbos A, Hernandez G, Cuesta J, de las Heras M, et al. Clinical cross-reactivity between amoxicillin and cephadroxil in patients allergic to amoxicillin and with good tolerance of penicillin. *Allergy* 1996; 51(6): 383–86.

- 465 28. Cederbrant K, Stejskal V, Broman P, Lindkvist L, Sundell K. In vitro lymphocyte
466 proliferation in the diagnosis of allergy to phenoxymethylpenicillin. *Allergy* 1998; 53(12):
467 1155–61.
- 468 29. Loza Cortina C. Adverse drug reactions in area I of Asturias. *An Esp Pediatr* 1998;
469 49(4): 359–63.
- 470 30. Blanca M, Mayorga C, Torres MJ, Reche M, Moya MC, Rodriguez JL, et al. Clinical
471 evaluation of Pharmacia CAP System (TM) RAST FEIA amoxicilloyl and benzylpenicilloyl
472 in patients with penicillin allergy. *Allergy* 2001; 56(9): 862–70.
- 473 31. Forrest DM, Schellenberg RR, Thien VVS, King S, Anis AH, Dodek PM.
474 Introduction of a practice guideline for penicillin skin testing improves the appropriateness of
475 antibiotic therapy. *Clin Infect Dis* 2001; 32(12): 1685–90.
- 476 32. Romano A, Viola M, Mondino C, Pettinato R, Di Fonso M, Papa G, et al. Diagnosing
477 nonimmediate reactions to penicillins by in vivo tests. *Int Arch Allergy Immunol* 2002;
478 129(2): 169–74.
- 479 33. Brusca I, Corrao S, Sceusa G, Barrale M, Dragotto G, Gristina T, et al. Flow
480 cytometric basophil activation test in patients with allergic or pseudoallergic reactions to
481 betalactamic antibiotics and non steroidal and inflammatory drugs. *Rivista Italiana della*
482 *Medicina di Laboratorio* 2007; 3(3): 196–202.
- 483 34. Goldberg A, Confino-Cohen R. Skin testing and oral penicillin challenge in patients
484 with a history of remote penicillin allergy. *Ann Allergy Asthma Immunol* 2008; 100(1): 37–
485 43.
- 486 35. Tantikul C, Dhana N, Jongjarearnprasert K, Visitsunthorn N, Vichyanond P,
487 Jirapongsananuruk O. The Utility of the World Health Organization-The Uppsala Monitoring
488 Centre (WHO-UMC) System for the Assessment of Adverse Drug Reactions in Hospitalized
489 Children. *Asian Pac J Allergy Immunol* 2008; 26(2-3): 77–82.

- 490 36. Caubet JC, Kaiser L, Lemaitre B, Fellay B, Gervaix A, Eigenmann PA. The role of
491 penicillin in benign skin rashes in childhood: A prospective study based on drug rechallenge.
492 J Allergy Clin Immunol 2010; 127(1): 218–22.
- 493 37. Macy E, Goldberg B, Poon KY. Use of commercial anti-penicillin IgE fluorometric
494 enzyme immunoassays to diagnose penicillin allergy. Ann Allergy Asthma Immunol 2010;
495 105(2): 136–41.
- 496 38. Richter AG, Wong G, Goddard S, Heslegrave J, Derbridge C, Srivastava S, et al.
497 Retrospective case series analysis of penicillin allergy testing in a UK specialist regional
498 allergy clinic. J Clin Pathol 2011; 64(11): 1014–8.
- 499 39. Kopac P, Zidarn M, Kosnik M. Epidemiology of hypersensitivity reactions to
500 penicillin in Slovenia. Acta Dermatovenerol Alp Pannonica Adriat 2012; 21(4): 5–7.
- 501 40. Caubet JC, Frossard C, Fellay B, Eigenmann PA. Skin tests and in vitro allergy tests
502 have a poor diagnostic value for benign skin rashes due to β -lactams in children. Pediatr
503 Allergy Immunol 2014; 26(1): 80–2.
- 504 41. Barni S, Mori F, Sarti L, Pucci N, Rossi EM, de Martino M, et al. Utility of skin
505 testing in children with a history of non-immediate reactions to amoxicillin. Clin Exp Allergy
506 2015; 45(9): 1472–4.
- 507 42. Mori F, Cianferoni A, Barni S, Pucci N, Rossi ME, Novembre E. Amoxicillin Allergy
508 in Children: Five-Day Drug Provocation Test in the Diagnosis of Nonimmediate Reactions. J
509 Allergy Clin Immunol Pract 2015; 3(3): 375–+.
- 510 43. Rosenfield L, Kalicinsky C, Warrington R. A retrospective comparison of false
511 negative skin test rates in penicillin allergy, using pencilloyl-poly-lysine and minor
512 determinants or Penicillin G, followed by open challenge. Allergy Asthma Clin Immunol
513 2015; 11: 34.

- 514 44. Mawhirt SL, Fonacier LS, Calixte R, Davis-Lorton M, Aquino MR. Skin testing and
515 drug challenge outcomes in antibiotic-allergic patients with immediate-type hypersensitivity.
516 *Ann Allergy Asthma Immunol* 2016; 118(1): 73–9.
- 517 45. Meng J, Thursfield D, Lukawska JJ. Allergy test outcomes in patients self-reported as
518 having penicillin allergy: Two-year experience. *Ann Allergy Asthma Immunol* 2016; 117(3):
519 273–9.
- 520 46. Confino-Cohen R, Rosman Y, Meir-Shafir K, Stauber T, Lachover-Roth I, Hershko
521 A, et al. Oral Challenge without Skin Testing Safely Excludes Clinically Significant Delayed-
522 Onset Penicillin Hypersensitivity. *J Allergy Clin Immunol Pract* 2017; 5(3): 669–75.
- 523 47. Tannert LK, Falkencrone S, Mortz CG, Bindslev-Jensen C, Skov PS. Is a positive
524 intracutaneous test induced by penicillin mediated by histamine? A cutaneous microdialysis
525 study in penicillin-allergic patients. *Clin Transl Allergy* 2017; 7: 40.
- 526 48. Ibanez MD, Del Rio PR, Lasa EM, Joral A, Ruiz-Hornillos J, Munoz C, et al.
527 Prospective assessment of diagnostic tests for pediatric penicillin allergy, from clinical
528 history to challenge tests. *Ann Allergy Asthma Immunol* 2018; 121(2):235-244.
- 529 49. Brown BC, Price EV, Moore MB, Jr. Penicilloyl-Polylysine as an Intradermal Test of
530 Penicillin Sensitivity. *JAMA* 1964; 189(8): 599–604.
- 531 50. Resnik SS, Shelley WB. Penicillin hypersensitivity: detection by basophil response to
532 challenge. *J Invest Dermatol* 1965; 45(4): 269–74.
- 533 51. Wide L, Juhlin L. Detection of penicillin allergy of the immediate type by
534 radioimmunoassay of reagins (IgE) to penicilloyl conjugates. *Clin Exp Allergy* 1971; 1(2):
535 171–7.
- 536 52. Basomba A, Villalmanzo IG, Campos A, Pelaez A, Berglund A. IgE antibodies
537 against penicillin as determined by Phadebas RAST. *Clin Allergy* 1979; 9(5): 515–25.

- 538 53. Sullivan TJ, Wedner HJ, Shatz GS, Yecies LD, Parker CW. Skin testing to detect
539 penicillin allergy. *J Allergy Clin Immunol* 1981; 68(3): 171–80.
- 540 54. VanArsdel Jr PP, Martonick GJ, Johnson LE, Sprenger JD, Altman LC, Henderson Jr
541 WR. The value of skin testing for penicillin allergy diagnosis. *West J Med* 1986; 144(3):
542 311–4.
- 543 55. Graff-Lonnevig V, Hedlin G, Lindfors A. Penicillin allergy- A rare paediatric
544 condition? *Arch Dis Child* 1988; 63(11): 1342–6.
- 545 56. Surtees SJ, Stockton MG, Gietzen TW. Allergy to penicillin: fable or fact? *BMJ* 1991;
546 302(6784): 1051–2.
- 547 57. Gadde J, Spence M, Wheeler B, Adkinson NF. Clinical-experience with penicillin
548 skin testing in a large inner-city STD clinic. *JAMA* 1993; 270(20): 2456–63.
- 549 58. Juchet A, Rancé F, Brémont F, Dutau G. Investigation of beta-lactam allergy in 45
550 children. *Revue française d'allergologie et d'immunologie clinique* 1994; 34(5): 369–75.
- 551 59. Macy E, Richter PK, Falkoff R, Zeiger R. Skin testing with penicilloate and penilloate
552 prepared by an improved method: Amoxicillin oral challenge in patients with negative skin
553 test responses to penicillin reagents. *J Allergy Clin Immunol* 1997; 100(5): 586–91.
- 554 60. Hervé M, De la Rocque F, Bouhanna A, Albengres E, Reinert P. Exploration of 112
555 children suspected of amoxicillin allergy. Indications and efficacy of oral provocation test.
556 *Arch Pediatr* 1998; 5(5): 503–9.
- 557 61. Pichichero ME, Pichichero DM. Diagnosis of penicillin, amoxicillin, and
558 cephalosporin allergy: Reliability of examination assessed by skin testing and oral challenge.
559 *J Pediatr* 1998; 132(1): 137–43.
- 560 62. Lebel B, Messaad D, Kvedariene V, Rongier M, Bousquet J, Demoly P. Cysteinyl-
561 leukotriene release test (CAST) in the diagnosis of immediate drug reactions. *Allergy* 2001;
562 56(7): 688–92.

- 563 63. López Tiro JJ, Orea Solano M, Flores Sandoval G, Gómez Vera J. Skin tests with
564 high and minor determinants in patients with doubtful allergy to penicillin. *Rev Alerg Mex*
565 2001; 48(3): 80–7.
- 566 64. Novalbos A, Sastre J, Cuesta J, De Las Heras M, Lluch-Bernal M, Bombín C, et al.
567 Lack of allergic cross-reactivity to cephalosporins among patients allergic to penicillins. *Clin*
568 *Exp Allergy* 2001; 31(3): 438–43.
- 569 65. Solensky R, Earl HS, Gruchalla RS. Lack of penicillin resensitization in patients with
570 a history of penicillin allergy after receiving repeated penicillin courses. *Arch Intern Med*
571 2002; 162(7): 822–6.
- 572 66. Torres MJ, Mayorga C, Leyva L, Guzman AE, Cornejo-García JA, Juarez C, et al.
573 Controlled administration of penicillin to patients with a positive history but negative skin
574 and specific serum IgE tests. *Clin Exp Allergy* 2002; 32(2): 270–6.
- 575 67. Borch JE, Andersen KE, Bindslev-Jensen C. The prevalence of suspected and
576 challenge-verified penicillin allergy in a university hospital population. *Basic Clin Pharmacol*
577 *Toxicol* 2006; 98(4): 357–62.
- 578 68. Wong BBL, Keith PK, Wasserman S. Clinical history as a predictor of penicillin skin
579 test outcome. *Ann Allergy Asthma Immunol* 2006; 97(2): 169–74.
- 580 69. Bousquet PJ, Kvedariene V, Co-Minh HB, Martins P, Rongier M, Arnoux B, et al.
581 Clinical presentation and time course in hypersensitivity reactions to β -lactams. *Allergy*
582 2007; 62(8): 872–6.
- 583 70. del Real GA, Rose ME, Hammel J, Gordon SM, Arroliga ME. The penicillin skin test
584 has a high negative predictive value and helps to modify the use of antibiotics in patients with
585 history of beta-lactam allergy. *J Allergy Clin Immunol* 2007; 117(2): S224–S.
- 586 71. Rebelo Gomes E, Fonseca J, Araujo L, Demoly P. Drug allergy claims in children:
587 From self-reporting to confirmed diagnosis. *Clin Exp Allergy* 2007; 38(1): 191–8.

- 588 72. Bousquet PJ, Pipet A, Bousquet-Rouanet L, Demoly P. Oral challenges are needed in
589 the diagnosis of β -lactam hypersensitivity. *Clin Exp Allergy* 2008; 38(1): 185–90.
- 590 73. Nolan RC, Puy R, Deckert K, O'Hehir RE, Douglass JA. Experience with a new
591 commercial skin testing kit to identify IgE-mediated penicillin allergy. *Intern Med J* 2008;
592 38(5): 357–61.
- 593 74. Padial A, Antunez C, Blanca-Lopez N, Fernandez TD, Cornejo-Garcia JA, Mayorga
594 C, et al. Non-immediate reactions to beta-lactams: diagnostic value of skin testing and drug
595 provocation test. *Clin Exp Allergy* 2008; 38(5): 822–8.
- 596 75. Rodriguez-Alvarez M, Santos-Magadan S, Rodriguez-Jimenez B, de Arellano IRR,
597 Vazquez-Cortes S, Martinez-Cocera C. Reproducibility of delayed-type reactions to
598 betalactams. *Allergol Immunopathol (Madr)* 2008; 36(4): 201–4.
- 599 76. Hershkovich J, Broides A, Kirjner L, Smith H, Gorodischer R. Beta lactam allergy
600 and resensitization in children with suspected beta lactam allergy. *Clin Exp Allergy* 2009;
601 39(5): 726–30.
- 602 77. Silva R, Cruz L, Botelho C, Cadinha S, Castro E, Rodrigues J, et al. Work up of
603 patients with history of beta-lactam hypersensitivity. *Allergol Immunopathol (Madr)* 2009;
604 37(4): 193–7.
- 605 78. Chambel M, Martins P, Silva I, Palma-Carlos S, Romeira AM, Pinto PL. Drug
606 provocation tests to betalactam antibiotics: experience in a paediatric setting. *Allergol*
607 *Immunopathol (Madr)* 2010; 38(6): 300–6.
- 608 79. Holm A, Mosbech H. Challenge Test Results in Patients With Suspected Penicillin
609 Allergy, but No Specific IgE. *Allergy Asthma Immunol Res* 2011; 3(2): 118–22.
- 610 80. Kamboj S, Yousef E, McGeady S, Hossain J. The prevalence of antibiotic skin test
611 reactivity in a pediatric population. *Allergy Asthma Proc* 2011; 32(2): 99–105.

- 612 81. Moral L, Garde J, Toral T, Fuentes MJ, Marco N. Short protocol for the study of
613 paediatric patients with suspected betalactam antibiotic hypersensitivity and low risk criteria.
614 Allergol Immunopathol (Madr) 2011; 39(6): 337–41.
- 615 82. Na HR, Lee JM, Jung JW, Lee SY. Usefulness of drug provocation tests in children
616 with a history of adverse drug reaction. Korean J Pediatr 2011; 54(7): 304–9.
- 617 83. Ponvert C, Perrin Y, Bados-Albiero A, Le Bourgeois M, Karila C, Delacourt C, et al.
618 Allergy to betalactam antibiotics in children: Results of a 20-year study based on clinical
619 history, skin and challenge tests. Pediatr Allergy Immunol 2011; 22(4): 411–8.
- 620 84. Seitz CS, Brocker EB, Trautmann A. Diagnosis of drug hypersensitivity in children
621 and adolescents: discrepancy between physician-based assessment and results of testing.
622 Pediatr Allergy Immunol 2011; 22(4): 405–10.
- 623 85. Celik GE, Aydin O, Dogu F, Cipe F, Boyvat A, Ikinçiogullari A, et al. Diagnosis of
624 immediate hypersensitivity to β -lactam antibiotics can be made safely with current
625 approaches. Int Arch Allergy Immunol 2012; 157(3): 311–7.
- 626 86. Garcia Nunez I, Barasona Villarejo MJ, Algaba Marmol MA, Moreno Aguilar C,
627 Guerra Pasadas F. Diagnosis of patients with immediate hypersensitivity to beta-lactams
628 using retest. J Investig Allergol Clin Immunol 2012; 22(1): 41–7.
- 629 87. Hjortlund J, Mortz CG, Skov PS, Eller E, Poulsen JM, Borch JE, et al. One-week oral
630 challenge with penicillin in diagnosis of penicillin allergy. Acta Derm Venereol 2012; 92(3):
631 307–12.
- 632 88. Iglesias-Souto J, Gonzalez R, Poza P, Sanchez-Machin I, Matheu V. Evaluating the
633 usefulness of retesting for beta-lactam allergy in children. Pediatr Infect Dis J 2012; 31(10):
634 1091–3.
- 635 89. Soto L, Guilarte M, Luengo O, Labrador M, Sala A. Drug hypersensitivity to
636 betalactams: a retrospective study to assess a diagnostic challenge. Allergy 2012; 67: 429–.

- 637 90. Erkocoglu M, Kaya A, Civelek E, Ozcan C, Cakır B, Akan A, et al. Prevalence of
638 confirmed immediate type drug hypersensitivity reactions among school children. *Pediatr*
639 *Allergy Immunol* 2013; 24(2): 160–7.
- 640 91. Hjortlund J, Mortz CG, Skov PS, Bindslev-Jensen C. Diagnosis of penicillin allergy
641 revisited: The value of case history, skin testing, specific IgE and prolonged challenge.
642 *Allergy* 2013; 68(8): 1057–64.
- 643 92. Macy E, Ngor EW. Safely diagnosing clinically significant penicillin allergy using
644 only penicilloyl-poly-lysine, penicillin, and oral amoxicillin. *J Allergy Clin Immunol Pract*
645 2013; 1(3): 258–63.
- 646 93. Rimawi RH, Cook PP, Gooch M, Kabchi B, Ashraf MS, Rimawi BH, et al. The
647 impact of penicillin skin testing on clinical practice and antimicrobial stewardship. *J Hosp*
648 *Med* 2013; 8(6): 341–5.
- 649 94. Al-Ahmad M, Rodriguez Bouza T, Arifhodzic N. Penicillin allergy evaluation:
650 experience from a drug allergy clinic in an Arabian Gulf Country, Kuwait. *Asia Pac Allergy*
651 2014; 4(2): 106–12.
- 652 95. Bourke J, Pavlos R, James I, Phillips E. Improving the Effectiveness of Penicillin
653 Allergy De-labeling. *J Allergy Clin Immunol Pract* 2014; 3(3): 365–74.e1.
- 654 96. Colas H, David V, Molle I, Bernier C, Magnan A, Pipet A. Drug provocation tests for
655 children: In-hospital or in an outpatient consultation? A protocol comparing the two
656 possibilities. *Revue Francaise D Allergologie* 2014; 54(4): 300–6.
- 657 97. Fox SJ, Park MA. Penicillin skin testing is a safe and effective tool for evaluating
658 penicillin allergy in the pediatric population. *J Allergy Clin Immunol Pract* 2014; 2(4): 439–
659 44.

98. Picard M, Paradis L, Bégin P, Paradis J, Des Roches A. Skin testing only with penicillin G in children with a history of penicillin allergy. *Ann Allergy Asthma Immunol* 2014; 113(1): 75–81.
99. Trautmann A, Seitz CS, Stoevesandt J, Kerstan A. Aminopenicillin-associated exanthem: lymphocyte transformation testing revisited. *Clin Exp Allergy* 2014; 44(12): 1531–8.
100. Zambonino MA, Corzo JL, Munoz C, Requena G, Ariza A, Mayorga C, et al. Diagnostic evaluation of hypersensitivity reactions to beta-lactam antibiotics in a large population of children. *Pediatr Allergy Immunol* 2014; 25(1): 80–7.
101. Blanca-Lopez N, Perez-Alzate D, Ruano F, Garcimartin M, de la Torre V, Mayorga C, et al. Selective immediate responders to amoxicillin and clavulanic acid tolerate penicillin derivative administration after confirming the diagnosis. *Allergy* 2015; 70(8): 1013–9.
102. Ferré-Ybarz L, Salinas Argente R, Gómez Galán C, Duocastella Selvas P, Nevot Falcó S. Analysis of profitability in the diagnosis of allergy to beta-lactam antibiotics. *Allergol Immunopathol (Madr)* 2015; 43(4): 369-75.
103. Abrams EM, Wakeman A, Gerstner TV, Warrington RJ, Singer AG. Prevalence of beta-lactam allergy: a retrospective chart review of drug allergy assessment in a predominantly pediatric population. *Allergy Asthma Clin Immunol* 2016; 12.
104. Atanaskovic-Markovic M, Gaeta F, Medjo B, Gavrovic-Jankulovic M, Cirkovic Velickovic T, Tmusic V, et al. Non-immediate hypersensitivity reactions to beta-lactam antibiotics in children - our 10-year experience in allergy work-up. *Pediatr Allergy Immunol* 2016; 27(5): 533–8.
105. Chen JR, Tarver SA, Alvarez KS, Tran T, Khan DA. A Proactive Approach to Penicillin Allergy Testing in Hospitalized Patients. *J Allergy Clin Immunol Pract* 2016; 5(3): 686–93.

- 685 106. Chiriac AM, Rerkpattanapipat T, Bousquet PJ, Molinari N, Demoly P. Optimal step
686 doses for drug provocation tests to prove beta-lactam hypersensitivity. *Allergy* 2016; 72(4):
687 552–61.
- 688 107. King EA, Challa S, Curtin P, Bielory L. Penicillin skin testing in hospitalized patients
689 with beta-lactam allergies: Effect on antibiotic selection and cost. *Ann Allergy Asthma*
690 *Immunol* 2016; 117(1): 67–71.
- 691 108. Mill C, Primeau MN, Medoff E, Lejtenyi C, O'Keefe A, Netchiporouk E, et al.
692 Assessing the Diagnostic Properties of a Graded Oral Provocation Challenge for the
693 Diagnosis of Immediate and Nonimmediate Reactions to Amoxicillin in Children. *JAMA*
694 *Pediatr* 2016; 170(6): e160033.
- 695 109. Manuyakorn W, Singvijarn P, Benjaponpitak S, Kamchaisatian W, Rerkpattanapipat
696 T, Sasisakulporn C, et al. Skin testing with β -lactam antibiotics for diagnosis of β -lactam
697 hypersensitivity in children. *Asian Pac J Allergy Immunol* 2016; 34(3): 242–7.
- 698 110. Moreno E, Laffond E, Munoz-Bellido F, Gracia MT, Macias E, Moreno A, et al.
699 Performance in real life of the European Network on Drug Allergy algorithm in immediate
700 reactions to beta-lactam antibiotics. *Allergy* 2016; 71(12): 1787–90.
- 701 111. Mota I, Gaspar A, Chambel M, Piedade S, Morais-Almeida M. Hypersensitivity to
702 beta-lactam antibiotics: a three-year study. *Eur Ann Allergy Clin Immunol* 2016; 48(6): 212–
703 9.
- 704 112. Ratzon R, Reshef A, Efrati O, Deutch M, Forschmidt R, Cukierman-Yaffe T, et al.
705 Impact of an extended challenge on the effectiveness of β -lactam hypersensitivity
706 investigation. *Ann Allergy Asthma Immunol* 2016; 116(4): 329–33.
- 707 113. Faitelson Y, Boaz M, Dalal I. Asthma, Family History of Drug Allergy, and Age
708 Predict Amoxicillin Allergy in Children. *J Allergy Clin Immunol Pract* 2017; 6(4): 1363–7.

- 709 114. Fazlollahi MR, Bidad K, Shokouhi R, Dashti R, Nabavi M, Movahedi M, et al.
710 Frequency and Risk Factors of Penicillin and Amoxicillin Allergy in Suspected Patients with
711 Drug Allergy. Arch Iran Med 2017; 20(1): 34–7.
- 712 115. Juri MC, Romero DSF, Larrauri B, Malbran E, Torre G, Malbran A. Allergy to drugs.
713 Experience in 771 procedures. Medicina-Buenos Aires 2017; 77(3): 180–4.
- 714 116. Marwood J, Aguirrebarrena G, Kerr S, Welch SA, Rimmer J. De-labelling self-
715 reported penicillin allergy within the emergency department through the use of skin tests and
716 oral drug provocation testing. Emerg Med Australas 2017; 29(5): 509–15.
- 717 117. McDanel DL, Azar AE, Dowden AM, Murray-Bainer S, Noiseux NO, Willenborg M,
718 et al. Screening for Beta-Lactam Allergy in Joint Arthroplasty Patients to Improve Surgical
719 Prophylaxis Practice. Journal Arthroplasty 2017; 32(9s): S101–s8.
- 720 118. Misirlioglu ED, Guvenir H, Toyran M, Vezir E, Capanoglu M, Civelek E, et al.
721 Frequency of selective immediate responders to aminopenicillins and cephalosporins in
722 Turkish children. Allergy Asthma Proc 2017; 38(5): 376–82.
- 723 119. Ramsey A, Staicu ML. Use of a Penicillin Allergy Screening Algorithm and Penicillin
724 Skin Testing for Transitioning Hospitalized Patients to First-Line Antibiotic Therapy. J
725 Allergy Clin Immunol Pract 2017; 6(4): 1349–55.
- 726 120. Soria A, Autegarden E, Amsler E, Gaouar H, Vial A, Francès C, et al. A clinical
727 decision-making algorithm for penicillin allergy. Ann Med 2017; 49(8): 710–7.
- 728 121. Sundquist BK, Bowen BJ, Otabor U, Celestin J, Sorum PC. Proactive penicillin
729 allergy testing in primary care patients labeled as allergic: outcomes and barriers. Postgrad
730 Med 2017; 129(8): 915–20.
- 731 122. Vyles D, Adams J, Chiu A, Simpson P, Nimmer M, Brousseau DC. Allergy Testing in
732 Children With Low-Risk Penicillin Allergy Symptoms. Pediatr 2017; 140(2).

