APS European Society for Paediatric Research SPR Society for Pediatric Research

SYSTEMATIC REVIEW

Beta-lactam allergy and drug challenge test in children: a systematic review and meta-analysis

Yasutaka Kuniyoshi[1,2]✉, Yasushi Tsujimoto[2,3,4], Masahiro Banno[2,5,6], Shunsuke Taito[2,7], Takashi Ariie[2,8], Takafumi Kubota[2,9], Natsuki Takahashi[1] and Haruka Tokutake[1]

BACKGROUND: Most cases of beta-lactam allergy in children are likely to be mislabeled. This study aimed to assess the prevalence of true positives, as determined by drug challenge tests, and the rate of false negatives in children with suspected allergies and confirm the safety of the drug challenge test.

METHODS: We conducted a systematic review and meta-analysis according to established procedures. Study participants were children with suspected beta-lactam allergy who underwent a drug challenge. PubMed MEDLINE, Dialog EMBASE, Cochrane Central Register of Controlled Trials, World Health Organization International Clinical Trials Registry Platform, and clinicaltrials.gov were searched from inception until March 5, 2021.

RESULTS: The pooled prevalence of (a) positive results in the first challenge was 0.049 (95% CI, 0.041–0.057; $I^2 = 71\%$) from 78 studies; (b) serious adverse events was 0.00 (95% CI, 0.00–0.00; $I^2 = 0.0\%$) from 62 studies; and (c) positive results in the second challenge after the first negative result was 0.028 (95% CI, 0.016–0.043; $I^2 = 38\%$) from 18 studies.

CONCLUSIONS: The prevalence of children with suspected beta-lactam allergy with true-positive results and false-negative results from the drug challenge test was very low. Serious adverse events resulting from drug challenge tests were also very rare.

Pediatric Research (2023) 93:22–30; https://doi.org/10.1038/s41390-022-02076-x

IMPACT:

- Most children with suspected beta-lactam allergy were likely to be mislabeled.
- Serious adverse events caused by the drug challenge test were rare.
- Few false-negative results were obtained from the drug challenge test.

INTRODUCTION

The diagnosis of beta-lactam allergy without allergic workup for confirmation is a public health problem which is mostly observed among children. Approximately 2–5% of the pediatric population have self-reported beta-lactam allergy.[1] However, most children with suspected beta-lactam allergy are likely to be overlabeled.[2,3] Misinterpretation of beta-lactam allergy could increase antibiotic resistance and healthcare costs due to the use of broad-spectrum antibiotics.[4,5]

The drug challenge test is considered the gold standard to diagnose drug allergy.[1,2] Although skin tests are commonly performed before the drug challenge, the diagnostic accuracy of skin tests is considered insufficient; in particular, they have low sensitivity.[6,7] A previous meta-analysis published in 2015 determined that children with suspected beta-lactam allergy were truly allergic if either the skin or drug challenge test was positive with an immediate reaction.[3] The meta-analysis focused only on immediate reactions and included only four original studies, which reported prevalence rates between 0.9% and 14.8%. Since 2015, an increasing number of studies have assessed the true-positive rate for subgroups with a history of immediate or nonimmediate reactions[8,9] using a direct drug challenge without performing a prior skin test.[10–12]

To the best of our knowledge, there has been no systematic review and meta-analysis that has examined the child prevalence of true beta-lactam allergy in suspected cases based on the results of drug challenges. Neither safety nor the predictive negative value of the drug challenge has been thoroughly evaluated in previous systematic reviews and meta-analyses. We, therefore