- 733 123. Marraccini P, Pignatti P, A DAA, Salimbeni R, Consonni D. Basophil Activation Test
734 Application in Drug Hypersensitivity Diagnosis: An Empirical Approach. *Int Arch Allergy*
735 *Immunol* 2018; 1–7.
- 736 124. Moussa Y, Shuster J, Matte G, Sullivan A, Goldstein RH, Cunningham D, et al. De-
737 labeling of beta-lactam allergy reduces intraoperative time and optimizes choice in antibiotic
738 prophylaxis. *Surgery* 2018.
- 739 125. Sedlackova L, Prucha M, Polakova I, Mikova B. Evaluation of IFN-gamma Enzyme-
740 linked Immunospot Assay (ELISPOT) as a First-line Test in the Diagnosis of Non-Immediate
741 Hypersensitivity to Amoxicillin and Penicillin. *Prague Med Rep* 2018; 119(1): 30–42.
- 742 126. Chiriac AM, Vasconcelos MJ, Izquierdo L, Ferrando L, Nahas O, Demoly P. To
743 challenge or not to challenge: Literature data on the positive predictive value of skin tests to
744 beta-lactams. *J Allergy Clin Immunol Pract* 2019;7(7):2404-2408.
- 745 127. Chiriac AM, Banerji A, Gruchalla RS, Thong BYH, Wickner P, Mertes PM,
746 Terreehorst I, Blumenthal KG. Controversies in Drug Allergy: Drug Allergy Pathways. *J*
747 *Allergy Clin Immunol Pract* 2019;7(1):46-60.
- 748 128. Blanca M, Romano A, Torres MJ, Fernández J, Mayorga C, Rodriguez J, et al. Update
749 on the evaluation of hypersensitivity reactions to betalactams. *Allergy* 2009; 64(2): 183–93.
- 750 129. Torres MJ, Celik GE, Whitaker P, Atanaskovic-Markovic M, Barbaud A, Bircher A, et
751 al. A EAACI drug allergy interest group survey on how European allergy specialists deal
752 with beta-lactam allergy. *Allergy* 2019.
- 753 130. Sousa-Pinto B, Blumenthal KG, Macy E, Bavbek S, Benić MS, Alves-Correia M, et
754 al. Diagnostic Testing for Penicillin Allergy: A Survey of Practices and Cost Perceptions.
755 *Allergy* 2019.

- 756 131. Iammatteo M, Alvarez Arango S, Ferastraoaru D, Akbar N, Lee AY, Cohen HW, et
757 al. Safety and Outcomes of Oral Graded Challenges to Amoxicillin without Prior Skin
758 Testing. *J Allergy Clin Immunol Pract* 2019; 7(1): 236–43.
- 759 132. Macy E, Vyles D. Who needs penicillin allergy testing? *Ann Allergy Asthma*
760 *Immunol* 2018; 121(5): 523–9.

Tables

Table I. Description of studies included in the systematic review and assessed by means of bivariate random-effects meta-analyses

Reference	Country	N participants (% females)	Participants' age group	Setting	Sampling method	Performed tests ^a	Drugs participants reported to be allergic	Timing of reactions	Index reactions of included participants
Levine BB, 1968	USA	218 ^b	Children and adults	Patients (not specified)	Consecutive sample ^c	Skin tests (SPT; IDT ^d)	Penicillins	Immediate and non-immediate	MPE; U/A; others unspecified
Warrington RJ, 1978	Canada	253 ^b	Children and adults	Inpatients and outpatients	Convenience sample	Skin tests (SPT; IDT ^d)	Penicillins	Immediate and non-immediate	MPE; U/A; others (including anaphylaxis and arthralgias/arthritis)
Blanca M, 1990	Spain	288 ^b	Children and adults	Outpatients	Convenience sample	Skin tests (IDT); sIgE	Penicillins	Immediate and non-immediate	^e
Sogn DD, 1992	USA	825 ^b	Adults	Inpatients	Consecutive sample ^c	Skin tests (SPT; IDT ^d)	Penicillins	Immediate and non-immediate	^e
Romano A, 1995	Italy	60 ^b	Children and adults	Patients (not specified)	Consecutive sample ^f	Skin tests (SPT; IDT); sIgE	Aminopenicillins	Non-immediate	MPE
Sastre J, 1996	Spain	576 ^b	Children and adults	Outpatients	Consecutive sample ^f	Skin tests (SPT; IDT ^d)	Penicillins	Immediate	^e
Cederbrant K, 1998	Sweden	17 (82)	Adults	Outpatients	Convenience sample	Skin tests (SPT; IDT); sIgE	Penicillins	Immediate and non-immediate	MPE; U/A; others (including GI reactions and pruritus)
Loza Cortina C, 1998	Spain	76 ^b	Children	Outpatients	Convenience sample	Skin tests (SPT; IDT ^d); sIgE	β -lactams	^g	Exclusion only of anaphylactic shock (most cases with MPE or U/A)
Blanca M, 2001	Spain	81 (67)	Children and adults	Patients (not specified)	Convenience sample	sIgE	Penicillins	Immediate and non-immediate	Anaphylaxis; U/A; MPE
Forrest DM, 2001	Canada	72 ^b	Adults	Inpatients	Convenience sample	Skin tests	Penicillins	Immediate	^e
Romano A, 2002	Italy	259 (72.6)	Children and adults	Patients (not specified)	Consecutive sample ^f	Skin tests (IDT)	Penicillins	Non-immediate	MPE; U/A
Brusca I, 2007	Italy	122 (57.4)	Children and adults	Outpatients	Consecutive sample ^f	sIgE	Penicillins ^h	Immediate	U/A
Goldberg A, 2008	Israel	169 (62.7)	Children and adults	Outpatients	Consecutive sample ^f	Skin tests (IDT)	Penicillins	Immediate and non-immediate	Exclusion only of life-threatening reactions
Tantikul C, 2008	Thailand	4 ^b	Children	Inpatients	Consecutive sample ^f	Skin tests (SPT)	Amoxicillin ^h	Immediate and non-immediate	MPE; U/A
Caubet JC, 2010	Switzerland	88 (50)	Children	ED	Consecutive sample ^f	sIgE	β -lactams	Non-immediate	MPE; U/A

Macy E, 2010	USA	150 (69.3)	Adults	Outpatients	Convenience sample	Skin tests (SPT; IDT ^d); sIgE	Penicillins	Immediate and non-immediate	Exclusion only of severe mucocutaneous reactions, hemolytic anemia or nephritis
Richter AG, 2011	UK	47 ^b	Adults	Outpatients	Consecutive sample ⁱ	Skin tests (SPT; IDT); sIgE	Penicillins ^h	Immediate and non-immediate	U/A; rash; pruritus; other isolated reactions; anaphylaxis
Kopac P, 2012	Slovenia	606 (75.9)	Children and adults	Outpatients	Consecutive sample ^f	Skin tests (SPT; IDT); sIgE	Penicillins	Immediate and non-immediate	^e
Caubet JC, 2014	Switzerland	96 ^b	Children	Outpatients and ED	Convenience sample	Skin tests (IDT)	β -lactams	Non-immediate	MPE; U/A
Barni S, 2015	Italy	352 (50.3)	Children	Outpatients	Convenience sample	Skin tests (SPT; IDT ^d)	Amoxicillin	Non-immediate	MPE; U/A
Mori F, 2015	Italy	190 (47.9)	Children	Outpatients	Consecutive sample ^f	Skin tests (SPT; IDT); sIgE	Amoxicillin ^h	Immediate and non-immediate	MPE; U/A; others (including anaphylaxis and undefined reactions)
Rosenfield L, 2015	Canada	240 (72.9)	Children and adults	Outpatients	Convenience sample	Skin tests (SPT; IDT ^d)	β -lactams	Immediate and non-immediate	Rash; pruritus; U/A; cardiovascular reactions; others (including anaphylaxis and unknown reactions)
Mawhirt SL, 2016	USA	49 ^b	Adults	Inpatients and outpatients	Convenience sample	Skin tests (SPT; IDT ^d)	Penicillins ^h	Immediate	Cutaneous or other isolated reactions; mucocutaneous reactions; anaphylaxis
Meng J, 2016	UK	84 (66)	^g	Outpatients	Consecutive sample ^f	Skin tests (SPT; IDT); sIgE	Penicillins	Immediate and non-immediate	Rash; U/A; anaphylaxis; other/unknown
Confino-Cohen R, 2017	Israel	642 (51.4)	Children and adults	Outpatients	Consecutive sample ^f	Skin tests (IDT)	Penicillins	Non-immediate	MPE (90%); U/A; other non-severe delayed reactions
Tannert LK, 2017	Denmark	21 (71)	Adults	Outpatients	Convenience sample	Skin tests (IDT); sIgE	Penicillins	Immediate and non-immediate	Anaphylaxis; U/A; rash
Ibáñez MD, 2018	Spain	732 (48.8)	Children	Outpatients and ED	Consecutive sample ^f	Skin tests (SPT; IDT); sIgE	Penicillin; amoxicillin-clavulanate; cloxacillin	Immediate and non-immediate	Exclusion only of severe cutaneous reactions or severe anaphylaxis

ED: Emergency department; IDT: Intradermal tests; MPE: Maculopapular exanthem; sIgE: Specific IgE; SPT: Skin prick tests; U/A: Urticaria/angioedema; UK: United Kingdom; USA: United States of America

^a For each primary study, only diagnostic tests whose accuracy could be calculated are listed; ^b Sex discrimination not available; ^c Consecutive sample consisting of all patients with history of penicillin allergy and in whom penicillin is indicated; ^d Skin prick tests and intradermal tests were not presented separately; ^e Not clearly specified for all patients; ^f Consecutive sample consisting of all participants with history/events compatible with penicillin allergy; ^g Information not available; ^h Participants with reported cephalosporin allergy are not being counted; ⁱ Consecutive sample consisting of all participants tested for penicillin allergy

Table II. Results of the bivariate diagnostic random-effects meta-analyses for penicillin allergy diagnostic tests

Diagnostic test	N studies	AUC-sROC	Partial AUC-sROC	Summary measures (95% CI)					Correlation coefficient ^a	I ²
				Sensitivity	Specificity	LR ⁺	LR ⁻	DOR		
Skin tests	20	0.841	0.686	30.7% (18.9-45.8%)	96.8% (94.2-98.3%)	10.1 (4.9-18.6)	0.7 (0.6-0.8)	14.7 (6.2-29.7)	0.363	38.2%
More than 50% participants ^b	8	0.842	0.667	39.2% (14.6-70.0%)	93.5% (83.0-97.8%)	6.6 (2.1-16.0)	0.6 (0.3-0.9)	11.6 (2.5-34.5)	0.545	45.5%
Generalized provocation tests ^c	4	0.694	0.480	41.2% (14.2-75.0%)	87.4% (54.0-97.6%)	3.6 (1.4-8.5)	0.7 (0.4-0.9)	5.5 (1.9-12.6)	1.000	36.4%
Consecutive sample ^d	6	0.655	0.449	23.3% (9.3-47.3%)	95.6% (83.8-98.9%)	6.5 (1.4-19.8)	0.8 (0.6-1.0)	8.7 (1.5-28.7)	0.416	65.5%
Adult participants	5	0.677	0.228	24.2% (11.7-43.5%)	97.6% (94.1-99.0%)	12.2 (2.7-34.6)	0.8 (0.6-0.9)	17.2 (3.0-55.9)	-1.000	0%
Pediatric participants	5	0.857	0.338	31.5% (10.6-64.1%)	96.0% (87.8-98.7%)	7.9 (3.9-14.5)	0.7 (0.4-0.9)	11.8 (5.0-23.8)	1.000	20.2%
Skin tests (non-immediate reactions) ^{e,f}	7	0.926	0.396	44.2% (14.2-79.1%)	95.1% (88.3-98.0%)	9.7 (2.6-23.5)	0.6 (0.2-0.9)	21.4 (3.0-76.7)	0.335	22.2%
Intradermal tests ^e	10	0.778	0.582	29.9% (15.0-50.8%)	95.1% (88.1-98.0%)	6.5 (2.6-13.6)	0.7 (0.5-0.9)	9.2 (3.2-21.1)	0.607	54.8%
Intradermal tests (non-immediate reactions) ^{e,f}	6	0.939	0.404	43.1% (12.8-79.5%)	95.1% (90.9-97.5%)	9.7 (2.1-23.5)	0.6 (0.2-0.9)	23.0 (2.3-94.1)	-0.194	35.3%
Specific IgE	11	0.760	0.420	19.3% (12.0-29.4%)	97.4% (95.2-98.6%)	7.7 (3.8-13.8)	0.8 (0.7-0.9)	8.8 (4.3-17.9)	1.000	8.5%
More than 50% participants ^b	6	0.910	0.132	17.4% (8.7-31.0%)	97.6% (96.0-98.7%)	7.9 (3.4-15.6)	0.8 (0.7-0.9)	9.7 (3.7-21.0)	1.000	0% ^g
Specific IgE (immediate reactions) ^e	4	0.279	0.262	26.2% (18.7-35.3%)	92.1% (64.7-98.7%)	5.2 (0.7-21.0)	0.8 (0.7-1.2)	6.8 (0.6-29.2)	-1.000	0% ^g

AUC-sROC: Area under the curve of the summary receiver operating characteristic curve; CI: Confidence interval; DOR: Diagnostic odds ratio; LR: Likelihood ratio; z=z-score

^a Correlation coefficient between sensitivity and false positive rate; ^b Sensitivity analysis restricted to studies in which at least 50% of both patients with positive and negative results on index tests received a drug provocation test; ^c Studies assessing a mix of participants with immediate and non-immediate reactions (excluding only those with very severe index reactions), and testing all with index tests and DPT; ^d Sensitivity analysis restricted to consecutive samples of patients specifically reporting a history of penicillin allergy and testing for penicillin polylysine, minor determinant mixture, benzylpenicillin and an aminopenicillin; ^e Insufficient number of studies to perform sensitivity analyses as pre-specified; ^f All participants had maculopapular exanthem or urticaria/angioedema as their index reaction; ^g Obtained with the Higgins estimator.

Table III. Variables identified as associated with heterogeneity in the bivariate diagnostic random-effects meta-analyses for penicillin allergy diagnostic tests

Diagnostic test	Variables associated with heterogeneity ^a	Variables associated with heterogeneity in summary sensitivity estimation ^b	Variables associated with heterogeneity in summary specificity estimation ^b
Skin tests	Year ($p=0.002$); Middle East ($p=0.002$)	Year ($z=-2.8$; $p=0.005$)	Middle East ($z=-3.4$; $p=0.001$)
Skin tests (non-immediate reactions)	Year ($p=0.02$); Europe ($p=0.05$); Middle East ($p=0.02$); Suspected drug testing ($p=0.02$)	Year ($z=-2.9$; $p=0.004$)	-
Intradermal tests	Europe ($p=0.001$); Middle East ($p=0.001$); Suspected drug testing ($p=0.02$)	-	Europe ($z=2.4$; $p=0.02$); Middle East ($z=-2.4$; $p=0.02$)
Intradermal tests (non-immediate reactions)	Year ($p=0.010$); Europe ($p=0.05$); Middle East ($p=0.05$); Suspected drug testing ($p=0.05$)	Year ($z=-2.7$; $p=0.006$)	Reported allergy to aminopenicillins ($z=-2.2$; $p=0.03$)
Specific IgE ^c	Adults ($p=0.04$)	Emergency department ($z=-2.0$; $p=0.05$)	Adults ($z=-2.6$; $p=0.009$)

^a As assessed by the likelihood ratio test; ^b As assessed by meta-regression; ^c Insufficient number of studies assessing only participants reporting immediate reactions to perform meta-regression; Adults=Assessment of adult patients; Year=Publication year; z =z-score

Table IV. Limitations of current evidence and key aspects for future studies

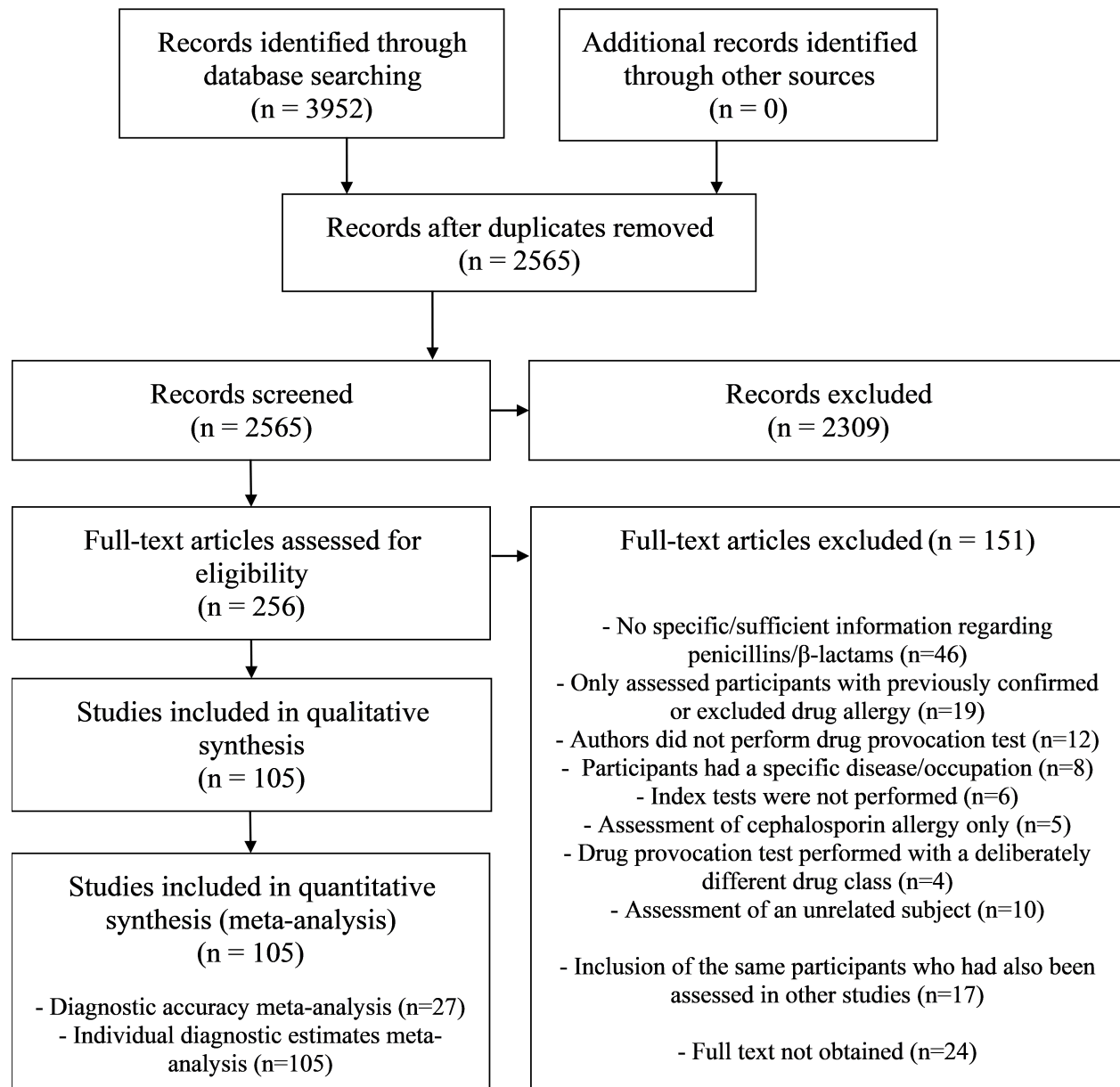
Limitations of current evidence	Key requirements needed for future studies
- Inconsistent eligibility criteria often creating heterogeneous patient mixes; - Inconsistent reporting of demographic and clinical details (e.g., timing or clinical manifestations of index reactions); - Limited ability to assess accuracy of skin tests and specific IgE for different clinical phenotypes or risk groups (except non-immediate maculopapular exanthema or urticaria/angioedema); - Potential selection bias related to using drug provocation tests; - Limited data from Latin America, Asia and Africa.	- Definition of more homogeneous inclusion criteria to have less patient variation; - Standardization of testing protocols; - Limitation selection bias through reporting of consecutive samples; - Inclusion of comprehensive details of index reactions (e.g., clinical manifestations and timing); - Presentation of results stratified by different reaction phenotypes and risk groups OR publication of anonymized patient-level data.

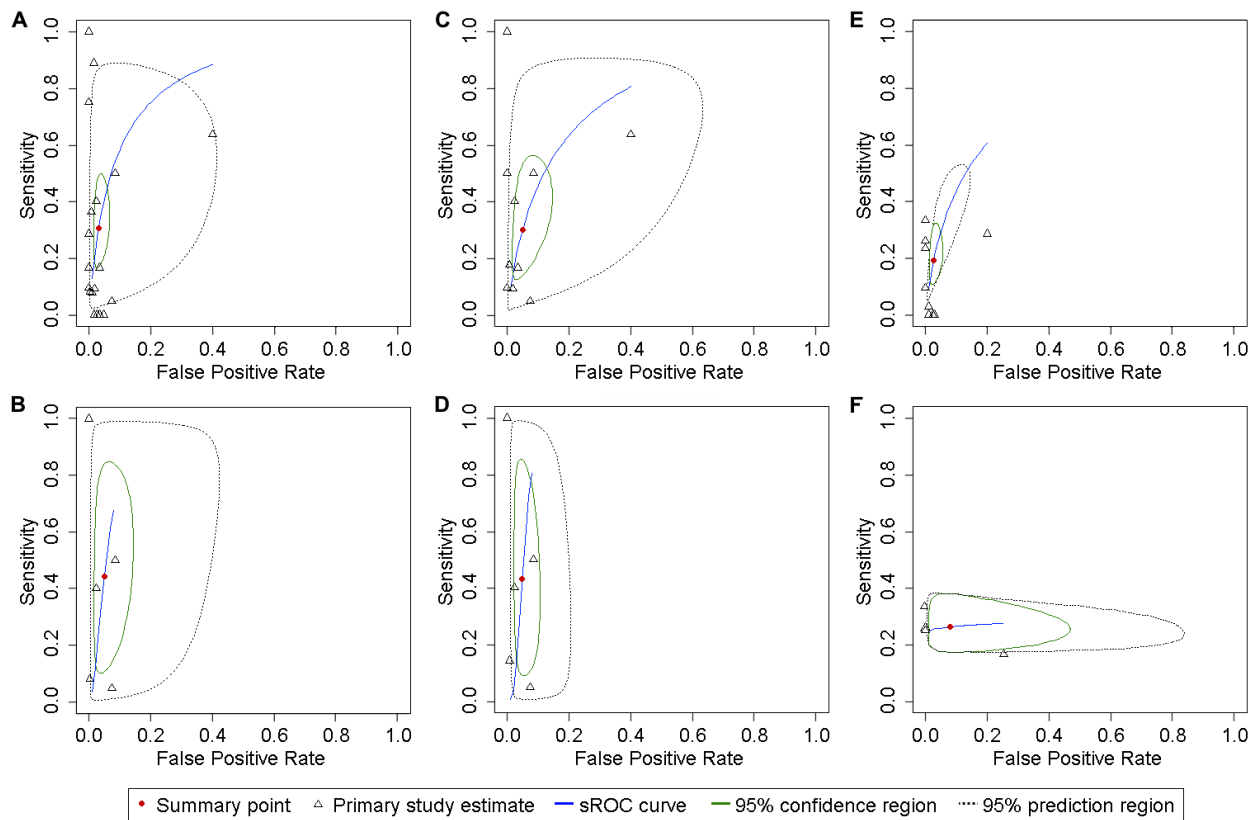
Figure legends

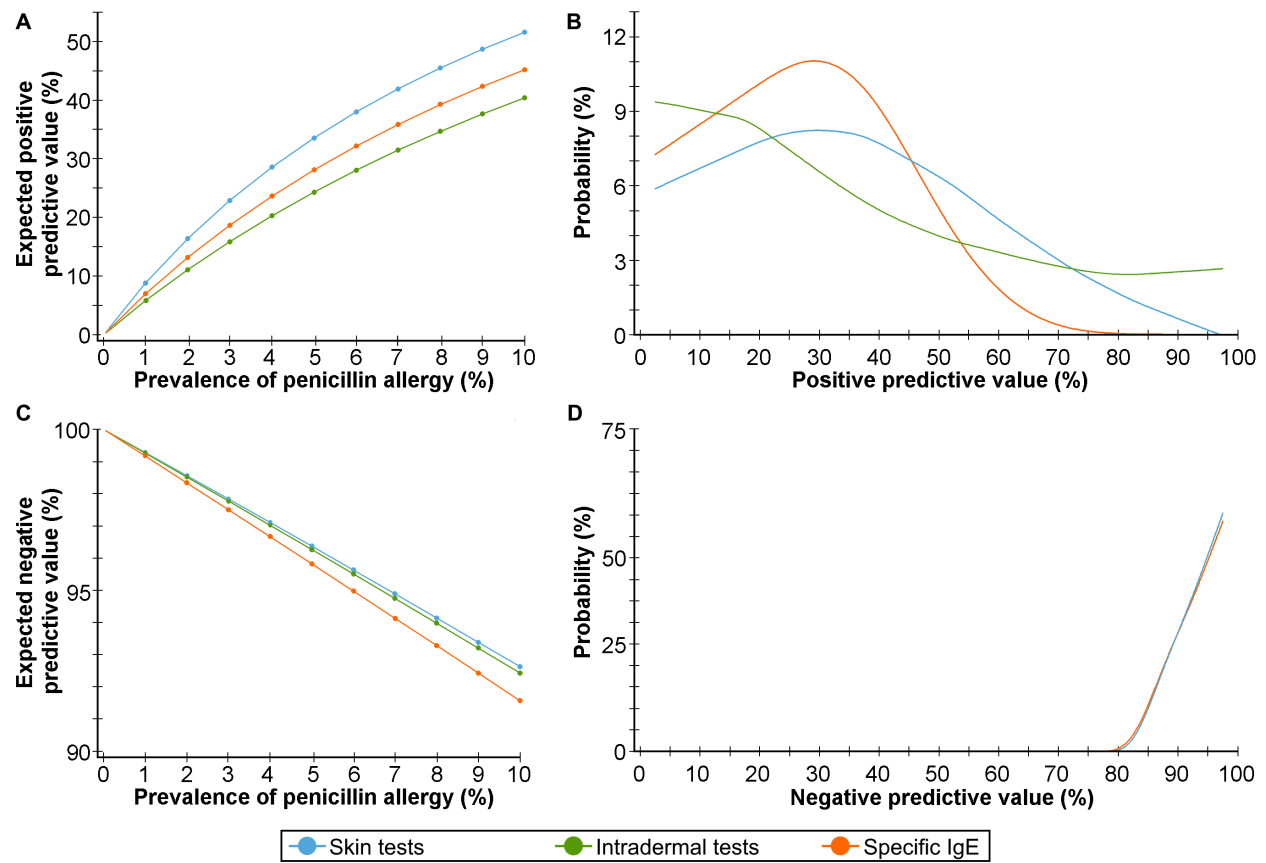
Figure 1. Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram illustrating study selection process.

Figure 2. Summary ROC curves for the diagnostic performance of penicillin skin tests (all reactions – A; non-immediate reactions - B), intradermal tests (all reactions – C; non-immediate reactions – D), and specific IgE quantification (all reactions – E; immediate reactions – F): For each test, the summary ROC curve plots the respective predicted sensitivity against the false positive rate. Within it, the summary point indicates the summary (“pooled”) sensitivity and specificity of the respective test (its precision is indicated by a 95% confidence region). Both the summary point and the summary ROC curve were calculated based on the diagnostic measures presented in the included primary studies. Between-studies variability is plotted through 95% prediction regions, which indicate the region where we are most confident the test diagnostic performance lies. For immediate reactions, we were not able to plot summary ROC curves, as we had not a sufficient number of studies presenting values for true positives, true negatives, false positives and false negatives.

Figure 3. Estimates of the positive and negative predictive values based on the pooled sensitivity and specificity obtained by bivariate meta-analysis. Figure 3A depicts the estimates of the positive predictive value across a range of penicillin allergy prevalences (univariate sensitivity analysis). Figure 3B depicts the probability distributions (“Probability (%)”) of the positive predictive values obtained by means of probabilistic sensitivity analyses. Figure 3C depicts the estimates of the negative predictive value across a range of penicillin allergy prevalences (univariate sensitivity analysis). Figure 3D depicts the probability distributions (“Probability (%)”) of the negative predictive values obtained by means of probabilistic sensitivity analyses.







Online repository - Accuracy of penicillin allergy diagnostic tests: A systematic review and meta-analysis

In this systematic review, we included 105 primary studies, of which 27 studies were included in bivariate random-effects meta-analyses, while the remainder were only quantitatively assessed by means of individual diagnostic parameter estimates meta-analysis (a less robust approach in the context of meta-analysis of diagnostic studies). The main manuscript describes the 27 studies included in bivariate random-effects meta-analyses, as well as the results of those analyses. In this online repository, we describe all included primary studies, and present the methods and results for individual diagnostic parameter meta-analyses.

Description of all included primary studies

The included studies were published between 1964 and 2018, and altogether assessed 31,761 participants (Table E2)¹⁻¹⁰⁵. Most studies had been performed in both children and adults ($n=45$)^{3,5,7,10,11,15-17,22,24,25,27,29-34,37,38,41,44,48,55,56,59,60,63-66,70,74,75,77,78,80,81,87-90,92,98,105}, but 30 assessed exclusively children^{9,14,19-21,36,42,43,45,46,49-52,54,58,61,67-69,72,73,76,79,83,86,91,96,101,102}, while 24 assessed exclusively adults^{2,12,13,18,23,28,35,39,40,47,53,57,62,71,82,84,93-95,97,99,100,103,104} (no information provided by six studies^{1,4,6,8,26,85}). The percentage of females was of 57.1%, based on the 70 studies which provided information on participants' gender distribution^{2,7,11,13,14,17-19,21,22,25,27,28,31,33-35,37-39,41,43-48,51,52,55,57-60,62-66,69-73,75-80,82,83,85-92,94,95,97-102,104,105}.

Most studies assessed outpatients ($n=78$)^{1,5,7,9-11,13,14,16-21,24-26,29,31-33,35-41,43,44,46-48,50-54,56-60,62,63,65-68,70,71,73-79,81,83-93,95,96,98-100,102-104}, with a minority assessing inpatients ($n=14$)^{5,7,8,12,23,30,34,42,63,64,80,82,84,97}, emergency department patients ($n=7$)^{19,45,67,83,94,101,102} or participants from the general population ($n=1$)⁶¹. For 13 studies^{2-4,6,15,22,27,28,49,55,69,72,105}, the setting of assessed patients was not specified. The sampling method was consecutive for 62 studies^{1,3,7,8,11-13,15,16,21,26,27,29,31-33,36-38,41,42,44-46,48,50,52-54,56-60,62,63,65,66,68-70,72,74-76,78-83,85-88,90,92,93,96,98,102,104}, while 41 studies adopted a convenience sample^{4-6,9,10,14,17-20,22-25,28,30,34,35,39,40,43,47,49,51,55,64,67,71,73,77,84,89,91,94,95,97,99-101,103,105}, one study assessed a cluster sample⁶¹ and another one a volunteer sample².

For most studies ($n=52$)^{1-5,7,8,10-12,16-18,22-31,33,35,38,41,47-49,51,53,59,62-64,66,70,80,84,85,90,93,94,97-103,105}, data were presented for participants reporting to be allergic to penicillins (excluding other β -lactam classes); in 43 studies, data were provided for reported allergy to overall β -lactams^{6,9,13,14,19-21,32,34,37,39,40,43-46,50,52,54-58,60,61,65,67,69,72,75,77-79,81-83,87-89,92,95,96,104}. Ten studies only included participants with reported allergy to aminopenicillins^{15,36,42,68,71,73,74,76,86,91}. Ninety-four studies provided information on the type of reaction participants reported^{3-6,9-13,15-19,21-24,27-68,70-92,94-103,105} – of these, 31 studies presented discriminated results for participants reporting immediate reactions^{4,11,16,18,22-24,29,33,36,37,42,51,52,54-56,61-63,72,74,76,83,84,86,87,91,96,100,101}, and 26 for participants claiming non-immediate reactions^{15,18,22,27,36,37,40-42,45,52,54,62,63,67,71-73,76,79,83,90,91,100,101,105}.

Quality of all included primary studies is summarized in Figure E1 and Table E5.

Meta-analysis of individual diagnostic parameter estimates: Methods

In addition to bivariate random-effects meta-analysis, we performed individual diagnostic parameter estimates meta-analysis. This allowed for obtaining pooled estimates with information from primary studies in which only a subset of participants were subject to drug provocation test (DPT). For example, there were several studies in which DPT had only been performed in patients with negative results on their index tests, only allowing for retrieving the frequency of true negatives and false negatives.

For each index test, we performed a quantitative synthesis of its frequency of positive results, sensitivity, specificity, negative predictive value (NPV) and positive predictive value (PPV), using raw data presented in Table E3. We performed random-effects meta-analyses (based on the restricted maximum likelihood approach), using logit-transformed proportions. Heterogeneity was assessed using the p -value of Cochran's Q statistics, and I^2 statistics: a p -value <0.10 and an $I^2 >50\%$ were considered to represent substantial heterogeneity. To identify moderator variables, we performed meta-regression analyses whenever we had more than 10 primary studies. Tested variables were the same assessed in meta-regression for bivariate random-effects meta-analysis (see Main Manuscript). These variables were

initially individually assessed in univariable analyses. We then performed multivariable meta-regression using Akaike Information Criteria for model selection¹⁰⁶ (with automated selection from all possible candidate models). Subgroup analyses were also performed when there were more than three primary studies. Exponentials of the coefficients and respective confidence intervals were calculated and interpreted as odds ratios.

Meta-analysis of individual diagnostic parameter estimates: Results

For all assessed diagnostic tests, Table E4 presents a summary of the pooled frequency of positive results and the separate pooling of sensitivity, specificity and predictive values (estimates were based in all studies, including those that did not simultaneously provide the number of true positives, true negatives, false positives and false negatives). Table E4 also presents a summary of each multivariable meta-regression model with the respective residual heterogeneity. Complete results of all performed meta-regression models and subgroup analyses can be explored in the App in the Online Repository.

Description of App

Online interactive app allowing for exploration of the meta-analysis results for each penicillin allergy diagnostic test. Results of bivariate meta-analyses and of single diagnosis measure meta-analyses for each diagnosis test can be accessed using the left side bar. Within each page, an interactive table containing raw data and the results of meta-analysis are displayed – data can be filtered by using the active fields in the top row of the interactive table; by doing so, users can perform subgroup analyses.

The page presenting the results for bivariate meta-analysis contains a drop-down menu, allowing to choose the diagnostic tests and reaction timing for which results are to be presented.

The pages presenting the results for single estimates contain a drop-down menu allowing to choose the diagnostic measure for which results are to be presented. In these pages, whenever there are more than 10 primary studies, it is possible to perform univariable and multivariable meta-regressions – users may select (and deselect) variables to include in meta-regression models by interacting with the bottom-left panel; meta-regression results appear on the bottom-right panel.

Figure E1. Risk-of-bias graph. Review of authors' judgments of each risk-of-bias and applicability domain presented as percentages across all included studies. Regarding the possibility of bias (risk of bias domains; top panel), primary studies were classified as being of low, unclear or high risk; regarding the applicability of their results (applicability domains; bottom panel), primary studies were classified of being of low, unclear or high concern. A low risk of bias/applicability concern corresponds to higher quality. Judgements were performed according to QUADAS-2 criteria.

Table E1. Queries used for searching MEDLINE/PubMed, Web of Science and Scopus

Database	Drug element	Allergy element	Diagnostic Tests element	Diagnostic Accuracy element
MEDLINE/Pub Med	antibiotic*[title/abstract] penicillin*[text word] beta-lactam*[text word] drug*[title]	allerg*[title/abstract] hypersensitivity[title/abstract]]	skin test*[title/abstract] cutaneous test*[title/abstract] prick [title/abstract] percutaneous [title/abstract] puncture [title/abstract] intradermal[title/abstract] patch[title/abstract] epicutaneous[title/abstract] in vitro test*[title/abstract] in vitro allerg* test*[title/abstract] ige[title/abstract] immunoglobuline[title/abstract] sige[title/abstract] bat[title/abstract] basophil*activation test[title/abstract] ltt[title/abstract] lymphocyt*transformation test[title/abstract] provocation[text word] challenge*[text word]	specificity[title/abstract] false positive[title/abstract] false negative[title/abstract] diagnosis[mesh subheading] sensitivity[text word] predictive value*[text word]) distinguish*[text word] differentiat*[text word] enhancement[text word]) identif*[text word] detect*[text word] diagnos*[text word] accura*[text word] comparison[text word]
Web of Science	TS=antibiotic* TS=penicillin* TS=beta-lactam* TI=drug*	TS=allerg* TS=hypersensitivity	TS=skin test* TS=cutaneous test* TS=prick TS=percutaneous TS=puncture TS=intradermal TS=patch TS=epicutaneous TS=in vitro test* TS=in vitro allerg* test* TS=IgE TS=Immunoglobulin E TS=sIgE TS=BAT TS=basophil* activation test TS=LTT TS=lymphocyt* transformation test TS=provocation TS=challenge*	TS=sensitivity TS=specificity TS=false positive TS=false negative TS=diagnos* TS=predictive value* TS=distinguish* TS=differentiat* TS=enhancement TS=identif* TS=detect* TS=accura* TS=comparison

Scopus	antibiotic* allerg* (TITLE-ABS) antibiotic* hypersensitivity (TITLE-ABS) penicillin* allerg* (TITLE-ABS) penicillin* hypersensitivity (TITLE-ABS) beta-lactam* allerg* (TITLE-ABS) beta-lactam* hypersensitivity (TITLE-ABS) drug* allerg* (TITLE) drug* hypersensitivity (TITLE)	skin test* (TITLE-ABS) cutaneous test* (TITLE-ABS) prick (TITLE-ABS) percutaneous (TITLE-ABS) puncture (TITLE-ABS) intradermal (TITLE-ABS) patch (TITLE-ABS) epicutaneous (TITLE-ABS) in vitro test* (TITLE-ABS) in vitro allerg* test* (TITLE-ABS) IgE (TITLE-ABS) Immunoglobulin E (TITLE-ABS) sIgE (TITLE-ABS) BAT (TITLE-ABS) basophil* activation test (TITLE-ABS) LTT (TITLE-ABS) lymphocyt* transformation test (TITLE-ABS) provocation (TITLE-ABS) challenge* (TITLE-ABS)	sensitivity (TITLE-ABS) specificity (TITLE-ABS) false positive (TITLE-ABS) false negative (TITLE-ABS) diagnos* (TITLE-ABS) predictive value* (TITLE-ABS) distinguish* (TITLE-ABS) differentiat* (TITLE-ABS) enhancement (TITLE-ABS) identif* (TITLE-ABS) detect* (TITLE-ABS) accura* (TITLE-ABS) comparison (TITLE-ABS)
--------	--	---	--

Table E2. Description of studies included in the systematic review

Reference	Country	N participants (% females)	Participants' age group	Setting	Sampling method	Performed tests ^a	Drugs participants reported to be allergic	Timing of reactions
Brown BC, 1964	USA	1085 ^b	^c	Outpatients	Consecutive sample ^d	Skin tests (IDT)	Penicillins	^e
Resnik SS, 1965	USA	16 (0)	Adults	Not specified	Volunteer sample	Skin tests (IDT)	Penicillins	^e
Levine BB, 1968	USA	218 ^b	Children and adults	Patients (not specified)	Consecutive sample ^d	Skin tests (SPT; IDT ^f)	Penicillins	Immediate and non-immediate
Wide L, 1971	Sweden	23 ^b	^c	Patients (not specified)	Convenience sample	Skin tests (IDT); sIgE	Penicillins	Immediate
Warrington RJ, 1978	Canada	253 ^b	Children and adults	Inpatients and outpatients	Convenience sample	Skin tests (SPT; IDT ^f)	Penicillins	Immediate and non-immediate
Basomba A, 1979	Spain	81 ^b	^c	Patients (not specified)	Convenience sample	Skin tests (SPT; IDT); sIgE	β -lactams	Immediate and non-immediate
Sullivan TJ, 1981	USA	290 (53.8)	Children and adults	Inpatients and outpatients	Consecutive sample ^g	Skin tests (SPT; IDT)	Penicillins	^e
Van Arsdel PP, 1986	USA	97 ^b	^c	Inpatients	Consecutive sample ^g	Skin tests (SPT; IDT ^f)	Penicillins	^e
Graff-Lonnevig V, 1988	Sweden	298 ^b	Children	Outpatients	Convenience sample	Skin tests (SPT)	β -lactams	Immediate and non-immediate
Blanca M, 1990	Spain	288 ^b	Children and adults	Outpatients	Convenience sample	Skin tests (IDT); sIgE	Penicillins	Immediate and non-immediate
Surtees SJ, 1991	UK	132 (57.6)	Children and adults	Outpatients	Consecutive sample ^g	sIgE	Penicillins	Immediate
Sogn DD, 1992	USA	825 ^b	Adults	Inpatients	Consecutive sample ^d	Skin tests (SPT; IDT ^f)	Penicillins	Immediate and non-immediate
Gadde J, 1993	USA	776 (45.4)	Adults	Outpatients	Consecutive sample ^d	Skin tests (SPT; IDT ^f)	β -lactams	Immediate and non-immediate
Juchet A, 1994	France	45 (44)	Children	Outpatients	Convenience sample	Skin tests (SPT; IDT ^f)	β -lactams	^e
Romano A, 1995	Italy	60 ^b	Children and adults	Patients (not specified)	Consecutive sample ^h	Skin tests (SPT; IDT); sIgE	Aminopenicillins	Non-immediate
Sastre J, 1996	Spain	576 ^b	Children and adults	Outpatients	Consecutive sample ^h	Skin tests (SPT; IDT ^f)	Penicillins	Immediate
Macy E, 1997	USA	348 (44.0)	Children and adults	Outpatients	Convenience sample	Skin tests (SPT; IDT)	Penicillins	Immediate and non-immediate
Cederbrant K, 1998	Sweden	17 (82)	Adults	Outpatients	Convenience sample	Skin tests (SPT; IDT); sIgE	Penicillins	Immediate and non-immediate
Hervé M, 1998	France	112 (43.8)	Children	Outpatients and ED	Convenience sample	Skin tests (SPT; IDT); sIgE	β -lactams	Immediate and non-immediate
Loza Cortina C, 1998	Spain	76 ^b	Children	Outpatients	Convenience sample	Skin tests (SPT; IDT ^f); sIgE	β -lactams	^e
Pichichero ME, 1998	USA	247 (46.2)	Children	Outpatients	Consecutive sample ^g	Skin tests (SPT; IDT ^f)	Penicillin; amoxicillin; cephalosporins	Immediate and non-immediate
Blanca M, 2001	Spain	81 (67)	Children and adults	Patients (not specified)	Convenience sample	sIgE	Penicillins	Immediate and non-immediate
Forrest DM, 2001	Canada	72 ^b	Adults	Inpatients	Convenience sample	Skin tests	Penicillins	Immediate
Lebel B, 2001	France	17 ^b	Children and adults	Outpatients	Convenience sample	Skin tests (IDT)	Penicillins ⁱ	Immediate
López Tiro JJ, 2001	Mexico	24 (88)	Children and adults	Outpatients	Convenience sample	Skin tests (SPT; IDT)	Penicillins	^e

Novalbos A, 2001	Spain	958 ^b	^c	Outpatients	Consecutive sample ^h	Skin tests (SPT; IDT)	Penicillins	^e
Romano A, 2002	Italy	259 (72.6)	Children and adults	Patients (not specified)	Consecutive sample ^h	Skin tests (IDT)	Penicillins	Non-immediate
Solensky R, 2002	USA	58 (62)	Adults	Patients (not specified), hospital staff and students	Convenience sample	Skin tests (SPT; IDT ^f)	Penicillins (but not exclusively amoxicillin or ampicillin)	Immediate and non-immediate
Torres MJ, 2002	Spain	330 ^b	Children and adults	Outpatients	Consecutive sample ^h	Skin tests (SPT; IDT ^f)	Penicillins	Immediate
Borch JE, 2006	Denmark	13 ^b	Children and adults	Inpatients	Convenience sample	Skin tests (SPT; IDT)	Penicillins	Immediate and non-immediate
Wong BBL, 2006	Canada	91 (36)	Children and adults	Outpatients	Consecutive sample ^h	Skin tests (SPT; IDT ^f)	Penicillins	Immediate and non-immediate
Bousquet PJ, 2007	France	973 ^b	Children and adults	Outpatients	Consecutive sample ^h	Skin tests (SPT; IDT)	β -lactams	Immediate and non-immediate
Brusca I, 2007	Italy	122 (57.4)	Children and adults	Outpatients	Consecutive sample ^h	sIgE	Penicillins ⁱ	Immediate
del Real GA, 2007	USA	596 (53.5)	Children and adults	Inpatients	Convenience sample	Skin tests (SPT; IDT ^f)	β -lactams	Immediate and non-immediate
Fontaine C, 2007	France	11 (46)	Adults	Outpatients	Convenience sample	sIgE	Penicillins ⁱ	Immediate and non-immediate
Rebello Gomes E, 2007	Portugal	16 ^b	Children	Outpatients	Consecutive sample ⁱ	Skin tests (SPT; IDT)	Amoxicillin/amoxicillin-clavulanate ⁱ	Immediate and non-immediate
Bousquet PJ, 2008	France	1218 (69.8)	Children and adults	Outpatients	Consecutive sample ^h	Skin tests (SPT; IDT ^f)	β -lactams	Immediate and non-immediate
Goldberg A, 2008	Israel	169 (62.7)	Children and adults	Outpatients	Consecutive sample ^h	Skin tests (IDT)	Penicillins	Immediate and non-immediate
Nolan RC, 2008	Australia	29 (69)	Adults	Outpatients	Convenience sample	Skin tests (SPT; IDT)	β -lactams	Immediate and non-immediate
Padial A, 2008	Spain	18 ^b	Adults	Outpatients	Convenience sample	Skin tests (IDT)	β -lactams	Non-immediate
Rodriguez-Alvarez M, 2008	Spain	23 (478)	Children and adults	Outpatients	Consecutive sample ^h	Skin tests (SPT; IDT)	Penicillins	Non-immediate
Tantikul C, 2008	Thailand	4 ^b	Children	Inpatients	Consecutive sample ^h	Skin tests (SPT)	Amoxicillin ⁱ	Immediate and non-immediate
Hershkovich J, 2009	Israel	166 (47.0)	Children	Outpatients	Convenience sample	Skin tests (SPT; IDT ^f)	β -lactams	Immediate and non-immediate
Silva R, 2009	Portugal	67 (81)	Children and adults	Outpatients	Consecutive sample ^h	Skin tests (SPT; IDT); sIgE	β -lactams	Immediate and non-immediate
Caubet JC, 2010	Switzerland	88 (50)	Children	ED	Consecutive sample ^h	sIgE	β -lactams	Non-immediate
Chambel M, 2010	Portugal	161 (40.4)	Children	Outpatients	Consecutive sample ^g	Skin tests (SPT; IDT); sIgE	β -lactams	Immediate and non-immediate
Macy E, 2010	USA	150 (69.3)	Adults	Outpatients	Convenience sample	Skin tests (SPT; IDT ^f); sIgE	Penicillins	Immediate and non-immediate
Holm A, 2011	Denmark	580 (76.6)	Children and adults	Outpatients	Consecutive sample ^h	sIgE	Penicillins	Immediate and non-immediate
Kamboj S, 2011	USA	60 ^b	Children	Patients (not specified)	Convenience sample	Skin tests (SPT; IDT)	Penicillins ⁱ	Immediate and non-immediate
Moral L, 2011	Spain	16 ^b	Children	Outpatients	Consecutive sample ^g	Skin tests (SPT; IDT); sIgE	β -lactams	Immediate and non-immediate
Na HR, 2011	South Korea	6 (17)	Children	Outpatients	Convenience sample	Skin tests (IDT)	Penicillins ⁱ	Immediate
Ponvert C, 2011	France	1431 (44.7)	Children	Outpatients	Consecutive sample ^g	Skin tests (SPT; IDT)	β -lactams	Immediate and non-immediate
Richter AG, 2011	UK	47 ^b	Adults	Outpatients	Consecutive sample ^g	Skin tests (SPT; IDT); sIgE	Penicillins ⁱ	Immediate and non-immediate

Seitz CS, 2011	Germany	22 ^b	Children	Outpatients	Consecutive sample ^h	Skin tests (SPT; IDT ^f); sIgE	β -lactams	Immediate and non-immediate
Celik GE, 2012	Turkey	80 (83)	Children and adults	Patients (not specified)	Convenience sample	Skin tests (SPT; IDT ^f)	β -lactams	Immediate
Garcia Nuñez I, 2012	Spain	1032 ^b	Children and adults	Outpatients	Consecutive sample ^h	Skin tests (SPT; IDT ^f)	β -lactams	Immediate
Hjortlund J, 2012	Denmark	405 (68.6)	Adults	Outpatients	Consecutive sample ^g	Skin tests (SPT; IDT); sIgE	β -lactams	Immediate and non-immediate
Iglesias-Souto J, 2012	Spain	756 (45.0)	Children	Outpatients	Consecutive sample ^h	Skin tests (SPT; IDT ^f)	β -lactams	Immediate and non-immediate
Kopac P, 2012	Slovenia	606 (75.9)	Children and adults	Outpatients	Consecutive sample ^h	Skin tests (SPT; IDT); sIgE	Penicillins	Immediate and non-immediate
Soto L, 2012	Spain	226 (69.9)	Children and adults	Outpatients	Consecutive sample ^g	Skin tests (SPT; IDT ^f); sIgE	β -lactams	Immediate and non-immediate
Erkoçoglu M, 2013	Turkey	70 ^b	Children	General population	Cluster sample	Skin tests (SPT; IDT); sIgE	β -lactams	Immediate
Hjortlund J, 2013	Denmark	342 (70.2)	Adults	Outpatients	Consecutive sample ^h	Skin tests (SPT; IDT); sIgE	Penicillins	Immediate and non-immediate
Macy E, 2013	USA	500 (63.4)	Children and adults	Inpatients and outpatients	Consecutive sample ^g	Skin tests (SPT; IDT)	Penicillins	Immediate and non-immediate
Rimawi RH, 2013	USA	146 (60.3)	Children and adults	Inpatients	Convenience sample	Skin tests (SPT; IDT)	Penicillins	Immediate and non-immediate
Al-Ahmad M, 2014	Kuwait	124 (68.6)	Children and adults	Outpatients	Consecutive sample ^h	Skin tests (SPT; IDT); sIgE	β -lactams	Immediate and non-immediate
Bourke J, 2014	Australia	341 (79.8)	Children and adults	Outpatients	Consecutive sample ^g	Skin tests (SPT; IDT ^f)	Penicillins ⁱ	Immediate and non-immediate
Caubet JC, 2014	Switzerland	96 ^b	Children	Outpatients and ED	Convenience sample	Skin tests (IDT)	β -lactams	Non-immediate
Colas H, 2014	France	90 ^b	Children	Outpatients	Consecutive sample ^g	Skin tests (SPT; IDT)	Amoxicillin/amoxicillin-clavulanate ⁱ	Immediate and non-immediate
Fox SJ, 2014	USA	778 (47.2)	Children	Patients (not specified)	Consecutive sample ^g	Skin tests (SPT; IDT ^f)	β -lactams	^e
Picard M, 2014	Canada	563 (51.3)	Children and adults	Outpatients	Consecutive sample ^g	Skin tests (SPT; IDT)	Penicillins	Immediate and non-immediate
Trautmann A, 2014	Germany	37 (82)	Adults	Outpatients	Convenience sample	Skin tests (IDT); sIgE	Ampicillin	Non-immediate
Zambonino MA, 2014	Spain	783 (52.4)	Children	Patients (not specified)	Consecutive sample ^h	Skin tests (SPT; IDT); sIgE	β -lactams	Immediate and non-immediate
Barni S, 2015	Italy	352 (50.3)	Children	Outpatients	Convenience sample	Skin tests (SPT; IDT ^f)	Amoxicillin	Non-immediate
Blanca-Lopez N, 2015	Spain	261 ^b	Children and adults	Outpatients	Consecutive sample ^h	Skin tests (SPT; IDT)	Amoxicillin/amoxicillin-clavulanate	Immediate
Ferré-Ybarz L, 2015	Spain	100 (52.0)	Children and adults	Outpatients	Consecutive sample ^h	Skin tests; sIgE	β -lactams	Immediate and non-immediate
Mori F, 2015	Italy	190 (47.9)	Children	Outpatients	Consecutive sample ^h	Skin tests (SPT; IDT); sIgE	Amoxicillin ⁱ	Immediate and non-immediate
Rosenfield L, 2015	Canada	240 (72.9)	Children and adults	Outpatients	Convenience sample	Skin tests (SPT; IDT ^f)	β -lactams	Immediate and non-immediate
Abrams EM, 2016	Canada	306 (51.0)	Children and adults	Outpatients	Consecutive sample ^h	Skin tests (IDT)	β -lactams	Immediate and non-immediate
Atanaskovic-Markovic M, 2016	Serbia	1026 (48.9)	Children	Outpatients	Consecutive sample ^h	Skin tests (IDT)	β -lactams	Non-immediate
Chen JR, 2016	USA	228 (60.1)	Children and adults	Inpatients	Consecutive sample ^g	Skin tests (SPT; IDT)	Penicillins	Immediate and non-immediate

Chiriac AM, 2016	France	2821 ^b	Children and adults	Outpatients	Consecutive sample ^g	Skin tests	β -lactams	Immediate and non-immediate
King EA, 2016	USA	50 (54)	Adults	Inpatients	Consecutive sample ^g	Skin tests (SPT; IDT)	β -lactams	Immediate and non-immediate
Manuyakorn W, 2016	Thailand	126 (33.3)	Children	Outpatients and ED	Consecutive sample ^g	Skin tests (SPT; IDT)	β -lactams	Immediate and non-immediate
Mawhirt SL, 2016	USA	49 ^b	Adults	Inpatients and outpatients	Convenience sample	Skin tests (SPT; IDT ^f)	Penicillins ⁱ	Immediate
Meng J, 2016	UK	84 (66)	^c	Outpatients	Consecutive sample ^h	Skin tests (SPT; IDT); sIgE	Penicillins	Immediate and non-immediate
Mill C, 2016	Canada	17 (41)	Children	Outpatients	Consecutive sample ^h	Skin tests (SPT; IDT)	Amoxicillin	Immediate
Moreno E, 2016	Spain	1779 (65.2)	Children and adults	Outpatients	Consecutive sample ^g	Skin tests (SPT; IDT ^f)	β -lactams	Immediate
Mota I, 2016	Portugal	234 (68.4)	Children and adults	Outpatients	Consecutive sample ^h	Skin tests (SPT; IDT)	β -lactams	Immediate and non-immediate
Ratzon R, 2016	Israel	50 (62)	Children and adults	Outpatients	Convenience sample	Skin tests (SPT; IDT ^f)	β -lactams	Immediate and non-immediate
Confino-Cohen R, 2017	Israel	642 (51.4)	Children and adults	Outpatients	Consecutive sample ^h	Skin tests (IDT)	Penicillins	Non-immediate
Faitelson Y, 2017	Israel	133 (47.4)	Children	Outpatients	Convenience sample	Skin tests (SPT)	Amoxicillin/amoxicillin-clavulanate	Immediate and non-immediate
Fazlollahi MR, 2017	Iran	51 (57)	Children and adults	Outpatients	Consecutive sample ^h	Skin tests (IDT)	β -lactams	Immediate and non-immediate
Juri MC, 2017	Argentina	178 ^b	Adults	Outpatients	Consecutive sample ^g	Skin tests	Penicillins ⁱ	^e
Marwood J, 2017	Australia	103 (54.4)	Adults	ED	Convenience sample	Skin tests (SPT; IDT)	Penicillins	Immediate and non-immediate
McDanel DL, 2017	USA	161 (64.6)	Adults	Outpatients	Convenience sample	Skin tests (SPT; IDT)	β -lactams	Immediate and non-immediate
Misirlioglu ED, 2017	Turkey	346 ^b	Children	Outpatients	Consecutive sample ^h	Skin tests (SPT; IDT)	β -lactams	Immediate
Ramsey A, 2017	USA	50 (54)	Adults	Inpatients	Convenience sample	Skin tests (SPT; IDT ^f)	Penicillins	Immediate and non-immediate
Soria A, 2017	France	325 (64.0)	Children and adults	Outpatients	Consecutive sample ^h	Skin tests (SPT; IDT)	Penicillins	Immediate and non-immediate
Sundquist BK, 2017	USA	37 (51)	Adults	Outpatients	Convenience sample	Skin tests (SPT; IDT)	Penicillin; amoxicillin/amoxicillin-clavulanate	Immediate and non-immediate
Tannert LK, 2017	Denmark	21 (71)	Adults	Outpatients	Convenience sample	Skin tests (IDT); sIgE	Penicillins	Immediate and non-immediate
Vyles D, 2017	USA	100 (35.0)	Children	ED	Convenience sample	Skin tests (SPT; IDT)	Penicillins	Immediate and non-immediate
Ibáñez MD, 2018	Spain	732 (48.8)	Children	Outpatients and ED	Consecutive sample ^h	Skin tests (SPT; IDT); sIgE	Penicillin; amoxicillin-clavulanate; cloxacillin	Immediate and non-immediate
Marraccini P, 2018	Italy	37 ^b	Adults	Outpatients	Convenience sample	Skin tests (SPT; IDT ^f); sIgE	Penicillins ⁱ	Immediate and non-immediate
Moussa Y, 2018	Canada	194 (66.5)	Adults	Outpatients	Consecutive sample ^h	Skin tests (SPT; IDT ^f)	β -lactams	^e
Sedlackova L, 2018	Czech Republic	24 (79)	Children and adults	Patients (not specified)	Convenience sample	Skin tests; sIgE	Penicillin; amoxicillin/amoxicillin-clavulanate	Non-immediate

^a For each primary study, only diagnostic tests whose accuracy could be calculated are listed; ^b Sex discrimination not available; ^c Information on patients' age group not available; ^d Consecutive sample consisting of all patients with history of penicillin allergy and in whom penicillin is indicated; ^e Information not provided; ^f Skin prick tests and intradermal tests were not presented separately; ^g Consecutive sample consisting of all participants tested for penicillin allergy; ^h Consecutive sample consisting of all participants with history/events compatible with penicillin allergy; ⁱ Participants with reported cephalosporin allergy are not being counted; ^j Consecutive sample consisting of all patients observed by seven clinical specialties of a hospital.

Table E3. Description *per* study of (i) the number of participants performing each diagnostic index test and drug provocation tests (with their respective results), and (ii) of the corresponding diagnostic measures (sensitivity, specificity and predictive values)

A. Skin tests														
Reference	Assessed drugs/determinants	N participants performing skin tests	Participants with positive skin tests				Participants with negative skin tests				Sn (%)	Sp (%)	PPV (%)	NPV (%)
			Total N (%)	No DPT performed	Positive DPT (TP)	Negative DPT (FP)	Total N (%)	No DPT performed	Positive DPT (FN)	Negative DPT (TN)				
Brown BC, 1964	PPL	1085	416 (38.3)	383	9	24	669 (61.7)	669	0	0	-	-	27.3	-
Resnik SS, 1965	PPL	16	6 (38)	0	0	6	10 (62)	0	0	10	-	62.5	0	100
Levine BB, 1968*	PPL; MDM	218	32 (14.7) ^a	21	8	3	185 (84.9) ^a	0	1	184	88.9	98.4	72.7	99.5
Wide L, 1971	PPL; PEN-G	23	9 (39)	9	0	0	14 (61)	0	0	14	-	-	-	100
Warrington RJ, 1978*	PPL; MDM; Unspecified aminopenicillins	253	29 (11.5)	23	0	6	224 (88.5)	55	2	167	0	96.5	0	98.8
Basomba A, 1979	PPL; PEN-G; AMP	81	65 (80)	65	0	0	16 (20)	8	2	6	-	-	-	75.0
Sullivan TJ, 1981	PPL; MDM; PEN-G; AMP; CAR; CEF	290	144 (49.7)	144	0	0	146 (50.3)	63	3	80	-	-	-	96.4
Van Arsdel PP, 1986	PPL; MDM	97	12 (12) ^a	12	0	0	79 (81) ^a	12	2	65	-	-	-	97.0
Graff-Lonnevig V, 1988	PEN-G; PEN-V	298	0 (0)	-	-	-	298 (100)	1	30	267	-	-	-	89.9
Blanca M, 1990*	PPL; MDM; AMX; AMP	278	36 (12.9)	34	2	0	242 (87.1)	1	19	222	9.5	100	100	92.1
Sogn DD, 1992*	PPL; MDM	825	146 (17.7) ^a	138	4	4	656 (79.5) ^a	90	7	559	36.4	99.3	50.0	98.8
Gadde J, 1993	PPL; MDM	776	55 (7.1) ^a	55	0	0	700 (90.2) ^a	51	34	615	-	-	-	94.8
Juchet A, 1994	PPL; MDM	45	17 (38)	17	0	0	28 (62)	17	2	9	-	-	-	81.8
Romano A, 1995*	PPL; MDM; PEN-G; AMP	60	33 (55)	15	18	0	27 (45)	11	0	16	100	100	100	100
Sastre J, 1996*	PPL; MDM; PEN-G; AMX	576	71 (12.3)	69	2	0	505 (87.7)	0	5	500	28.6	100	100	99.0
Macy E, 1997	PPL; MDM; PEN-G; AMX	348	60 (17.2)	60	0	0	288 (82.8)	73	13	202	-	-	-	94.0
Cederbrant K, 1998*	PPL; MDM; PEN-G; PEN-V; AMP; CLO; BAC	17	4 (24)	3	1	0	13 (76)	2	5	6	16.7	100	100	54.5
Hervé M, 1998	PPL; MDM; AMX; AMX-CLV	108	39 (36.1)	39	0	0	69 (63.9)	17	6	46	-	-	-	88.5
Loza Cortina C, 1998*	PPL; MDM; PEN-G; AMX; AMP	21	4 (19)	1	3	0	17 (81)	1	1	15	75.0	100	100	93.8
Pichichero ME, 1998	PPL; MDM; PEN-G ^b	247	53 (21.5)	53	0	0	194 (78.5)	0	5	189	-	-	-	97.4

Forrest DM, 2001*	Unspecified	72	4 (6)	3	0	1	68 (94)	7	1	60	0	98.4	0	98.4
López Tiro JJ, 2001	PPL; MDM; PEN-G	24	23 (96)	23	0	0	1 (4)	0	1	0	-	-	-	0
Novalbos A, 2001	PPL; MDM; PEN-G; AMX ^c	958	38 (4.0)	38	0	0	920 (96.0)	0	3	917	-	-	-	99.7
Romano A, 2002*	PPL; MDM; PEN-G; AMX; AMP	259	105 (40.5)	100	2	3	154 (59.5)	29	3	122	40.0	97.6	40.0	97.6
Solensky R, 2002	PPL; MDM; PEN-G	58	5 (9)	5	0	0	53 (91)	4	1	48	-	-	-	98.0
Torres MJ, 2002	MDM; PEN-G; AMX; AMP	330	203 (61.5)	203	0	0	127 (38.5)	38	49	40	-	-	-	44.9
Borch JE, 2006	PPL; MDM; PEN-G; AMP	13	1 (8)	1	0	0	12 (92)	0	0	12	-	-	-	100
Wong BBL, 2006	PPL; MDM; PEN-G; AMP	91	16 (18)	16	0	0	75 (82)	3	2	70	-	-	-	97.2
del Real GA, 2007	PPL; PEN-G	596	49 (8.2) ^a	49	0	0	527 (88.4) ^a	237	5	285	-	-	-	98.3
Rebello Gomes E, 2007	Unspecified	16	0 (0)	-	-	-	16 (100)	0	1	15	-	-	-	93.8
Goldberg A, 2008*	PPL; MDM; PEN-G; AMX; AMP	169	73 (43.2)	6	7	60	96 (56.8)	2	4	90	63.6	60.0	10.4	95.7
Nolan RC, 2008	PPL; MDM	29	4 (14)	4	0	0	25 (86)	0	4	21	-	-	-	84.0
Padial A, 2008	MDM; PEN-G; AMX; AMP; CLO; Unspecified cephalosporins	18	2 (11)	0	2	0	16 (89)	0	16	0	11.1	-	100	0
Rodriguez-Alvarez M, 2008	PPL; MDM; Unspecified penicillins; AMX; AMP; Unspecified cephalosporins	23	1 (4)	1	0	0	22 (95.7)	1	4	17	-	-	-	81.0
Tantikul C, 2008*	AMX; LEX	4	1 (25)	0	1	0	3 (75)	0	0	3	100	100	100	100
Hershkovich J, 2009	PPL; MDM; PEN-G	166	4 (2.4)	4	0	0	162 (97.6)	0	6	156	-	-	-	96.3
Silva R, 2009	PPL; MDM; AMX; AMP; Suspected drug	67	6 (9)	6	0	0	61 (91)	28	2	31	-	-	-	93.9
Chambel M, 2010	PPL; MDM; Unspecified penicillins; AMX; AMX- CLV; CXM	47	8 (17)	8	0	0	39 (83)	9	1	29	-	-	-	96.7
Macy E, 2010	MDM; AMX	150	6 (4.0)	6	0	0	144 (96.0)	4	3	137	-	-	-	97.9
Kamboj S, 2011	PEN-G; AMP; Unspecified cephalosporins	60	4 (7)	4	0	0	56 (93)	0	4	52	-	-	-	92.9
Moral L, 2011	PPL; MDM; Unspecified penicillins; Suspected drug	16	0 (0)	-	-	-	16 (100)	0	0	16	-	-	-	100
Na HR, 2011	Suspected drug	5	2 (40)	0	2	0	3 (60)	0	3	0	40.0	-	100	0
Ponvert C, 2011	Suspected drug	1431	99 (6.9)	99	0	0	1332 (93.1)	17	111	1204	-	-	-	91.6
Richter AG, 2011*	PPL; MDM; AMX ^d	47	8 (17)	6	1	1	39 (83)	6	5	28	16.7	96.6	50.0	84.8

Seitz CS, 2011	Unspecified	22	1 (5)	1	0	0	21 (95)	0	0	21	-	-	-	100
Celik GE, 2012	PPL; MDM	80	24 (30)	24	0	0	56 (70)	16	5	35	-	-	-	87.5
Garcia Nuñez I, 2012	PPL; MDM; PEN-G; AMX-CLV; AMP	1004	79 (7.9) ^e	79	0	0	898 (89.4) ^e	0	23	875	-	-	-	97.4
Hjortlund J, 2012	PEN-G; AMP	405	21 (5.2)	21	0	0	384 (94.8)	44	46	294	-	-	-	86.5
Iglesias-Souto J, 2012	PPL; MDM; PEN-G; AMX; CFZ	756	19 (2.5)	19	0	0	737 (97.5)	0	3	734	-	-	-	99.6
Kopac P, 2012	PPL; MDM; Suspected drug	274	14 (5.1)	14	0	0	260 (94.9)	0	21	239	-	-	-	91.9
Soto L, 2012	Unspecified	226	39 (17.3)	39	0	0	187 (82.7)	46	4	137	-	-	-	97.2
Erkoçoglu M, 2013	PPL; MDM; PEN-G; AMX; AMP; CEC; Unspecified cephalosporins	70	3 (4)	3	0	0	67 (96)	5	0	62	-	-	-	100
Hjortlund J, 2013	PPL; MDM	342	35 (10.2)	35	0	0	307 (89.8)	16	50	241	-	-	-	82.8
Macy E, 2013	PPL; MDM; AMX	500	4 (0.8)	4	0	0	496 (99.2)	0	4	492	-	-	-	99.2
Rimawi RH, 2013	PPL; PEN-G	146	1 (0.7)	1	0	0	145 (99.3)	0	0	145	-	-	-	100
Al-Ahmad M, 2014	PPL; MDM; AMX; AMP	124	44 (35.5)	44	0	0	80 (64.5)	1	1	78	-	-	-	98.7
Bourke J, 2014	PPL; MDM; PEN-G; AMX; CRO; CFZ; Indicated drug	341	42 (12.3)	42	0	0	299 (87.7)	0	7	292	-	-	-	97.7
Caubet JC, 2014*	PPL; MDM; AMX; CXM ^f	96	14 (15)	0	7	7	82 (85)	0	7	75	50.0	91.5	50.0	91.5
Colas H, 2014	Unspecified penicillins; Unspecified cephalosporins	90	2 (2)	2	0	0	88 (98)	0	2	86	-	-	-	97.7
Fox SJ, 2014	PPL; PEN-G; AMX	778	66 (8.5) ^a	66	0	0	703 (90.4) ^a	334	14	355	-	-	-	96.2
Picard M, 2014	PEN-G	563	185 (32.9)	185	0	0	378 (67.1)	3	18	357	-	-	-	95.2
Trautmann A, 2014	PEN-G; AMP; CRO	37	10 (27)	10	0	0	27 (73)	0	1	26	-	-	-	96.3
Zambonino MA, 2014	PPL; MDM; AMX; AMX-CLV; CXM	781	4 (0.5)	4	0	0	777 (99.5)	0	56	721	-	-	-	92.8
Barni S, 2015*	AMX	352	3 (0.9)	0	2	1	349 (99.1)	14	23	312	8.0	99.7	66.7	93.1
Blanca-Lopez N, 2015	PPL; MDM; AMX; CLV	261	29 (11.1)	29	0	0	232 (88.9)	0	11	221	-	-	-	95.3
Ferré-Ybarz L, 2015	PPL; MDM; AMX; CXM ^g	82	3 (4)	3	0	0	79 (96)	2	1	76	-	-	-	98.7
Rosenfield L, 2015*	PPL; MDM; PEN-G;	240	17 (7.1)	14	1	2	223	1	12	210	7.7	99.1	33.3	94.6

	Suspected drug						(92.9)							
Abrams EM, 2016	PPL; PEN-G; AMP	106	1 (0.9)	1	0	0	105 (99.1)	0	5	100	-	-	-	95.2
Atanaskovic-Markovic M, 2016	PPL; MDM; PEN-G; AMX; AMP; Suspected drug ^h	1026	61 (5.9)	61	0	0	965 (94.1)	0	19	946	-	-	-	98.0
Chen JR, 2016	PPL; MDM; PEN-G	228	4 (1.8)	4	0	0	224 (98.2)	0	1	223	-	-	-	99.6
Chiriac AM, 2016	PPL; MDM; AMX; AMP; Suspected drug	2821	466 (16.5)	466	0	0	2355 (83.5)	0	171	2184	-	-	-	92.7
King EA, 2016	PPL; PEN-G	50	0 (0)	-	-	-	50 (100)	0	1	49	-	-	-	98.0
Manuyakorn W, 2016	PEN-G; AMX-CLV; AMP; Suspected drug	126	12 (9.5)	12	0	0	114 (90.5)	0	10	104	-	-	-	91.2
Mawhirt SL, 2016*	PPL; MDM; Drugs received on DPT	49	4 (8)	3	0	1	45 (92)	3	5	37	0	97.4	0	88.1
Meng J, 2016*	PPL; MDM; PEN-G; AMX ⁱ	84	22 (26)	19	0	3	62 (74)	0	2	60	0	95.2	0	96.8
Mill C, 2016	PPL; PEN-G	17	1 (6)	0	1	0	16 (94)	0	16	0	5.9	-	100	0
Moreno E, 2016	PPL; MDM; PEN-G; AMX	1779	403 (22.7)	403	0	0	1376 (77.3)	12	42	1322	-	-	-	96.9
Mota I, 2016	PPL; MDM; PEN-G; AMX; CLV; CXM; Suspected drug	219	30 (13.7)	30	0	0	189 (86.3)	0	5	184	-	-	-	97.4
Ratzon R, 2016	Suspected drug	50	0 (0)	0	0	0	50 (100)	0	3	47	-	-	-	94.0
Confino-Cohen R, 2017*	PPL; MDM; PEN-G; AMX; Suspected drug	642	34 (5.3) ^a	4	2	28	400 (62.3) ^a	10	40	350	4.8	92.6	6.7	89.7
Faitelson Y, 2017	PEN-G; AMX	128	0 (0)	0	0	0	128 (100)	0	10	118	-	-	-	92.2
Fazlollahi MR, 2017	PPL; MDM; AMX	44	8 (18)	8	0	0	36 (82)	22	1	13	-	-	-	92.9
Juri MC, 2017	Unspecified	178	33 (18.5)	33	0	0	145 (81.5)	18	13	114	-	-	-	89.8
Marwood J, 2017	PPL; MDM; AMX	103	16 (15.5)	16	0	0	87 (84.5)	3	3	81	-	-	-	96.4
McDanel DL, 2017	PPL; PEN-G	140	0 (0) ^a	0	0	0	139 (99.3) ^a	120	0	19	-	-	-	100
Misirlioglu ED, 2017	PPL; MDM; PEN-G; AMX-CLV; Suspected drug	346	28 (8.1)	28	0	0	318 (91.9)	0	17	301	-	-	-	94.7
Ramsey A, 2017	PPL; PEN-G	50	2 (4) ^a	2	0	0	47 (94) ^a	19	0	28	-	-	-	100
Soria A, 2017	Unspecified penicillins; AMX; AMX-CLV; OXA; CLO	259	33 (12.7)	33	0	0	226 (87.3)	1	7	218	-	-	-	96.9
Sundquist BK, 2017	PPL; MDM	37	0 (0)	0	0	0	37 (100)	1	0	36	-	-	-	100
Tannert LK, 2017	Suspected drug	21	21 (100)	9	7	5	0 (0)	0	0	0	-	-	58.3	-
Vyles D, 2017	PPL; MDM	100	3 (3.0)	0	0	3	97 (97.0)	0	0	97	-	97.0	0	100

Ibáñez MD, 2018*	PPL; MDM; PEN-G; PEN-G; AMX; AMX-CLV	726	15 (2.1)	0	3	12	711 (97.9)	0	30	681	9.1	98.3	20.0	95.8
Marraccini P, 2018	PPL; MDM; AMX	26	1 (4)	1	0	0	25 (96)	0	1	24	-	-	-	96.0
Moussa Y, 2018	PPL	194	4 (2.1)	4	0	0	190 (97.9)	44	7	139	-	-	-	95.2
Sedlackova L, 2018	PPL; MDM; PEN-G; AMX-CLV	21	5 (24)	5	0	0	16 (76)	6	1	9	-	-	-	90.0
A.1. Skin tests (immediate reactions)														
Reference	Assessed drugs/determinants	N participants performing skin tests	Participants with positive skin tests				Participants with negative skin tests				Sn (%)	Sp (%)	PPV (%)	NPV (%)
			Total N (%)	No DPT performed	Positive DPT (TP)	Negative DPT (FP)	Total N (%)	No DPT performed	Positive DPT (FN)	Negative DPT (TN)				
Wide L, 1971	PPL; PEN-G	23	9 (39)	9	0	0	14 (61)	0	0	14	-	-	-	100
Sastre J, 1996	PPL; MDM; PEN-G; AMX	576	71 (12.3)	69	2	0	505 (87.7)	0	5	500	28.6	100	100	99.0
Cederbrant K, 1998	PPL; MDM; PEN-G; PEN-V; AMP; CLO; BAC	8	2 (25)	2	0	0	6 (75)	1	4	1	-	-	-	20.0
Torres MJ, 2002	MDM; PEN-G; AMX; AMP	330	203 (61.5)	203	0	0	127 (38.5)	38	49	40	-	-	-	44.9
Rebelo Gomes E, 2007	Unspecified	2	0 (0)	-	-	-	2 (100)	0	0	2	-	-	-	100
Bousquet PJ, 2008	PPL; MDM; AMX; AMP; Suspected drug	205	74 (36.1)	74	0	0	131 (63.9)	0	20	111	-	-	-	84.7
Tantikul C, 2008	AMX; LEX	2	0 (0)	-	-	-	2 (100)	0	0	2	-	-	-	100
Na HR, 2011	Suspected drug	5	2 (40)	0	2	0	3 (60)	0	3	0	40.0	-	100	0
Ponvert C, 2011	Suspected drug	162	43 (26.5)	43	0	0	119 (73.5)	0	7	112	-	-	-	94.1
Seitz CS, 2011	Unspecified	10	0 (0)	-	-	-	10 (100)	0	0	10	-	-	-	100
Celik GE, 2012	PPL; MDM	80	24 (30)	24	0	0	56 (70)	16	5	35	-	-	-	87.5
Garcia Nuñez I, 2012	PPL; MDM; PEN-G; AMX-CLV; AMP	1004	79 (7.9) ^e	79	0	0	898 (89.4) ^e	0	23	875	-	-	-	97.4
Erkoçoglu M, 2013	PPL; MDM; PEN-G; AMX; AMP; CEC; Unspecified cephalosporins	70	3 (4)	3	0	0	67 (96)	5	0	62	-	-	-	100
Hjortlund J, 2013	PPL; MDM	36	5 (14)	5	0	0	31 (86)	1	11	19	-	-	-	63.3
Macy E, 2013	PPL; MDM; AMX	52	1 (2)	1	0	0	51 (98)	0	1	50	-	-	-	98.0
Zambonino MA, 2014	PPL; MDM; AMX; AMX-CLV; CXM	64	2 (3)	2	0	0	62 (97)	0	5	57	-	-	-	91.9
Blanca-Lopez N, 2015	PPL; MDM; AMX; CLV	261	29 (11.1)	29	0	0	232 (88.9)	0	11	221	-	-	-	95.3
Mori F, 2015	AMX	38	3 (8)	2	1	0	35 (92)	0	2	33	33.3	100	100	94.3
Manuyakorn W,	PEN-G; AMX-CLV;	40	9 (23)	9	0	0	31 (78)	0	6	25	-	-	-	80.6

2016	AMP; Suspected drug													
Mawhirt SL, 2016	PPL; MDM; Drugs received on DPT	49	4 (8)	3	0	1	45 (92)	3	5	37	0	97.4	0	88.1
Mill C, 2016	PPL; PEN-G	17	1 (6)	0	1	0	16 (94)	0	16	0	5.9	-	100	0
Moreno E, 2016	PPL; MDM; PEN-G; AMX	1779	403 (22.7)	403	0	0	1376 (77.3)	12	42	1322	-	-	-	96.9
Faitelson Y, 2017	PEN-G; AMX	6	0 (0)	-	-	-	6 (100)	0	0	6	-	-	-	100
Misirlioglu ED, 2017	PPL; MDM; PEN-G; AMX-CLV; Suspected drug	346	28 (8.1)	28	0	0	318 (91.9)	0	17	301	-	-	-	94.7
Tannert LK, 2017	Suspected drug	17	17 (100)	7	6	4	0 (0)	-	-	-	-	-	60.0	-
A.2. Skin tests (non-immediate reactions)														
Reference	Assessed drugs/determinants	N participants performing skin tests	Participants with positive skin tests				Participants with negative skin tests				Sn (%)	Sp (%)	PPV (%)	NPV (%)
			Total N (%)	No DPT performed	Positive DPT (TP)	Negative DPT (FP)	Total N (%)	No DPT performed	Positive DPT (FN)	Negative DPT (TN)				
Romano A, 1995*	PPL; MDM; PEN-G; AMP	60	33 (55)	15	18	0	27 (45)	11	0	16	100	100	100	100
Cederbrant K, 1998*	PPL; MDM; PEN-G; PEN-V; AMP; CLO; BAC	9	2 (22)	1	1	0	7 (78)	1	1	5	50.0	100	100	83.3
Romano A, 2002*	PPL; MDM; PEN-G; AMX; AMP	259	105 (40.5)	100	2	3	154 (59.5)	29	3	122	40.0	97.6	40.0	97.6
Rebello Gomes E, 2007	Unspecified	14	0 (0)	-	-	-	14 (100)	0	1	13	-	-	-	92.9
Bousquet PJ, 2008	PPL; MDM; AMX; AMP; Suspected drug	658	71 (10.8)	71	0	0	587 (89.2)	0	36	551	-	-	-	93.9
Padial A, 2008	MDM; PEN-G; AMX; AMP; CLO; Unspecified cephalosporins	18	2 (11)	0	2	0	16 (89)	0	16	0	11.1	-	100	0
Rodriguez-Alvarez M, 2008	PPL; MDM; Unspecified penicillins; AMX; AMP; Unspecified cephalosporins	23	1 (4)	1	0	0	22 (96)	1	4	17	-	-	-	81.0
Tantikul C, 2008*	AMX; LEX	2	1 (50)	0	1	0	1 (50)	0	0	1	100	100	100	100
Ponvert C, 2011	Suspected drug	1269	56 (4.4)	56	0	0	1213 (95.6)	17	104	1092	-	-	-	91.3
Seitz CS, 2011	Unspecified	12	1 (8)	1	0	0	11 (92)	0	0	11	-	-	-	100
Hjortlund J, 2013	PPL; MDM	235	24 (10.2)	24	0	0	211 (89.8)	9	36	166	-	-	-	82.2
Macy E, 2013	PPL; MDM; AMX	296	3 (1.0)	3	0	0	293 (99.0)	0	3	290	-	-	-	99.0
Caubet JC, 2014*	PPL; MDM; AMX; CXM ^f	96	14 (15)	0	7	7	82 (85)	0	7	75	50.0	91.5	50.0	91.5
Trautmann A, 2014	PEN-G; AMP; CRO	37	10 (27)	10	0	0	27 (73)	0	1	26	-	-	-	96.3
Zambonino MA,	PPL; MDM; AMX; AMX-	717	2 (0.3)	2	0	0	715	0	51	664	-	-	-	92.9

2014	CLV; CXM						(99.7)							
Barni S, 2015*	AMX	352	3 (0.9)	0	2	1	349 (99.1)	14	23	312	8.0	99.7	66.7	93.1
Atanaskovic-Markovic M, 2016	PPL; MDM; PEN-G; AMX; AMP; Suspected drug ^h	1026	61 (5.9)	61	0	0	965 (94.1)	0	19	946	-	-	-	98.0
Manuyakorn W, 2016	PEN-G; AMX-CLV; AMP; Suspected drug	86	3 (4)	3	0	0	83 (96)	0	4	79	-	-	-	95.2
Confino-Cohen R, 2017*	PPL; MDM; PEN-G; AMX; Suspected drug	642	34 (5.3) ^a	4	2	28	400 (62.3) ^a	10	40	350	4.8	92.6	6.7	89.7
Faitelson Y, 2017	PEN-G; AMX	101	0 (0)	-	-	-	101 (100)	0	10	91	-	-	-	90.1
Sedlackova L, 2018	PPL; MDM; PEN-G; AMX-CLV	21	5 (24)	5	0	0	16 (76)	6	1	9	-	-	-	90.0

B. Skin prick tests

Reference	Assessed drugs/determinants	N participants performing skin prick tests	Participants with positive skin prick tests				Participants with negative skin prick tests				Sn (%)	Sp (%)	PPV (%)	NPV (%)
			Total N (%)	No DPT performed	Positive DPT (TP)	Negative DPT (FP)	Total N (%)	No DPT performed	Positive DPT (FN)	Negative DPT (TN)				
Basomba A, 1979	PPL; PEN-G; AMP	81	31 (38)	31	0	0	50 (62)	35	5	10	-	-	-	66.7
Sullivan TJ, 1981	PPL; MDM; PEN-G; AMP; CAR; CEF	290	34 (12)	34	0	0	256 (88.3)	173	3	80	-	-	-	96.4
Graff-Lonnevig V, 1988	PEN-G; PEN-V	298	0 (0)	-	-	-	298 (100)	1	30	267	-	-	-	89.9
Romano A, 1995	PPL; MDM; PEN-G; AMP	18	0 (0)	-	-	-	18 (100)	0	18	0	-	-	-	0
Macy E, 1997	PPL; MDM; PEN-G; AMX	348	6 (1.7)	6	0	0	342 (98.3)	127	13	202	-	-	-	94.0
Cederbrant K, 1998	PPL; MDM; PEN-G; PEN-V; AMP; CLO; BAC	17	1 (6)	0	1	0	16 (94)	5	5	6	16.7	100	100	54.5
Hervé M, 1998	PPL; MDM; AMX; AMX-CLV	108	3 (2.8)	3	0	0	105 (97.2)	53	6	46	-	-	-	88.5
López Tiro JJ, 2001	PPL; MDM; PEN-G	24	12 (50)	12	0	0	12 (50)	11	1	0	-	-	-	0
Novalbos A, 2001	PPL; MDM; PEN-G; AMX ^c	958	8 (0.8)	8	0	0	950 (99.2)	30	3	917	-	-	-	99.7
Borch JE, 2006	PPL; MDM; PEN-G; AMP; AMX	13	0 (0)	-	-	-	13 (100)	1	0	12	-	-	-	100
Bousquet PJ, 2007	PPL; MDM; AMX; AMP; Suspected drug	973	21 (2.2)	21	0	0	952 (97.8)	120	69	763	-	-	-	91.7
Rebelo Gomes E, 2007	Unspecified	16	0 (0)	-	-	-	16 (100)	0	1	15	-	-	-	93.8
Nolan RC, 2008	PPL; MDM; AMX	29	0 (0)	-	-	-	29 (100)	4	4	21	-	-	-	84.0
Rodriguez-Alvarez M, 2008	PPL; MDM; Unspecified penicillins; AMX; AMP; Unspecified	23	0 (0)	-	-	-	23 (100)	2	4	17	-	-	-	81.0

Tantikul C, 2008	cephalosporins AMX; LEX	4	1 (25)	0	1	0	3 (75)	0	0	3	100	100	100	100
Silva R, 2009	PPL; MDM; AMX; AMP; Suspected drug	67	3 (5)	3	0	0	64 (95)	31	2	31	-	-	-	93.9
Chambel M, 2010	PPL; MDM; Unspecified penicillins; AMX; AMX- CLV; CXM	47	0 (0)	-	-	-	47 (100)	17	1	29	-	-	-	96.7
Kamboj S, 2011	PEN-G; AMP; Unspecified cephalosporins	60	3 (5)	3	0	0	57 (95)	1	4	52	-	-	-	92.9
Moral L, 2011	PPL; MDM; Unspecified penicillins; Suspected drug	16	0 (0)	-	-	-	16 (100)	0	0	16	-	-	-	100
Ponvert C, 2011	Suspected drug	1431	2 (0.1)	2	0	0	1429 (99.9)	114	111	1204	-	-	-	91.6
Richter AG, 2011	PPL; MDM; AMX ^d	47	1 (2)	1	0	0	46 (98)	5	6	35	-	-	-	85.4
Hjortlund J, 2012	PEN-G; AMP; AMX	405	1 (0.2)	1	0	0	404 (99.8)	64	46	294	-	-	-	86.5
Kopac P, 2012	PPL; MDM; Suspected drug	274	6 (2.2)	6	0	0	268 (97.8)	8	19	241	-	-	-	92.7
Erkoçoglu M, 2013	PPL; MDM; PEN-G; AMX; AMP; CEC; Unspecified cephalosporins	70	2 (3)	2	0	0	68 (97)	6	0	62	-	-	-	100
Hjortlund J, 2013	PPL; MDM; PEN-V; AMX	342	1 (0.3)	1	0	0	341 (99.7)	50	50	241	-	-	-	82.8
Macy E, 2013	PPL; MDM; AMX	500	0 (0)	-	-	-	500 (100)	4	4	492	-	-	-	99.2
Rimawi RH, 2013	PPL; PEN-G	146	0 (0)	-	-	-	146 (100)	1	0	145	-	-	-	100
Al-Ahmad M, 2014	PPL; MDM; AMX; AMP	124	15 (12.1)	15	0	0	109 (87.9)	30	1	78	-	-	-	98.7
Colas H, 2014	Unspecified penicillins; Unspecified cephalosporins	90	0 (0)	-	-	-	90 (100)	2	2	86	-	-	-	97.7
Picard M, 2014	PEN-G	563	1 (0.2)	1	0	0	562 (99.8)	187	18	357	-	-	-	95.2
Zambonino MA, 2014	PPL; MDM; AMX; AMX- CLV; CXM	781	1 (0.1)	1	0	0	780 (99.9)	3	56	721	-	-	-	92.8
Blanca-Lopez N, 2015	PPL; MDM; AMX; CLV	261	9 (3.4)	9	0	0	252 (96.6)	20	11	221	-	-	-	95.3
Mori F, 2015	AMX	190	1 (0.5)	1	0	0	189 (99.5)	12	17	160	-	-	-	90.4
Chen JR, 2016	PPL; MDM; PEN-G	228	4 (1.8)	4	0	0	224 (98.2)	0	1	223	-	-	-	99.6
King EA, 2016	PPL; PEN-G	50	0 (0)	-	-	-	50 (100)	0	1	49	-	-	-	98.0
Manuyakorn W,	PEN-G; AMX-CLV;	126	4 (3.2)	4	0	0	122	8	10	104	-	-	-	91.2

2016	AMP; Suspected drug						(96.8)							
Meng J, 2016	PPL; MDM; PEN-G; AMX ¹	84	4 (5)	4	0	0	80 (95)	18	2	60	-	-	-	96.8
Mill C, 2016	PPL; PEN-G	17	1 (6)	0	1	0	16 (94)	0	16	0	5.9	-	100	0
Mota I, 2016	PPL; MDM; PEN-G; AMX; CLV; CXM; Suspected drug	219	5 (2.3)	5	0	0	214 (97.7)	25	5	184	-	-	-	97.4
Faitelson Y, 2017	PEN-G; AMX	128	0 (0)	-	-	-	128 (100)	0	10	118	-	-	-	92.2
Marwood J, 2017	PPL; MDM; AMX	103	2 (1.9)	2	0	0	101 (98.1)	17	3	81	-	-	-	96.4
McDanel DL, 2017	PPL; PEN-G	140	0 (0) ^a	-	-	-	139 (99.3) ^a	120	0	19	-	-	-	100
Misirlioglu ED, 2017	PPL; MDM; PEN-G; AMX-CLV; Suspected drug	346	4 (1.2)	4	0	0	342 (98.8)	24	17	301	-	-	-	94.7
Soria A, 2017	Unspecified penicillins; AMX; AMX-CLV; OXA; CLO	259	14 (5.4)	14	0	0	245 (94.6)	20	7	218	-	-	-	96.9
Sundquist BK, 2017	PPL; MDM	37	0 (0)	-	-	-	37 (100)	1	0	36	-	-	-	100
Vyles D, 2017	PPL; MDM	100	3 (3.0)	0	0	3	97 (97.0)	0	0	97	-	97.0	0	100
Ibáñez MD, 2018	PPL; MDM; PEN-G; PEN-G; AMX; AMX-CLV	726	1 (0.1)	0	1	0	725 (99.9)	0	32	693	3.0	100	100	95.6

B.1. Skin prick tests (immediate reactions)

Reference	Assessed drugs/determinants	N participants performing skin prick tests	Participants with positive skin prick tests				Participants with negative skin prick tests				Sn (%)	Sp (%)	PPV (%)	NPV (%)
			Total N (%)	No DPT performed	Positive DPT (TP)	Negative DPT (FP)	Total N (%)	No DPT performed	Positive DPT (FN)	Negative DPT (TN)				
Cederbrant K, 1998	PPL; MDM; PEN-G; PEN-V; AMP; CLO; BAC	8	0 (0)	-	-	-	8 (100)	3	4	1	-	-	-	20.0
Rebelo Gomes E, 2007	Unspecified	2	0 (0)	-	-	-	2 (100)	0	0	2	-	-	-	100
Tantikul C, 2008	AMX; LEX	2	0 (0)	-	-	-	2 (100)	0	0	2	-	-	-	100
Ponvert C, 2011	Suspected drug	162	2 (1.2)	2	0	0	160 (98.8)	41	7	112	-	-	-	94.1
Erkoçoglu M, 2013	PPL; MDM; PEN-G; AMX; AMP; CEC; Unspecified cephalosporins	70	2 (3)	2	0	0	68 (97)	6	0	62	-	-	-	100
Hjortlund J, 2013	PPL; MDM; PEN-V; AMX	36	0 (0)	-	-	-	36 (100)	6	11	19	-	-	-	63.3
Macy E, 2013	PPL; MDM; AMX	52	0 (0)	-	-	-	52 (100)	1	1	50	-	-	-	98.0
Zambonino MA, 2014	PPL; MDM; AMX; AMX-CLV; CXM	64	1 (2)	1	0	0	63 (98)	1	5	57	-	-	-	91.9

Blanca-Lopez N, 2015	PPL; MDM; AMX; CLV	261	9 (3.4)	9	0	0	252 (96.6)	20	11	221	-	-	-	95.3
Mori F, 2015	AMX	38	1 (3)	1	0	0	37 (97)	1	3	33	-	-	-	91.7
Manuyakorn W, 2016	PEN-G; AMX-CLV; AMP; Suspected drug	40	4 (10)	4	0	0	36 (90)	5	6	25	-	-	-	80.6
Mill C, 2016	PPL; PEN-G	17	1 (6)	0	1	0	16 (94)	0	16	0	5.9	-	100	0
Faitelson Y, 2017	PEN-G; AMX	6	0 (0)	0	0	6	6 (100)	0	0	6	-	100	-	100
Misirlioglu ED, 2017	PPL; MDM; PEN-G; AMX-CLV; Suspected drug	346	4 (1.2)	4	0	0	342 (98.8)	24	17	301	-	-	-	94.7

B.2. Skin prick tests (non-immediate reactions)

Reference	Assessed drugs/determinants	N participants performing skin prick tests	Participants with positive skin prick tests				Participants with negative skin prick tests				Sn (%)	Sp (%)	PPV (%)	NPV (%)
			Total N (%)	No DPT performed	Positive DPT (TP)	Negative DPT (FP)	Total N (%)	No DPT performed	Positive DPT (FN)	Negative DPT (TN)				
Romano A, 1995	PPL; MDM; PEN-G; AMP	18	0 (0)	-	-	-	18 (100)	0	18	0	-	-	-	0
Cederbrant K, 1998	PPL; MDM; PEN-G; PEN-V; AMP; CLO; BAC	9	1 (11)	0	1	0	8 (89)	2	1	5	50.0	100	100	83.3
Rebello Gomes E, 2007	Unspecified	14	0 (0)	-	-	-	14 (100)	0	1	13	-	-	-	92.9
Rodriguez-Alvarez M, 2008	PPL; MDM; Unspecified penicillins; AMX; AMP; Unspecified cephalosporins	23	0 (0)	-	-	-	23 (100)	2	4	17	-	-	-	81.0
Tantikul C, 2008	AMX; LEX	2	1 (50)	0	1	0	1 (50)	0	0	1	100	100	100	100
Ponvert C, 2011	Suspected drug	1269	0 (0)	-	-	-	1269 (100)	73	104	1092	-	-	-	91.3
Hjortlund J, 2013	PPL; MDM; PEN-V; AMX	235	1 (0.4)	1	0	0	234 (99.6)	32	36	166	-	-	-	82.2
Macy E, 2013	PPL; MDM; AMX	296	0 (0)	-	-	-	296 (100)	3	3	290	-	-	-	99.0
Zambonino MA, 2014	PPL; MDM; AMX; AMX-CLV; CXM	717	0 (0)	-	-	-	717 (100)	2	51	664	-	-	-	92.9
Mori F, 2015	AMX	152	0 (0)	-	-	-	152 (100)	11	14	127	-	-	-	90.1
Manuyakorn W, 2016	PEN-G; AMX-CLV; AMP; Suspected drug	86	0 (0)	-	-	-	86 (100)	3	4	79	-	-	-	95.2
Faitelson Y, 2017	PEN-G; AMX	101	0 (0)	-	-	-	101 (100)	0	10	91	-	-	-	90.1

C. Intradermal tests

Reference	Assessed drugs/determinants	N participants performing intradermal tests	Participants with positive intradermal tests				Participants with negative intradermal tests				Sn (%)	Sp (%)	PPV (%)	NPV (%)
			Total N (%)	No DPT performed	Positive DPT (TP)	Negative DPT (FP)	Total N (%)	No DPT performed	Positive DPT (FN)	Negative DPT (TN)				
Brown BC, 1964	PPL	1085	416 (38.3)	383	9	24	669 (61.7)	669	0	0	-	-	27.3	-
Resnik SS, 1965	PPL	16	6 (38)	0	0	6	10 (63)	0	0	10	-	62.5	0	100

Wide L, 1971	PPL; PEN-G	23	9 (39)	9	0	0	14 (61)	0	0	14	-	-	-	100
Basomba A, 1979	PPL; PEN-G; AMP	80	55 (69)	55	0	0	25 (31)	19	2	4	-	-	-	66.7
Sullivan TJ, 1981	PPL; MDM; PEN-G; AMP; CAR; CEF	256	110 (43.0)	110	0	0	146 (57.0)	63	2	81	-	-	-	97.6
Blanca M, 1990*	PPL; MDM; AMX; AMP	278	36 (12.5)	34	2	0	242 (84.0)	1	19	222	9.5	100	100	92.1
Romano A, 1995*	PPL; MDM; PEN-G; AMP	60	33 (55)	15	18	0	27 (45)	11	0	16	100	100	100	100
Macy E, 1997	PPL; MDM; PEN-G; AMX	342	54 (15.8)	54	0	0	288 (84.2)	73	13	202	-	-	-	94.0
Cederbrant K, 1998*	PPL; MDM; PEN-G; PEN- V; AMP; CLO; BAC	13	4 (31)	3	1	0	9 (69)	2	1	6	50.0	100	100	85.7
Hervé M, 1998	PPL; MDM; AMX; AMX- CLV	105	36 (34.3)	36	0	0	69 (65.7)	17	6	46	-	-	-	88.5
López Tiro JJ, 2001	PPL; MDM; PEN-G	20	19 (95)	19	0	0	1 (5)	0	1	0	-	-	-	0
Novalbos A, 2001	PPL; MDM; PEN-G; AMX ^c	950	30 (3.2)	30	0	0	920 (96.8)	13	3	904	-	-	-	99.7
Romano A, 2002*	PPL; MDM; PEN-G; AMX; AMP	259	105 (40.5)	100	2	3	154 (59.5)	29	3	122	40.0	97.6	40.0	97.6
Borch JE, 2006	PPL; MDM; PEN-G; AMP	13	1 (8)	1	0	0	12 (92)	0	0	12	-	-	-	100
Bousquet PJ, 2007	PPL; MDM; PEN- G; AMX; AMP; Suspected drug	952	120 (12.6)	120	0	0	832 (87.4)	0	69	763	-	-	-	91.7
Rebello Gomes E, 2007	Unspecified	16	0 (0)	0	0	0	16 (100)	0	1	15	-	-	-	93.8
Goldberg A, 2008*	PPL; MDM; PEN-G; AMX; AMP	169	73 (43.2)	6	7	60	96 (56.8)	2	4	90	63.6	60.0	10.4	95.7
Nolan RC, 2008	PPL; MDM	29	4 (14)	4	0	0	25 (86)	0	4	21	-	-	-	84.0
Padial A, 2008	MDM; PEN-G; AMX; AMP; CLO; Unspecified cephalosporins	18	2 (11)	0	2	0	16 (89)	0	16	0	11.1	-	100	0
Rodriguez-Alvarez M, 2008	PPL; MDM; Unspecified penicillins; AMX; AMP; Unspecified cephalosporins	23	1 (4)	1	0	0	22 (96)	1	4	17	-	-	-	81.0
Silva R, 2009	PPL; MDM; AMX; AMP; Suspected drug	64	3 (5)	3	0	0	61 (95)	28	2	31	-	-	-	93.9
Chambel M, 2010	PPL; MDM; Unspecified penicillins; AMX; AMX- CLV; CXM	47	8 (17)	8	0	0	39 (83)	9	1	29	-	-	-	96.7
Kamboj S, 2011	PEN-G; AMP; Unspecified cephalosporins	60	4 (7)	4	0	0	56 (93)	0	4	52	-	-	-	92.9
Moral L, 2011	PPL; MDM; Unspecified penicillins; Suspected drug	16	0 (0)	0	0	0	16 (100)	0	0	16	-	-	-	100

Na HR, 2011	Suspected drug	5	2 (40)	0	2	0	3 (60)	0	3	0	40.0	-	100	0
Ponvert C, 2011	Suspected drug	1431	97 (6.8)	97	0	0	1334 (93.2)	19	111	1204	-	-	-	91.6
Richter AG, 2011*	PPL; MDM; AMX ^d	47	8 (17)	6	1	1	39 (83)	6	5	28	16.7	96.6	50.0	84.8
Hjortlund J, 2012	PEN-G; AMP	360	20 (5.6)	20	0	0	340 (94.4)	0	46	294	-	-	-	86.5
Kopac P, 2012	PPL; MDM; Suspected drug	268	8 (3.0)	8	0	0	260 (97.0)	0	19	241	-	-	-	92.7
Erkoçoglu M, 2013	PPL; MDM; PEN-G; AMX; AMP; CEC; Unspecified cephalosporins	70	1 (1)	1	0	0	69 (99)	7	0	62	-	-	-	100
Hjortlund J, 2013	PPL; MDM; PEN-G; AMP; DCX; MEC	342	35 (10.2)	35	0	0	307 (89.8)	16	50	241	-	-	-	82.8
Macy E, 2013	PPL; MDM; AMX	500	4 (0.8)	4	0	0	496 (99.2)	0	4	492	-	-	-	99.2
Rimawi RH, 2013	PPL; PEN-G	146	1 (0.7)	1	0	0	145 (99.3)	0	0	145	-	-	-	100
Al-Ahmad M, 2014	PPL; MDM; AMX; AMP	109	29 (26.6)	29	0	0	80 (73.4)	1	1	78	-	-	-	98.7
Caubet JC, 2014*	PPL; MDM; AMX; CXM ^f	96	14 (15)	0	7	7	82 (85)	0	7	75	50.0	91.5	50.0	91.5
Colas H, 2014	Unspecified penicillins; Unspecified cephalosporins	90	2 (2)	2	0	0	88 (98)	0	2	86	-	-	-	97.7
Picard M, 2014	PEN-G	562	184 (32.7)	184	0	0	378 (67.3)	3	18	357	-	-	-	95.2
Trautmann A, 2014	PEN-G; AMP; CRO	37	10 (27)	10	0	0	27 (73)	0	1	26	-	-	-	96.3
Zambonino MA, 2014	PPL; MDM; AMX; AMX-CLV; CXM	780	3 (0.4)	3	0	0	777 (99.6)	0	56	721	-	-	-	92.8
Blanca-Lopez N, 2015	PPL; MDM; AMX; CLV	252	20 (7.9)	20	0	0	232 (92.1)	0	11	221	-	-	-	95.3
Mori F, 2015*		189	8 (4.2)	4	3	1	181 (95.8)	8	14	159	17.6	99.4	75.0	91.9
Abrams EM, 2016	PPL; PEN-G; AMP	106	1 (0.9)	1	0	0	105 (99.1)	0	5	100	-	-	-	95.2
Atanaskovic-Markovic M, 2016	PPL; MDM; PEN-G; AMX; AMP; Suspected drug ^h	1026	61 (5.9)	61	0	0	965 (94.1)	0	19	946	-	-	-	98.0
Chen JR, 2016	PPL; MDM; PEN-G	224	0 (0)	0	0	0	224 (100)	0	1	223	-	-	-	99.6
King EA, 2016	PPL; PEN-G	50	0 (0)	0	0	0	50 (100)	0	1	49	-	-	-	98.0
Manuyakorn W, 2016	PEN-G; AMX-CLV; AMP; Suspected drug	122	8 (6.6)	8	0	0	114 (93.4)	0	10	104	-	-	-	91.2
Meng J, 2016	PPL; MDM; PEN-G;	80	18 (23)	18	0	0	62 (78)	0	2	60	-	-	-	96.8

	AMX ¹													
Mill C, 2016	PPL; PEN-G	17	1 (6)	0	1	0	16 (94)	0	16	0	5.9	-	100	0
Mota I, 2016	PPL; MDM; PEN-G; AMX; CLV; CXM; Suspected drug	214	25 (11.7)	25	0	0	189 (88.3)	0	5	184	-	-	-	97.4
Confino-Cohen R, 2017*	PPL; MDM; PEN-G; AMX; Suspected drug	642	34 (5.3) ^a	4	2	28	400 (62.3) ^a	10	40	350	4.8	92.6	6.7	89.7
Fazlollahi MR, 2017	PPL; MDM; AMX	44	8 (18)	8	0	0	36 (82)	22	1	13	-	-	-	92.9
Marwood J, 2017	PPL; MDM; AMX	101	14 (13.9)	14	0	0	87 (86.1)	3	3	81	-	-	-	96.4
McDanel DL, 2017	PPL; PEN-G	140	0 (0)	0	0	0	139 (99.3)	120	0	19	-	-	-	100
Misirlioglu ED, 2017	PPL; MDM; PEN-G; AMX-CLV; Suspected drug	342	24 (7.0)	24	0	0	318 (93.0)	0	17	301	-	-	-	94.7
Soria A, 2017	Unspecified penicillins; AMX; AMX-CLV; OXA; CLO	259	23 (8.9)	23	0	0	236 (91.1)	11	7	218	-	-	-	96.9
Sundquist BK, 2017	PPL; MDM	37	0 (0)	0	0	0	37 (100)	1	0	36	-	-	-	100
Tannert LK, 2017	Suspected drug	21	21 (100)	9	7	5	0 (0)	0	0	0	-	-	58.3	-
Vyles D, 2017	PPL; MDM	97	0 (0)	0	0	0	97 (100)	0	0	97	-	-	-	100
Ibáñez MD, 2018*	PPL; MDM; PEN-G; PEN- G; AMX; AMX-CLV	726	15 (2.1)	0	3	12	711 (97.9)	0	30	681	9.1	98.3	20.0	95.8

C.1. Intradermal tests (immediate reactions)

Reference	Assessed drugs/determinants	N participants performing intradermal tests	Participants with positive intradermal tests				Participants with negative intradermal tests				Sn (%)	Sp (%)	PPV (%)	NPV (%)
			Total N (%)	No DPT performed	Positive DPT (TP)	Negative DPT (FP)	Total N (%)	No DPT performed	Positive DPT (FN)	Negative DPT (TN)				
Wide L, 1971	PPL; PEN-G	23	9 (39)	9	0	0	14 (61)	0	0	14	-	-	-	100
Cederbrant K, 1998	PPL; MDM; PEN-G; PEN- V; AMP; CLO; BAC	5	2 (40)	2	0	0	3 (60)	1	1	1	-	-	-	50.0
Lebel B, 2001	PPL; MDM; PEN-G; AMX; AMP	17	12 (71)	12	0	0	5 (29)	0	5	0	-	-	-	0
Rebello Gomes E, 2007	Unspecified	2	0 (0)	0	0	0	2 (100)	0	0	2	-	-	-	100
Na HR, 2011	Suspected drug	5	2 (40)	0	2	0	3 (60)	0	3	0	40.0	-	100	0
Ponvert C, 2011	Suspected drug	162	41 (25.3)	41	0	0	121 (74.7)	2	7	112	-	-	-	94.1
Erkoçoğlu M, 2013	PPL; MDM; PEN-G; AMX; AMP; CEC; Unspecified cephalosporins	70	1 (1)	1	0	0	69 (99)	7	0	62	-	-	-	100

Hjortlund J, 2013	PPL; MDM; PEN-G; AMP; DCX; MEC	36	5 (14)	5	0	0	31 (86)	1	11	19	-	-	-	63.3
Macy E, 2013	PPL; MDM; AMX	52	1 (2)	1	0	0	51 (98)	0	1	50	-	-	-	98.0
Zambonino MA, 2014	PPL; MDM; AMX; AMX-CLV; CXM	63	1 (2)	1	0	0	62 (98)	0	5	57	-	-	-	91.9
Blanca-Lopez N, 2015	PPL; MDM; AMX; CLV	252	20 (7.9)	20	0	0	232 (92.1)	0	11	221	-	-	-	95.3
Mori F, 2015		37	2 (5)	1	1	0	35 (95)	0	2	33	33.3	100	100	94.3
Manuyakorn W, 2016	PEN-G; AMX-CLV; AMP; Suspected drug	36	5 (14)	5	0	0	31 (86)	0	6	25	-	-	-	80.6
Mill C, 2016	PPL; PEN-G	17	1 (6)	0	1	0	16 (94)	0	16	0	5.9	-	100	0
Misirlioglu ED, 2017	PPL; MDM; PEN-G; AMX-CLV; Suspected drug	342	24 (7.0)	24	0	0	318 (93.0)	0	17	301	-	-	-	94.7
Tannert LK, 2017	Suspected drug	17	17 (100)	7	6	4	0 (0)	0	0	0	-	-	60.0	-

C.2. Intradermal tests (non-immediate reactions)

Reference	Assessed drugs/determinants	N participants performing intradermal tests	Participants with positive intradermal tests				Participants with negative intradermal tests				Sn (%)	Sp (%)	PPV (%)	NPV (%)
			Total N (%)	No DPT performed	Positive DPT (TP)	Negative DPT (FP)	Total N (%)	No DPT performed	Positive DPT (FN)	Negative DPT (TN)				
Romano A, 1995*	PPL; MDM; PEN-G; AMP	60	33 (55)	15	18	0	27 (45)	11	0	16	100	100	100	100
Cederbrant K, 1998*	PPL; MDM; PEN-G; PEN-V; AMP; CLO; BAC	8	2 (25)	1	1	0	6 (75)	1	0	5	100	100	100	100
Romano A, 2002*	PPL; MDM; PEN-G; AMX; AMP	259	105 (40.5)	100	2	3	154 (59.5)	29	3	122	40	97.6	40.0	97.6
Rebello Gomes E, 2007	Unspecified	14	0 (0)	0	0	0	14 (100)	0	1	13	-	-	-	92.9
Padial A, 2008	MDM; PEN-G; AMX; AMP; CLO; Unspecified cephalosporins	18	2 (11)	0	2	0	16 (89)	0	16	0	11.1	-	100	0
Rodriguez-Alvarez M, 2008	PPL; MDM; Unspecified penicillins; AMX; AMP; Unspecified cephalosporins	23	1 (4)	1	0	0	22 (96)	1	4	17	-	-	-	81.0
Ponvert C, 2011	Suspected drug	1269	56 (4.4)	56	0	0	1213 (95.6)	17	104	1092	-	-	-	91.3
Hjortlund J, 2013	PPL; MDM; PEN-G; AMP; DCX; MEC	235	24 (10.2)	24	0	0	211 (89.8)	9	36	166	-	-	-	82.2
Macy E, 2013	PPL; MDM; AMX	296	3 (1.0)	3	0	0	293 (99.0)	0	3	290	-	-	-	99.0
Caubet JC, 2014*	PPL; MDM; AMX; CXM ^f	96	14 (15)	0	7	7	82 (85)	0	7	75	50.0	91.5	50.0	91.5
Trautmann A, 2014	PEN-G; AMP; CRO	37	10 (27)	10	0	0	27 (73)	0	1	26	-	-	-	96.3

Zambonino MA, 2014	PPL; MDM; AMX; AMX-CLV; CXM	717	2 (0.3)	2	0	0	715 (99.7)	0	51	664	-	-	-	92.9
Mori F, 2015*		152	6 (3.9)	3	2	1	146 (96.1)	8	12	126	14.3	99.2	66.7	91.3
Atanaskovic-Markovic M, 2016	PPL; MDM; PEN-G; AMX; AMP; Suspected drug ^h	1026	61 (5.9)	61	0	0	965 (94.1)	0	19	946	-	-	-	98.0
Manuyakorn W, 2016	PEN-G; AMX-CLV; AMP; Suspected drug	86	3 (4)	3	0	0	83 (97)	0	4	79	-	-	-	95.2
Confino-Cohen R, 2017*	PPL; MDM; PEN-G; AMX; Suspected drug	642	34 (5.3) ^c	4	2	28	400 (62.3) ^c	10	40	350	4.8	92.6	6.7	89.7

D. Specific IgE quantification

Reference	Assessed drugs/determinants	N participants performing IgE	Participants with positive IgE				Participants with negative IgE				Sn (%)	Sp (%)	PPV (%)	NPV (%)
			Total N (%)	No DPT performed	Positive DPT (TP)	Negative DPT (FP)	Total N (%)	No DPT performed	Positive DPT (FN)	Negative DPT (TN)				
Wide L, 1971	PEN-V	23	9 (39)	9	0	0	14 (61)	0	0	14	-	-	-	100
Basomba A, 1979	PEN-G; PEN-V	81	37 (46)	37	0	0	44 (54)	33	2	9	-	-	-	81.8
Blanca M, 1990*	PEN-G; AMX; AMP	288	28 (9.7)	26	2	0	260 (90.3)	19	19	222	9.5	100	100	92.1
Surtees SJ, 1991	PEN-G; PEN-V	132	4 (3.0)	4	0	0	128 (97.0)	0	0	128	-	-	-	100
Romano A, 1995	PEN-G; PEN-V	18	18 (100)	0	18	0	0 (0)	0	0	0	-	-	100	-
Cederbrant K, 1998*	PEN-G; PEN-V	17	7 (41)	5	2	0	10 (59)	0	4	6	33.3	100	100	60.0
Hervé M, 1998	Unspecified penicillins; AMX	109	16 (14.7)	16	0	0	93 (85.3)	41	6	46	-	-	-	88.5
Loza Cortina C, 1998	Unspecified	28	0 (0)	-	-	-	28 (100)	2	4	22	-	-	-	84.6
Blanca M, 2001*	Suspected drug	81	12 (15)	0	12	0	69 (85)	0	39	30	23.5	100	100	43.5
Brusca I, 2007*	PEN-G; PEN-V; AMX; AMP	122	24 (19.7)	0	24	0	98 (80.3)	0	68	30	26.1	100	100	30.6
Fontaine C, 2007	PEN-G; AMX; AMP; CEC	11	0 (0)	-	-	-	11 (100)	0	0	11	-	-	-	100
Silva R, 2009	PEN-G; PEN-V; AMX; AMP	67	11 (16)	11	0	0	56 (84)	23	2	31	-	-	-	93.9
Caubet JC, 2010*	PEN-G; PEN-V; AMX	88	2 (2)	0	0	2	86 (98)	0	6	80	0	97.6	0	93.0
Chambel M, 2010	PEN-G; PEN-V; AMX	106	2 (1.9)	2	0	0	104 (98.1)	74	1	29	-	-	-	96.7
Macy E, 2010*	PEN-G; PEN-V; AMX	150	4 (2.7)	0	0	4	146 (97.3)	6	3	137	0	97.2	0	97.9
Holm A, 2011	PEN-G; PEN-V; AMX; AMP	580	0 (0)	-	-	-	580 (100)	0	14	566	-	-	-	97.6
Moral L, 2011	PEN-G; AMX	16	0 (0)	-	-	-	16 (100)	0	0	16	-	-	-	100

Richter AG, 2011*	PEN-G; PEN-V; AMX	47	1 (2)	0	0	1	46 (98)	6	6	34	0	97.1	0	85.0
Seitz CS, 2011	PEN-G; PEN-V; AMX; AMP	8	0 (0)	-	-	-	8 (100)	0	0	8	-	-	-	100
Hjortlund J, 2012	PEN-G; PEN-V; AMX; AMP	405	44 (10.9)	44	0	0	361 (89.1)	21	46	294	-	-	-	86.5
Kopac P, 2012*	PEN-G; PEN-V; AMX; AMP	606	49 (8.1)	45	0	4	557 (91.9)	131	21	405	0	99.0	0	95.1
Soto L, 2012	PEN-G; PEN-V; AMX	226	5 (2.2)	5	0	0	221 (97.8)	80	4	137	-	-	-	97.2
Erkoçoglu M, 2013	PEN-G; PEN-V; AMX; AMP; CEC	70	0 (0)	-	-	-	70 (100)	8	0	62	-	-	-	100
Hjortlund J, 2013	PEN-G; PEN-V; AMX; AMP	342	9 (2.6)	9	0	0	333 (97.4)	42	50	241	-	-	-	82.8
Al-Ahmad M, 2014	Unspecified penicillins	124	1 (0.8)	1	0	0	123 (99.2)	44	1	78	-	-	-	98.7
Trautmann A, 2014	PPL; PEN-G; PEN-V; AMX; AMP	37	0 (0)	-	-	-	37 (100)	10	1	26	-	-	-	96.3
Zambonino MA, 2014	PEN-G; AMX	66	2 (3)	2	0	0	64 (97)	2	5	57	-	-	-	91.9
Ferré-Ybarz L, 2015	PEN-G; PEN-V; AMX; CLV; AMP	82	2 (2)	2	0	0	80 (98)	3	1	76	-	-	-	98.7
Mori F, 2015*	PEN-G; PEN-V; AMX; AMP	38	3 (8)	2	1	0	35 (92)	0	2	33	33.3	100	100	94.3
Meng J, 2016	PEN-G; PEN-V; AMX	16	3 (19)	3	0	0	13 (81)	6	2	5	-	-	-	71.4
Tannert LK, 2017*	PEN-G; PEN-V; AMX; AMP	21	5 (24)	2	2	1	16 (76)	7	5	4	28.6	80.0	66.7	44.4
Ibáñez MD, 2018*	PEN-V; AMX	732	8 (1.1)	0	1	7	724 (98.9)	0	34	690	2.9	99.0	12.5	95.3
Marraccini P, 2018	PEN-G; PEN-V; AMX; AMP; CEC	37	9 (24)	9	0	0	28 (76)	3	1	24	-	-	-	96.0
Sedlackova L, 2018	PEN-G; PEN-V; AMX; AMP	24	1 (4)	1	0	0	23 (96)	13	1	9	-	-	-	90.0

D.1. Specific IgE quantification (immediate reactions)

Reference	Assessed drugs/determinants	N participants performing IgE	Participants with positive IgE				Participants with negative IgE				Sn (%)	Sp (%)	PPV (%)	NPV (%)
			Total N (%)	No DPT performed	Positive DPT (TP)	Negative DPT (FP)	Total N (%)	No DPT performed	Positive DPT (FN)	Negative DPT (TN)				
Wide L, 1971	PEN-V	23	9 (39)	9	0	0	14 (61)	0	0	14	-	-	-	100
Surtees SJ, 1991	PEN-G; PEN-V	132	4 (3.0)	4	0	0	128 (97.0)	0	0	128	-	-	-	100
Cederbrant K, 1998	PEN-G; PEN-V	8	4 (50)	3	1	0	4 (50)	0	3	1	25.0	100	100	25.0
Blanca M, 2001	Suspected drug	26	11 (42)	0	11	0	15 (58)	0	15	0	42.3	-	100	0
Brusca I, 2007	PEN-G; PEN-V; AMX;	122	24 (19.7)	0	24	0	98 (80.3)	0	68	30	26.1	100	100	30.6

	AMP													
Seitz CS, 2011	PEN-G; PEN-V; AMX; AMP	8	0 (0)	-	-	-	8 (100)	0	0	8	-	-	-	100
Erkoçoglu M, 2013	PEN-G; PEN-V; AMX; AMP; CEC	70	0 (0)	-	-	-	70 (100)	8	0	62	-	-	-	100
Hjortlund J, 2013	PEN-G; PEN-V; AMX; AMP	36	4 (11)	4	0	0	32 (89)	2	11	19	-	-	-	63.3
Zambonino MA, 2014	PEN-G; AMX	66	2 (3)	2	0	0	64 (97)	2	5	57	-	-	-	91.9
Mori F, 2015	PEN-G; PEN-V; AMX; AMP	38	3 (8)	2	1	0	35 (92)	0	2	33	33.3	100	100	94.3
Tannert LK, 2017	PEN-G; PEN-V; AMX; AMP	17	4 (24)	2	1	1	13 (77)	5	5	3	16.7	75.0	50.0	37.5

D.2. Specific IgE quantification (non-immediate reactions)

Reference	Assessed drugs/determinants	N participants performing IgE	Participants with positive IgE				Participants with negative IgE				Sn (%)	Sp (%)	PPV (%)	NPV (%)
			Total N (%)	No DPT performed	Positive DPT (TP)	Negative DPT (FP)	Total N (%)	No DPT performed	Positive DPT (FN)	Negative DPT (TN)				
Romano A, 1995	PEN-G; PEN-V	18	18 (100)	0	18	0	0 (0)	-	-	-	-	-	100	-
Cederbrant K, 1998	PEN-G; PEN-V	9	3 (33)	2	1	0	6 (67)	0	1	5	50.0	100	100	83.3
Blanca M, 2001	Suspected drug	25	1 (4)	0	1	0	24 (96)	0	24	0	4.0	-	100	0
Caubet JC, 2010	PEN-G; PEN-V; AMX	88	2 (2)	0	0	2	86 (98)	0	6	80	0	97.6	0	93.0
Hjortlund J, 2013	PEN-G; PEN-V; AMX; AMP	235	12 (5.1)	12	0	0	223 (94.9)	21	36	166	-	-	-	82.2
Trautmann A, 2014	PPL; PEN-G; PEN-V; AMX; AMP	37	0 (0)	-	-	-	37 (100)	10	1	26	-	-	-	96.3
Tannert LK, 2017	PEN-G; PEN-V; AMX; AMP	2	0 (0)	-	-	-	2 (100)	2	0	0	-	-	-	-
Sedlackova L, 2018	PEN-G; PEN-V; AMX; AMP	24	1 (4)	1	0	0	23 (96)	13	1	9	-	-	-	90.0

DPT: Drug provocation test; FN: False negative; FP: False positive; NPV: Negative predictive value; PPV: Positive predictive value; Sn: Sensitivity; Sp: Specificity

Determinants: AMP: Ampicillin; AMX: Amoxicillin; BAC: Bacampicillin; CAR: Carbenicillin; CEC: Cefaclor; CEF: Cephalothin; CFZ: Cefazolin; CLO: Cloxacillin; CLV: Clavulanate; CRO: Ceftriaxone; CXM: Cefuroxime; DCX : Dicloxacillin ; LEX: Cephalexin; MDM: Penicillin minor determinant; MEC : Mecillinam ; OXA: Oxacillin; PEN-G: Penicillin G; PEN-V: Penicillin V; PPL: Penicillin major determinant

* Assessed in the bivariate random-effects meta-analysis

^a The percentages of participants with positive and negative results do not add up to 100%, as some participants had equivocal/undetermined test results; ^b In 130 participants, ampicillin, cefazolin, cefuroxime and ceftriaxone were also tested; ^c In some patients, cefazolin, cefuroxime and ceftriaxone were also tested; ^d Flucloxacillin and benzylpenicillin were tested in some patients; ^e 27 participants had negative results in the 1st skin tests but positive results when retested; ^f Cephalosporins were tested if suspected; ^g Penicillin G, amoxicillin-clavulanate and ampicillin were tested in some patients; ^h After the first 100 participants, tests were only performed with the suspected drug; ⁱ Flucloxacillin and clavulanate if suspected

Table E4. Summary of meta-analysis and meta-regression results for individual diagnostic measures of penicillin allergy tests

A. Skin tests				
	N studies	Pooled meta-analytic value (95%CI) [I^2 ; Q Cochran p value]	Multivariable meta-regression model	
			Variables in the model (OR; 95%CI) [p value]	Residual heterogeneity
% positive results	95	10.4% (8.3-13.2%) [98.7%; <0.001]	- Year of publication (1.0; 0.9-1.0) [<0.001] - Europe (region) (0.6; 0.4-1.2) [0.15] - North America (region) (0.3; 0.2-0.6) [<0.001] - Reported Allergy to Penicillins (1.4; 0.9-2.1) [0.12]	$I^2=97.9\%$; <0.001
Sensitivity	23	27.9% (18.2-42.7%) [89.6%; <0.001]	- Year of publication (0.9; 0.9-1.0) [<0.001] - Europe (region) (0.4; 0.2-1.0) [0.04] - North America (region) (0.2; 0.1-0.7) [0.01] - Reported Allergy to Penicillins (0.5; 0.2-0.9) [0.03]	$I^2=54.6\%$; 0.02
Specificity	22	95.8% (93.2-98.6%) [99.0%; <0.001]	- Middle East (region) (0.9; 0.9-0.9) [<0.001] - Cephalosporins testing (0.9; 0.9-1.0) [0.02]	$I^2=68.4\%$; <0.001
PPV	27	50.2% (38.4-65.7%) [80.8%; <0.001]	- Year of publication (1.0; 1.0-1.0) [0.11] - North America (region) (0.6; 0.4-0.9) [0.02] - Middle East (region) (0.2; 0.1-0.3) [<0.001] - Reported Allergy to Penicillins (0.8; 0.5-1.1) [0.09]	$I^2=25.6\%$; 0.18
NPV	93	95.8% (95.0-96.5%) [91.3%; <0.001]	- North America (region) (1.0; 1.0-1.1) [<0.001] - Number of tested drugs/determinants (1.0; 1.0-1.0) [0.05] - Suspected Drug Testing (1.0; 1.0-1.0) [0.15]	$I^2=88.2\%$; <0.001
A.1. Skin tests (immediate reactions)				
	N studies	Pooled meta-analytic value (95%CI) [I^2 ; Q Cochran p value]	Multivariable meta-regression model	
			Variables in the model (OR; 95%CI) [p value]	Residual heterogeneity
% positive results	25	16.6% (11.4-24.2%) [98.0%; <0.001]	- Europe (region) (2.1; 1.0-4.2) [0.04] - North America (region) (0.5; 0.1-1.6) [0.22] - Convenience Sample (2.9; 1.5-5.6) [0.001] - Suspected Drug Testing (2.2; 1.1-4.1) [0.02]	$I^2=94.4\%$; <0.001
Sensitivity	5	25.6% (13.5-48.9%) [0%; 0.44]	- ^a	- ^a
Specificity	3	99.9% (99.6-100%) [0%; 0.51]	- ^a	- ^a
PPV	6	73.8% (57.0-95.6%) [0%; 0.87]	- ^a	- ^a
NPV	24	90.4% (85.7-95.4%) [98.4%; <0.001]	- ^b	- ^b
A.2. Skin tests (non-immediate reactions)				
	N studies	Pooled meta-analytic value (95%CI) [I^2 ; Q Cochran p value]	Multivariable meta-regression model	
			Variables in the model (OR; 95%CI) [p value]	Residual heterogeneity
% positive results	21	6.7% (3.4-12.9%) [97.8%; <0.001]	- Year of Publication (0.9; 0.8-1.0) [0.008]	$I^2=96.5\%$; <0.001

			- North America (region) (0.2; 0.001-2.4) [0.19]	
Sensitivity	8	33.0% (15.4-70.9%) [90.7%; <0.001]	- ^c	- ^c
Specificity	7	95.9% (92.7-99.2%) [79.3%; <0.001]	Not performed. No substantial heterogeneity was detected when performing subgroup analysis (i) of the 3 studies assessing only participants reporting allergy to aminopenicillins (99.7%; 95%CI=99.0-100%; $I^2=0\%$; $p=0.64$), or (ii) of the 2 studies in which cephalosporin allergy was tested (91.3%; 95%CI=85.5-97.6%; $I^2=0\%$; $p=0.63$)	
PPV	8	61.6% (42.1-90.3%) [71.5%; 0.001]	Not performed. No substantial heterogeneity was detected when performing subgroup analysis (i) of the 2 studies assessing only adult patients (80.9%; 95%CI=52.7-100%; $I^2=0\%$; $p=0.83$), (ii) of the 3 studies assessing only children (58.7%; 95%CI=39.9-86.2%; $I^2=0\%$; $p=0.66$), (iii) of the 4 studies assessing a convenience sample (66.7%; 95%CI=49.1-90.6%; $I^2=0\%$; $p=0.58$), (iv) of the 3 studies assessing only participants reporting allergy to aminopenicillins (96.8%; 95%CI=90.0-100%; $I^2=0\%$; $p=0.54$), or (v) of the 3 studies in which cephalosporin allergy was tested (66.8%; 95%CI=46.4-96.2%; $I^2=15.3\%$; $p=0.37$)	
NPV	21	93.4% (91.4-95.5%) [87.8%; <0.001]	- Middle East (region) (1.0; 0.9-1.0) [0.04] - Consecutive testing (1.0; 1.0-1.1) [0.07] - Aminopenicillins testing (1.1; 1.0-1.1) [0.001]	$I^2=62.0\%$; 0.001

B. Skin prick tests

	N studies	Pooled meta-analytic value (95%CI) [I^2 ; Q Cochran p value]	Multivariable meta-regression model	
			Variables in the model (OR; 95%CI) [p value]	Residual heterogeneity
% positive results	47	2.0% (1.3-3.2%) [89.5%; <0.001]	- Year of publication (0.9; 0.9-1.0) [<0.001] - Europe (region) (0.3; 0.1-0.7) [0.005] - North America (region) (0.3; 0.1-1.0) [0.06] - Consecutive sampling (0.4; 0.2-1.1) [0.08]	$I^2=78.1\%$; <0.001
Sensitivity	4	14.6% (3.2-65.6%) [73.8%; 0.003]	Not performed. No substantial heterogeneity was detected when performing subgroup analysis of the 3 studies assessing outpatients (7.0%; 95%CI=2.4-20.6%; $I^2=0\%$; $p=0.44$)	
Specificity	4	98.8% (96.1-100%) [40.7%; 0.28]	- ^a	- ^a
PPV	5	71.9% (48.4-100%) [0%; 0.77]	- ^a	- ^a
NPV	47	94.8% (93.6-96.1%) [93.1%; <0.001]	- Year of publication (1.0; 1.0-1.0) [0.04] - North America (region) (1.0; 1.0-1.1) [0.001] - Cephalosporins testing (1.0; 1.0-1.1) [0.16]	$I^2=87.0\%$; <0.001

B.1. Skin prick tests (immediate reactions)

	N studies	Pooled meta-analytic value (95%CI) [I^2 ; Q Cochran p value]	Multivariable meta-regression model	
			Variables in the model (OR; 95%CI) [p value]	Residual heterogeneity
% positive results	14	3.3% (1.9-5.5%) [34.3%; 0.17]	- Australia and Far East (region) (6.1; 2.2-16.8) [0.001] - Reported Allergy to Aminopenicillins (2.5; 1.1-5.4) [0.03]	$I^2=0\%$; 0.96
Sensitivity	1	5.9% (0.9-39.4%)	- ^d	- ^d
Specificity	1	92.9% (75.6-100%)	- ^d	- ^d
PPV	1	75.0% (33.7-100%)	- ^d	- ^d
NPV	14	95.1% (92.9-97.3%) [45.8%; 0.001]	- Adults (age group) (0.7; 0.5-0.9) [0.002] - Consecutive sampling (1.0; 1.0-1.0) [0.03]	$I^2=0.1\%$; 0.13

- Number of tested drugs/determinants (1.0; 1.0-1.0) [0.04]

B.2. Skin prick tests (non-immediate reactions)

	N studies	Pooled meta-analytic value (95%CI) [I^2 ; Q Cochran p value]	Multivariable meta-regression model	
			Variables in the model (OR; 95%CI) [p value]	Residual heterogeneity
% positive results	12	1.0% (0.3-3.7%) [71.0%; <0.001]	- Year of publication (0.8; 0.7-0.9) [<0.001] - Europe (region) (0.04; 0.01-0.2) [<0.001] - Consecutive testing (0.1; 0.01-0.4) [0.003]	$I^2=0\%$; 0.48
Sensitivity	2	67.8% (33.9-100%) [0%; 0.62]	- ^a	- ^a
Specificity	2	90.1% (71.6-100%) [0%; 0.64]	- ^a	- ^a
PPV	2	75.0% (42.6-100%) [0%; >0.99]	- ^a	- ^a
NPV	12	91.4% (88.0-95.0%) [88.9%; <0.001]	- Adults (age group) (0.9; 0.8-1.0) [0.001] - North America (region) (1.1; 1.1-1.1) [<0.001]	$I^2=4.9\%$; 0.18

C. Intradermal tests

	N studies	Pooled meta-analytic value (95%CI) [I^2 ; Q Cochran p value]	Multivariable meta-regression model	
			Variables in the model (OR; 95%CI) [p value]	Residual heterogeneity
% positive results	59	10.1% (7.2-14.1%) [98.9%; <0.001]	- Year of publication (0.; 0.9-1.0) [<0.001] - Europe (region) (0.5; 0.3-1.1) [0.08] - North America (region) (0.2; 0.1-0.5) [0.001] - Convenience Sample (1.6; 0.9-2.9) [0.12]	$I^2=97.4\%$; <0.001
Sensitivity	13	25.2% (14.4-43.8%) [86.9%; <0.001]	- ^b	- ^b
Specificity	11	91.1% (83.5-99.4%) [99.5%; <0.001]	- Europe (region) (1.3; 1.1-1.5) [0.008] - North America (region) (0.8; 0.5-1.3) [0.35] - Middle East (region) (0.3; 0.2-0.7) [0.005]	$I^2=99.1\%$; 0.006
PPV	16	47.2% (32.5-68.5%) [86.9%; <0.001]	- Convenience Sample (2.6; 1.6-4.3) [<0.001] - Cephalosporins Testing (0.3 0.1-0.5) [<0.001] - Reported Allergy to Penicillins (0.3; 0.2-0.4) [<0.001]	$I^2=0\%$; 0.77
NPV	57	95.2% (94.1-96.3%) [91.3%; <0.001]	- North America (region) (1.0; 1.0-1.1) [0.03] - Suspected Drug Testing (1.0; 1.0-1.0) [0.24]	$I^2=86.1\%$; <0.001

C.1. Intradermal tests (immediate reactions)

	N studies	Pooled meta-analytic value (95%CI) [I^2 ; Q Cochran p value]	Multivariable meta-regression model	
			Variables in the model (OR; 95%CI) [p value]	Residual heterogeneity
% positive results	16	15.0% (8.1-27.6%) [96.3%; <0.001]	- Europe (region) (3.2; 2.2-4.7) [<0.001] - Convenience Sample (10.4; 7.3-14.9) [<0.001] - Cephalosporins Testing (0.1; 0.03-0.4) [0.002] - Reported Allergy to Beta-lactams (2.9; 1.9-4.5) [<0.001]	$I^2=0\%$; 0.65
Sensitivity	3	24.8% (9.1-67.5%) [28.7%; 0.22]	- ^a	- ^a

Specificity	1	98.5% (94.6-100%)	- ^d	- ^d
PPV	4	71.9% (53.1-97.3%) [0%; 0.84]	- ^a	- ^a
NPV	15	95.4% (93.4-97.5%) [38.0%; <0.001]	- Europe (region) (1.0; 0.9-1.0) [0.002] - Australia and Far East (region) (0.9; 0.7-1.0) [0.06] - Suspected Drug Testing (1.0; 0.9-1.0) [0.009]	$I^2=0.4\%$; 0.02

C.2. Intradermal tests (non-immediate reactions)

	N studies	Pooled meta-analytic value (95% CI) [I^2 ; Q Cochran p value]	Multivariable meta-regression model	
			Variables in the model (OR; 95% CI) [p value]	Residual heterogeneity
% positive results	16	7.5% (3.9-14.4%) [97.6%; <0.001]	- Children (age group) (0.1; 0.04-0.3) [<0.001] - Adults (age group) (0.2; 0.1-0.8) [0.02] - North America (region) (0.03; 0.01-0.2) [<0.001] - Middle East (region) (0.1; 0.04-0.6) [0.009] - Convenience Sample (3.1; 1.1-8.6) [0.03]	$I^2=87.9\%$; <0.001
Sensitivity	7	31.4% (14.1-69.8%) [90.0%; <0.001]	- ^c	- ^c
Specificity	6	95.8% (92.8-99.0%) [75.3%; 0.001]	Not performed. No substantial heterogeneity was detected when performing subgroup analysis (i) of the 2 studies assessing only participants reporting allergy to aminopenicillins (99.1%; 95%CI=97.6-100%; $I^2=0\%$; $p=0.61$), or (ii) of the 2 studies assessing a convenience sample (91.5%; 95%CI=85.8-97.5%; $I^2=0\%$; $p=0.99$)	
PPV	7	57.6% (35.9-92.2%) [80.1%; <0.001]	Not performed. No substantial heterogeneity was detected when performing subgroup analysis (i) of the 2 studies assessing only adult patients (80.9%; 95%CI=52.7-100%; $I^2=0\%$; $p=0.33$), (ii) of the 2 studies assessing only children (54.5%; 95%CI=35.2-84.5%; $I^2=0\%$; $p=0.56$), (iii) of the 3 studies assessing a convenience sample (66.8%; 95%CI=46.4-96.2%; $I^2=15.3\%$; $p=0.37$), or (iv) of the 2 studies assessing only participants reporting allergy to aminopenicillins (97.1%; 95%CI=90.2-100%; $I^2=0\%$; $p=0.43$)	
NPV	16	93.4% (90.9-95.9%) [91.2%; <0.001]	- Adults (age group) (1.0; 0.9-1.0) [0.14] - North America (region) (1.1; 1.0-1.1) [0.16]	$I^2=84.9\%$; <0.001

E. Specific IgE

	N studies	Pooled meta-analytic value (95% CI) [I^2 ; Q Cochran p value]	Multivariable meta-regression model	
			Variables in the model (OR; 95% CI) [p value]	Residual heterogeneity
% positive results	34	6.8% (4.3-10.9%) [96.6%; <0.001]	- Year of publication (1.0; 0.9-1.0) [0.003] - Europe (region) (5.1; 1.1-24.7) [0.04] - Consecutive Testing (0.4; 0.1-1.1) [0.07]	$I^2=93.8\%$; <0.001
Sensitivity	11	22.9% (17.8-29.5%) [0%; 0.29]	- ^b	- ^b
Specificity	11	99.1% (98.5-99.7%) [19.2%; 0.54]	- ^b	- ^b
PPV	12	96.9% (93.0-100%) [0.1%; 0.15]	- ^b	- ^b
NPV	33	89.7% (85.1-94.6%) [98.5%; <0.001]	- Europe (region) (0.9; 0.9-1.0) [0.06] - Suspected Drug Testing (0.5; 0.4-0.6) [<0.001]	$I^2=88.9\%$; <0.001

E.1. Specific IgE (immediate reactions)

	N studies	Pooled meta-analytic value (95% CI) [I^2 ; Q	Multivariable meta-regression model
--	-----------	---	-------------------------------------

Cochran <i>p</i> value]			Variables in the model (OR; 95%CI) [<i>p</i> value]	Residual heterogeneity
% positive results	11	13.5% (6.8-26.7%) [88.3%; <0.001]	- Consecutive History (6.2; 2.4-16.2) [<0.001] - Convenience Sample (13.5; 5.2-35.1) [<0.001] - Reported Allergy to Penicillins (3.4; 1.5-7.9) [0.004]	$I^2=0\%$; 0.73
Sensitivity	5	31.0% (21.5-44.8%) [23.5%; 0.50]	- ^a	- ^a
Specificity	4	98.3% (95.4-100%) [0%; 0.72]	- ^a	- ^a
PPV	5	97.3% (92.5-100%) [0%; 0.77]	- ^a	- ^a
NPV	11	73.1% (55.6-96.0%) [99.6%; <0.001]	- Aminopenicillins Testing (0.5; 0.4-0.6) [<0.001] - Reported Allergy to Penicillins (0.5; 0.4-0.6) [<0.001]	$I^2=35.1\%$; 0.001
E.2. Specific IgE (non-immediate reactions)				
<i>N</i> studies	Pooled meta-analytic value (95%CI) [I^2 ; Q Cochran <i>p</i> value]	Multivariable meta-regression model		
		Variables in the model (OR; 95%CI) [<i>p</i> value]		Residual heterogeneity
% positive results	8	9.4% (3.1-28.3%) [92.3%; <0.001]	Not performed. No substantial heterogeneity was detected when performing subgroup analysis of the 5 studies in which aminopenicillins allergy was tested (4.6%; 95%CI=2.8-7.4%; $I^2=0\%$; $p=0.56$)	
Sensitivity	3	13.4% (2.4-74.2%) [57.7%; 0.09]	- ^e	- ^e
Specificity	2	97.4% (94.2-100%) [0%; 0.62]	- ^a	- ^a
PPV	4	96.8% (90.0-100%) [0%; 0.45]	- ^a	- ^a
NPV	6	89.6% (83.0-96.8%) [62.8%; 0.001]	Not performed. No substantial heterogeneity was detected when performing subgroup analysis (i) of the 4 studies assessing a convenience sample (94.9%; 95%CI=88.6-100%; $I^2=0.01\%$; $p=0.04$), (iii) of the 4 studies assessing participants specifically reporting allergy to a penicillin (82.7%; 95%CI=77.8-87.9%; $I^2=0.05\%$; $p=0.05$), or (iii) of the 2 studies assessing only participants reporting allergy to aminopenicillins (95.6%; 95%CI=89.1-100%; $I^2=0\%$; $p=0.55$)	

NPV: Negative predictive value; PPV: Positive predictive value; ^a Meta-regression not performed (≤ 10 primary studies) but heterogeneity is not substantial; ^b No variable selected for inclusion in a multivariable model; ^c Meta-regression not performed (< 10 primary studies) and subgroup analysis did not identify any subgroup with non-substantial heterogeneity; ^d Results calculated based on a single primary study; ^e Meta-regression and subgroup analyses not performed (despite substantial heterogeneity) due to a small number of included primary studies (≤ 3).

Table E5. Risk-of-bias summary for the included primary studies according to QUADAS-2 criteria. Green, yellow and red dots correspond, respectively, to low, unclear and high risk of bias

Study	Risk of bias				Applicability Concerns		
	Patient Selection	Index Test	Reference Standard	Flow and Timing	Patient Selection	Index Test	Reference Standard
Brown BC	●	●	●	●	●	●	●
Resnik SS	●	●	●	●	●	●	●
Levine BB*	●	●	●	●	●	●	●
Wide L	●	●	●	●	●	●	●
Warrington RJ*	●	●	●	●	●	●	●
Basomba A	●	●	●	●	●	●	●
Sullivan TJ	●	●	●	●	●	●	●
van Arsdell PP	●	●	●	●	●	●	●
Graff-Lonnevig V	●	●	●	●	●	●	●
Blanca M (1990)*	●	●	●	●	●	●	●
Surtees SJ	●	●	●	●	●	●	●
Sogn DD*	●	●	●	●	●	●	●
Gadde J	●	●	●	●	●	●	●
Juchet A	●	●	●	●	●	●	●
Romano A (1995)*	●	●	●	●	●	●	●
Sastre J*	●	●	●	●	●	●	●
Macy E (1997)	●	●	●	●	●	●	●
Cederbrant K*	●	●	●	●	●	●	●
Herve M	●	●	●	●	●	●	●
Loza Cortina C*	●	●	●	●	●	●	●
Pichichero ME	●	●	●	●	●	●	●
Blanca M (2001)*	●	●	●	●	●	●	●
Forrest DM*	●	●	●	●	●	●	●
Lebel B	●	●	●	●	●	●	●
Lopez Tiro JJ	●	●	●	●	●	●	●
Novalbos A	●	●	●	●	●	●	●
Romano A (2002)*	●	●	●	●	●	●	●
Solensky R	●	●	●	●	●	●	●
Torres MJ	●	●	●	●	●	●	●
Borch JE	●	●	●	●	●	●	●
Wong BBL	●	●	●	●	●	●	●
Bousquet PJ (2007)	●	●	●	●	●	●	●
Brusca I*	●	●	●	●	●	●	●
del Real GA	●	●	●	●	●	●	●
Fontaine C	●	●	●	●	●	●	●
Rebello Gomes E	●	●	●	●	●	●	●

Bousquet PJ (2008)	●	●	●	●	●	●	●
Goldberg, A*	●	●	●	●	●	●	●
Nolan RC	●	●	●	●	●	●	●
Padial A	●	●	●	●	●	●	●
Rodriguez-Alvarez M	●	●	●	●	●	●	●
Tantikul C*	●	●	●	●	●	●	●
Hershkovich J	●	●	●	●	●	●	●
Silva R	●	●	●	●	●	●	●
Caubet JC (2010)*	●	●	●	●	●	●	●
Chambel M	●	●	●	●	●	●	●
Macy E (2010)*	●	●	●	●	●	●	●
Holm A	●	●	●	●	●	●	●
Kamboj S	●	●	●	●	●	●	●
Moral L	●	●	●	●	●	●	●
Na HR	●	●	●	●	●	●	●
Ponvert C	●	●	●	●	●	●	●
Richter AG*	●	●	●	●	●	●	●
Seitz CS	●	●	●	●	●	●	●
Celik GE	●	●	●	●	●	●	●
Garcia Nunez I	●	●	●	●	●	●	●
Hjortlund J (2012)	●	●	●	●	●	●	●
Iglesias-Souto J	●	●	●	●	●	●	●
Kopac P*	●	●	●	●	●	●	●
Soto L	●	●	●	●	●	●	●
Erkocoglu M	●	●	●	●	●	●	●
Hjortlund J (2013)	●	●	●	●	●	●	●
Macy E (2013)	●	●	●	●	●	●	●
Rimawi RH	●	●	●	●	●	●	●
Al-Ahmad M	●	●	●	●	●	●	●
Bourke J	●	●	●	●	●	●	●
Caubet JC (2014)*	●	●	●	●	●	●	●
Colas H	●	●	●	●	●	●	●
Fox SJ	●	●	●	●	●	●	●
Picard M	●	●	●	●	●	●	●
Trautmann A	●	●	●	●	●	●	●
Zambonino MA	●	●	●	●	●	●	●
Barni S*	●	●	●	●	●	●	●
Blanca-Lopez N	●	●	●	●	●	●	●
Ferre-Ybarz L	●	●	●	●	●	●	●
Mori F*	●	●	●	●	●	●	●

Rosenfield L*	●	●	●	●	●	●	●
Abrams EM	●	●	●	●	●	●	●
Atanaskovic-Markovic M	●	●	●	●	●	●	●
Chen JR	●	●	●	●	●	●	●
Chiriac AM	●	●	●	●	●	●	●
King EA	●	●	●	●	●	●	●
Manuyakorn W	●	●	●	●	●	●	●
Mawhirt SL*	●	●	●	●	●	●	●
Meng J*	●	●	●	●	●	●	●
Mill C	●	●	●	●	●	●	●
Moreno E	●	●	●	●	●	●	●
Mota I	●	●	●	●	●	●	●
Ratzon R	●	●	●	●	●	●	●
Confino-Cohen R*	●	●	●	●	●	●	●
Faitelson Y	●	●	●	●	●	●	●
Fazlollahi MR	●	●	●	●	●	●	●
Juri MC	●	●	●	●	●	●	●
Marwood J	●	●	●	●	●	●	●
McDanel DL	●	●	●	●	●	●	●
Misirlioglu ED	●	●	●	●	●	●	●
Ramsey A	●	●	●	●	●	●	●
Soria A	●	●	●	●	●	●	●
Sundquist BK	●	●	●	●	●	●	●
Tannert LK*	●	●	●	●	●	●	●
Vyles D	●	●	●	●	●	●	●
Ibanez MD*	●	●	●	●	●	●	●
Marraccini P	●	●	●	●	●	●	●
Moussa Y	●	●	●	●	●	●	●
Sedlackova L	●	●	●	●	●	●	●

* Studies included in bivariate random-effects meta-analyses

References

1. Brown BC, Price EV, Moore MB, Jr. Penicilloyl-Polylysine as an Intradermal Test of Penicillin Sensitivity. *JAMA* 1964; 189(8): 599–604.
2. Resnik SS, Shelley WB. Penicillin hypersensitivity: detection by basophil response to challenge. *J Invest Dermatol* 1965; 45(4): 269–74.
3. Levine BB, Zolov DM. Prediction of penicillin allergy by immunological tests. *J Allergy* 1969; 43(4): 231–44.
4. Wide L, Juhlin L. Detection of penicillin allergy of the immediate type by radioimmunoassay of reagins (IgE) to penicilloyl conjugates. *Clin Exp Allergy* 1971; 1(2): 171–7.
5. Warrington RJ, Simons FER, Ho HW, Gorski BA. Diagnosis of penicillin allergy by skin testing: The manitoba experience. *CMAJ* 1978; 118(7): 787–91.
6. Basomba A, Villalmanzo IG, Campos A, Pelaez A, Berglund A. IgE antibodies against penicillin as determined by Phadebas RAST. *Clin Allergy* 1979; 9(5): 515–25.
7. Sullivan TJ, Wedner HJ, Shatz GS, Yecies LD, Parker CW. Skin testing to detect penicillin allergy. *J Allergy Clin Immunol* 1981; 68(3): 171–80.
8. VanArsdel Jr PP, Martonick GJ, Johnson LE, Sprenger JD, Altman LC, Henderson Jr WR. The value of skin testing for penicillin allergy diagnosis. *West J Med* 1986; 144(3): 311–4.
9. Graff-Lonnevig V, Hedlin G, Lindfors A. Penicillin allergy- A rare paediatric condition? *Arch Dis Child* 1988; 63(11): 1342–6.
10. Blanca M, Vega JM, Garcia J, et al. Allergy to penicillin with good tolerance to other penicillins; study of the incidence in subjects allergic to betalactams. *Clin Exp Allergy* 1990; 20(5): 475–81.
11. Surtees SJ, Stockton MG, Gietzen TW. Allergy to penicillin: fable or fact? *BMJ* 1991; 302(6784): 1051–2.
12. Sogn DD, Evans R, III, 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. *Arch Intern Med* 1992; 152(5): 1025–32.
13. Gadde J, Spence M, Wheeler B, Adkinson NF. Clinical-experience with penicillin skin testing in a large inner-city STD clinic. *JAMA* 1993; 270(20): 2456–63.
14. Juchet A, Rancé F, Brémont F, Dutau G. Investigation of beta-lactam allergy in 45 children. *Revue française d'allergologie et d'immunologie clinique* 1994; 34(5): 369–75.
15. Romano A, Di Fonso M, Papa G, et al. Evaluation of adverse cutaneous reactions to aminopenicillins with emphasis on those manifested by maculopapular rashes. *Allergy* 1995; 50(2): 113–8.
16. Sastre J, Quijano LD, Novalbos A, et al. Clinical cross-reactivity between amoxicillin and cephadroxil in patients allergic to amoxicillin and with good tolerance of penicillin. *Allergy* 1996; 51(6): 383–86.
17. Macy E, Richter PK, Falkoff R, Zeiger R. Skin testing with penicilloate and penilloate prepared by an improved method: Amoxicillin oral challenge in patients with negative skin test responses to penicillin reagents. *J Allergy Clin Immunol* 1997; 100(5): 586–91.
18. Cederbrant K, Stejskal V, Broman P, Lindkvist L, Sundell K. In vitro lymphocyte proliferation in the diagnosis of allergy to phenoxymethylpenicillin. *Allergy* 1998; 53(12): 1155–61.
19. Hervé M, De la Rocque F, Bouhanna A, Albengres E, Reinert P. Exploration of 112 children suspected of amoxicillin allergy. Indications and efficacy of oral provocation test. *Arch Pediatr* 1998; 5(5): 503–9.
20. Loza Cortina C. Adverse drug reactions in area I of Asturias. *An Esp Pediatr* 1998; 49(4): 359–63.
21. Pichichero ME, Pichichero DM. Diagnosis of penicillin, amoxicillin, and cephalosporin allergy: Reliability of examination assessed by skin testing and oral challenge. *J Pediatr* 1998; 132(1): 137–43.
22. Blanca M, Mayorga C, Torres MJ, et al. Clinical evaluation of Pharmacia CAP System (TM) RAST FEIA amoxicilloyl and benzylpenicilloyl in patients with penicillin allergy. *Allergy* 2001; 56(9): 862–70.
23. Forrest DM, Schellenberg RR, Thien VVS, King S, Anis AH, Dodek PM. Introduction of a practice guideline for penicillin skin testing improves the appropriateness of antibiotic therapy. *Clin Infect Dis* 2001; 32(12): 1685–90.
24. Lebel B, Messaad D, Kvedariene V, Rongier M, Bousquet J, Demoly P. Cysteinyl-leukotriene release test (CAST) in the diagnosis of immediate drug reactions. *Allergy* 2001; 56(7): 688–92.
25. López Tiro JJ, Orea Solano M, Flores Sandoval G, Gómez Vera J. Skin tests with high and minor determinants in patients with doubtful allergy to penicillin. *Rev Alerg Mex* 2001; 48(3): 80–7.

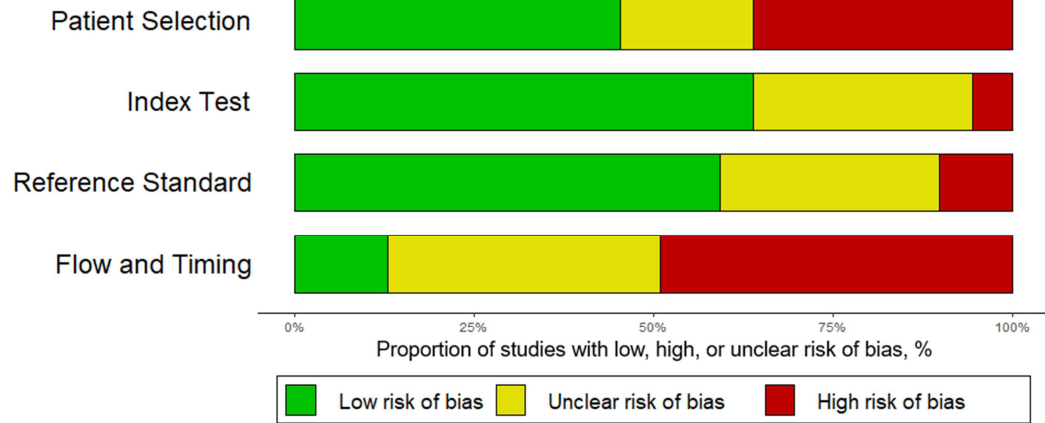
26. Novalbos A, Sastre J, Cuesta J, et al. Lack of allergic cross-reactivity to cephalosporins among patients allergic to penicillins. *Clin Exp Allergy* 2001; 31(3): 438–43.
27. Romano A, Viola M, Mondino C, et al. Diagnosing nonimmediate reactions to penicillins by in vivo tests. *Int Arch Allergy Immunol* 2002; 129(2): 169–74.
28. Solensky R, Earl HS, Gruchalla RS. Lack of penicillin resensitization in patients with a history of penicillin allergy after receiving repeated penicillin courses. *Arch Intern Med* 2002; 162(7): 822–6.
29. Torres MJ, Mayorga C, Leyva L, et al. Controlled administration of penicillin to patients with a positive history but negative skin and specific serum IgE tests. *Clin Exp Allergy* 2002; 32(2): 270–6.
30. Borch JE, Andersen KE, Bindslev-Jensen C. The prevalence of suspected and challenge-verified penicillin allergy in a university hospital population. *Basic Clin Pharmacol Toxicol* 2006; 98(4): 357–62.
31. Wong BBL, Keith PK, Waserman S. Clinical history as a predictor of penicillin skin test outcome. *Ann Allergy Asthma Immunol* 2006; 97(2): 169–74.
32. Bousquet PJ, Kvedariene V, Co-Minh HB, et al. Clinical presentation and time course in hypersensitivity reactions to β -lactams. *Allergy* 2007; 62(8): 872–6.
33. Brusca I, Corrao I, Sceusa G, et al. Flow cytometric basophil activation test in patients with allergic or pseudoallergic reactions to betalactamic antibiotics and non steroidal and inflammatory drugs. *Rivista Italiana della Medicina di Laboratorio* 2007; 3(3): 196–202.
34. del Real GA, Rose ME, Hammel J, Gordon SM, Arroliga ME. The penicillin skin test has a high negative predictive value and helps to modify the use of antibiotics in patients with history of beta-lactam allergy. *J Allergy Clin Immunol* 2007; 117(2): S224–S.
35. Fontaine C, Mayorga C, Bousquet PJ, et al. Relevance of the determination of serum-specific IgE antibodies in the diagnosis of immediate β -lactam allergy. *Allergy* 2007; 62(1): 47–52.
36. Rebelo Gomes E, Fonseca J, Araujo L, Demoly P. Drug allergy claims in children: From self-reporting to confirmed diagnosis. *Clin Exp Allergy* 2007; 38(1): 191–8.
37. Bousquet PJ, Pipet A, Bousquet-Rouanet L, Demoly P. Oral challenges are needed in the diagnosis of β -lactam hypersensitivity. *Clin Exp Allergy* 2008; 38(1): 185–90.
38. Goldberg A, Confino-Cohen R. Skin testing and oral penicillin challenge in patients with a history of remote penicillin allergy. *Ann Allergy Asthma Immunol* 2008; 100(1): 37–43.
39. Nolan RC, Puy R, Deckert K, O'Hehir RE, Douglass JA. Experience with a new commercial skin testing kit to identify IgE-mediated penicillin allergy. *Intern Med J* 2008; 38(5): 357–61.
40. Padial A, Antunez C, Blanca-Lopez N, et al. Non-immediate reactions to beta-lactams: diagnostic value of skin testing and drug provocation test. *Clin Exp Allergy* 2008; 38(5): 822–8.
41. Rodriguez-Alvarez M, Santos-Magadan S, Rodriguez-Jimenez B, de Arellano IRR, Vazquez-Cortes S, Martinez-Cocera C. Reproducibility of delayed-type reactions to betalactams. *Allergol Immunopathol (Madr)* 2008; 36(4): 201–4.
42. Tantikul C, Dhana N, Jongjarearnprasert K, Visitsunthorn N, Vichyanond P, Jirapongsananuruk O. The Utility of the World Health Organization-The Uppsala Monitoring Centre (WHO-UMC) System for the Assessment of Adverse Drug Reactions in Hospitalized Children. *Asian Pac J Allergy Immunol* 2008; 26(2-3): 77–82.
43. Hershkovich J, Broides A, Kirjner L, Smith H, Gorodischer R. Beta lactam allergy and resensitization in children with suspected beta lactam allergy. *Clin Exp Allergy* 2009; 39(5): 726–30.
44. Silva R, Cruz L, Botelho C, et al. Work up of patients with history of beta-lactam hypersensitivity. *Allergol Immunopathol (Madr)* 2009; 37(4): 193–7.
45. Caubet JC, Kaiser L, Lemaitre B, Fellay B, Gervaix A, Eigenmann PA. The role of penicillin in benign skin rashes in childhood: A prospective study based on drug rechallenge. *J Allergy Clin Immunol* 2010; 127(1): 218–22.
46. Chambel M, Martins P, Silva I, Palma-Carlos S, Romeira AM, Pinto PL. Drug provocation tests to betalactam antibiotics: experience in a paediatric setting. *Allergol Immunopathol (Madr)* 2010; 38(6): 300–6.
47. Macy E, Goldberg B, Poon KY. Use of commercial anti-penicillin IgE fluorometric enzyme immunoassays to diagnose penicillin allergy. *Ann Allergy Asthma Immunol* 2010; 105(2): 136–41.
48. Holm A, Mosbech H. Challenge Test Results in Patients With Suspected Penicillin Allergy, but No Specific IgE. *Allergy Asthma Immunol Res* 2011; 3(2): 118–22.
49. Kamboj S, Yousef E, McGeady S, Hossain J. The prevalence of antibiotic skin test reactivity in a pediatric population. *Allergy Asthma Proc* 2011; 32(2): 99–105.
50. Moral L, Garde J, Toral T, Fuentes MJ, Marco N. Short protocol for the study of paediatric patients with suspected betalactam antibiotic hypersensitivity and low risk criteria. *Allergol Immunopathol (Madr)* 2011; 39(6): 337–41.
51. Na HR, Lee JM, Jung JW, Lee SY. Usefulness of drug provocation tests in children with a history of adverse drug reaction. *Korean J Pediatr* 2011; 54(7): 304–9.

52. Ponvert C, Perrin Y, Bados-Albiero A, et al. Allergy to betalactam antibiotics in children: Results of a 20-year study based on clinical history, skin and challenge tests. *Pediatr Allergy Immunol* 2011; 22(4): 411–8.
53. Richter AG, Wong G, Goddard S, et al. Retrospective case series analysis of penicillin allergy testing in a UK specialist regional allergy clinic. *J Clin Pathol* 2011; 64(11): 1014–8.
54. Seitz CS, Brocker EB, Trautmann A. Diagnosis of drug hypersensitivity in children and adolescents: discrepancy between physician-based assessment and results of testing. *Pediatr Allergy Immunol* 2011; 22(4): 405–10.
55. Celik GE, Aydin O, Dogu F, et al. Diagnosis of immediate hypersensitivity to β -lactam antibiotics can be made safely with current approaches. *Int Arch Allergy Immunol* 2012; 157(3): 311–7.
56. Garcia Nunez I, Barasona Villarejo MJ, Algaba Marmol MA, Moreno Aguilar C, Guerra Pasadas F. Diagnosis of patients with immediate hypersensitivity to beta-lactams using retest. *J Investig Allergol Clin Immunol* 2012; 22(1): 41–7.
57. Hjortlund J, Mortz CG, Skov PS, et al. One-week oral challenge with penicillin in diagnosis of penicillin allergy. *Acta Derm Venereol* 2012; 92(3): 307–12.
58. Iglesias-Souto J, Gonzalez R, Poza P, Sanchez-Machin I, Matheu V. Evaluating the usefulness of retesting for beta-lactam allergy in children. *Pediatr Infect Dis J* 2012; 31(10): 1091–3.
59. Kopac P, Zidarn M, Kosnik M. Epidemiology of hypersensitivity reactions to penicillin in Slovenia. *Acta Dermatovenereol Alp Pannonica Adriat* 2012; 21(4): 5–7.
60. Soto L, Guilarte M, Luengo O, Labrador M, Sala A. Drug hypersensitivity to betalactams: a retrospective study to assess a diagnostic challenge. *Allergy* 2012; 67: 429–.
61. Erkocoglu M, Kaya A, Civelek E, et al. Prevalence of confirmed immediate type drug hypersensitivity reactions among school children. *Pediatr Allergy Immunol* 2013; 24(2): 160–7.
62. Hjortlund J, Mortz CG, Skov PS, Bindselev-Jensen C. Diagnosis of penicillin allergy revisited: The value of case history, skin testing, specific IgE and prolonged challenge. *Allergy* 2013; 68(8): 1057–64.
63. Macy E, Ngor EW. Safely diagnosing clinically significant penicillin allergy using only penicilloyl-poly-lysine, penicillin, and oral amoxicillin. *J Allergy Clin Immunol Pract* 2013; 1(3): 258–63.
64. Rimawi RH, Cook PP, Gooch M, et al. The impact of penicillin skin testing on clinical practice and antimicrobial stewardship. *J Hosp Med* 2013; 8(6): 341–5.
65. Al-Ahmad M, Rodriguez Bouza T, Arifhodzic N. Penicillin allergy evaluation: experience from a drug allergy clinic in an Arabian Gulf Country, Kuwait. *Asia Pac Allergy* 2014; 4(2): 106–12.
66. Bourke J, Pavlos R, James I, Phillips E. Improving the Effectiveness of Penicillin Allergy De-labeling. *J Allergy Clin Immunol Pract* 2014; 3(3): 365–74.e1.
67. Caubet JC, Frossard C, Fellay B, Eigenmann PA. Skin tests and in vitro allergy tests have a poor diagnostic value for benign skin rashes due to β -lactams in children. *Pediatr Allergy Immunol* 2014; 26(1): 80–2.
68. Colas H, David V, Molle I, Bernier C, Magnan A, Pipet A. Drug provocation tests for children: In-hospital or in an outpatient consultation? A protocol comparing the two possibilities. *Revue Francaise D Allergologie* 2014; 54(4): 300–6.
69. Fox SJ, Park MA. Penicillin skin testing is a safe and effective tool for evaluating penicillin allergy in the pediatric population. *J Allergy Clin Immunol Pract* 2014; 2(4): 439–44.
70. Picard M, Paradis L, Bégin P, Paradis J, Des Roches A. Skin testing only with penicillin G in children with a history of penicillin allergy. *Ann Allergy Asthma Immunol* 2014; 113(1): 75–81.
71. Trautmann A, Seitz CS, Stoevesandt J, Kerstan A. Aminopenicillin-associated exanthem: lymphocyte transformation testing revisited. *Clin Exp Allergy* 2014; 44(12): 1531–8.
72. Zambonino MA, Corzo JL, Munoz C, et al. Diagnostic evaluation of hypersensitivity reactions to beta-lactam antibiotics in a large population of children. *Pediatr Allergy Immunol* 2014; 25(1): 80–7.
73. Barni S, Mori F, Sarti L, et al. Utility of skin testing in children with a history of non-immediate reactions to amoxicillin. *Clin Exp Allergy* 2015; 45(9): 1472–4.
74. Blanca-Lopez N, Perez-Alzate D, Ruano F, et al. Selective immediate responders to amoxicillin and clavulanic acid tolerate penicillin derivative administration after confirming the diagnosis. *Allergy* 2015; 70(8): 1013–9.
75. Ferré-Ybarz L, Salinas Argente R, Gómez Galán C, Duocastella Selvas P, Nevot Falcó S. Analysis of profitability in the diagnosis of allergy to beta-lactam antibiotics. *Allergol Immunopathol (Madr)* 2015; 43(4): 369–75.
76. Mori F, Cianferoni A, Barni S, Pucci N, Rossi ME, Novembre E. Amoxicillin Allergy in Children: Five-Day Drug Provocation Test in the Diagnosis of Nonimmediate Reactions. *J Allergy Clin Immunol Pract* 2015; 3(3): 375–+.

77. Rosenfield L, Kalicinsky C, Warrington R. A retrospective comparison of false negative skin test rates in penicillin allergy, using pencilloyl-poly-lysine and minor determinants or Penicillin G, followed by open challenge. *Allergy Asthma Clin Immunol* 2015; 11: 34.
78. Abrams EM, Wakeman A, Gerstner TV, Warrington RJ, Singer AG. Prevalence of beta-lactam allergy: a retrospective chart review of drug allergy assessment in a predominantly pediatric population. *Allergy Asthma Clin Immunol* 2016; 12.
79. Atanaskovic-Markovic M, Gaeta F, Medjo B, et al. Non-immediate hypersensitivity reactions to beta-lactam antibiotics in children - our 10-year experience in allergy work-up. *Pediatr Allergy Immunol* 2016; 27(5): 533–8.
80. Chen JR, Tarver SA, Alvarez KS, Tran T, Khan DA. A Proactive Approach to Penicillin Allergy Testing in Hospitalized Patients. *J Allergy Clin Immunol Pract* 2016; 5(3): 686–93.
81. Chiriac AM, Rerkpattanapipat T, Bousquet PJ, Molinari N, Demoly P. Optimal step doses for drug provocation tests to prove beta-lactam hypersensitivity. *Allergy* 2016; 72(4): 552–61.
82. King EA, Challa S, Curtin P, Bielory L. Penicillin skin testing in hospitalized patients with beta-lactam allergies: Effect on antibiotic selection and cost. *Ann Allergy Asthma Immunol* 2016; 117(1): 67–71.
83. Manuyakorn W, Singvijarn P, Benjaponpitak S, et al. Skin testing with β -lactam antibiotics for diagnosis of β -lactam hypersensitivity in children. *Asian Pac J Allergy Immunol* 2016; 34(3): 242–7.
84. Mawhirt SL, Fonacier LS, Calixte R, Davis-Lorton M, Aquino MR. Skin testing and drug challenge outcomes in antibiotic-allergic patients with immediate-type hypersensitivity. *Ann Allergy Asthma Immunol* 2016; 118(1): 73–9.
85. Meng J, Thursfield D, Lukawska JJ. Allergy test outcomes in patients self-reported as having penicillin allergy: Two-year experience. *Ann Allergy Asthma Immunol* 2016; 117(3): 273–9.
86. Mill C, Primeau MN, Medoff E, et al. Assessing the Diagnostic Properties of a Graded Oral Provocation Challenge for the Diagnosis of Immediate and Nonimmediate Reactions to Amoxicillin in Children. *JAMA Pediatr* 2016; 170(6): e160033.
87. Moreno E, Laffond E, Munoz-Bellido F, et al. Performance in real life of the European Network on Drug Allergy algorithm in immediate reactions to beta-lactam antibiotics. *Allergy* 2016; 71(12): 1787–90.
88. Mota I, Gaspar A, Chambel M, Piedade S, Morais-Almeida M. Hypersensitivity to beta-lactam antibiotics: a three-year study. *Eur Ann Allergy Clin Immunol* 2016; 48(6): 212–9.
89. Ratzon R, Reshef A, Efrati O, et al. Impact of an extended challenge on the effectiveness of β -lactam hypersensitivity investigation. *Ann Allergy Asthma Immunol* 2016; 116(4): 329–33.
90. Confino-Cohen R, Rosman Y, Meir-Shafir K, et al. Oral Challenge without Skin Testing Safely Excludes Clinically Significant Delayed-Onset Penicillin Hypersensitivity. *J Allergy Clin Immunol Pract* 2017; 5(3): 669–75.
91. Faitelson Y, Boaz M, Dalal I. Asthma, Family History of Drug Allergy, and Age Predict Amoxicillin Allergy in Children. *J Allergy Clin Immunol Pract* 2017; 6(4): 1363–7.
92. Fazlollahi MR, Bidad K, Shokouhi R, et al. Frequency and Risk Factors of Penicillin and Amoxicillin Allergy in Suspected Patients with Drug Allergy. *Arch Iran Med* 2017; 20(1): 34–7.
93. Juri MC, Romero DSF, Larrauri B, Malbran E, Torre G, Malbran A. Allergy to drugs. Experience in 771 procedures. *Medicina-Buenos Aires* 2017; 77(3): 180–4.
94. Marwood J, Aguirrebarrena G, Kerr S, Welch SA, Rimmer J. De-labelling self-reported penicillin allergy within the emergency department through the use of skin tests and oral drug provocation testing. *Emerg Med Australas* 2017; 29(5): 509–15.
95. McDanel DL, Azar AE, Dowden AM, et al. Screening for Beta-Lactam Allergy in Joint Arthroplasty Patients to Improve Surgical Prophylaxis Practice. *Journal Arthroplasty* 2017; 32(9s): S101–s8.
96. Misirlioglu ED, Guvenir H, Toyran M, et al. Frequency of selective immediate responders to aminopenicillins and cephalosporins in Turkish children. *Allergy Asthma Proc* 2017; 38(5): 376–82.
97. Ramsey A, Staicu ML. Use of a Penicillin Allergy Screening Algorithm and Penicillin Skin Testing for Transitioning Hospitalized Patients to First-Line Antibiotic Therapy. *J Allergy Clin Immunol Pract* 2017; 6(4): 1349–55.
98. Soria A, Autegarden E, Amsler E, et al. A clinical decision-making algorithm for penicillin allergy. *Ann Med* 2017; 49(8): 710–7.
99. Sundquist BK, Bowen BJ, Otabor U, Celestin J, Sorum PC. Proactive penicillin allergy testing in primary care patients labeled as allergic: outcomes and barriers. *Postgrad Med* 2017; 129(8): 915–20.
100. Tannert LK, Falkencrone S, Mortz CG, Bindeslev-Jensen C, Skov PS. Is a positive intracutaneous test induced by penicillin mediated by histamine? A cutaneous microdialysis study in penicillin-allergic patients. *Clin Transl Allergy* 2017; 7: 40.

101. Vyles D, Adams J, Chiu A, Simpson P, Nimmer M, Brousseau DC. Allergy Testing in Children With Low-Risk Penicillin Allergy Symptoms. *Pediatr* 2017; 140(2).
102. Ibanez MD, Del Rio PR, Lasa EM, et al. Prospective assessment of diagnostic tests for pediatric penicillin allergy, from clinical history to challenge tests. *Ann Allergy Asthma Immunol* 2018.
103. Marraccini P, Pignatti P, A DAA, Salimbeni R, Consonni D. Basophil Activation Test Application in Drug Hypersensitivity Diagnosis: An Empirical Approach. *Int Arch Allergy Immunol* 2018; 1–7.
104. Moussa Y, Shuster J, Matte G, et al. De-labeling of beta-lactam allergy reduces intraoperative time and optimizes choice in antibiotic prophylaxis. *Surgery* 2018.
105. Sedlackova L, Prucha M, Polakova I, Mikova B. Evaluation of IFN-gamma Enzyme-linked Immunospot Assay (ELISPOT) as a First-line Test in the Diagnosis of Non-Immediate Hypersensitivity to Amoxicillin and Penicillin. *Prague Med Rep* 2018; 119(1): 30–42.
106. Zucchini W. An introduction to model selection. *J Math Psychol* 2000;44:41-61.

QUADAS-2 Domain



QUADAS-2 Domain

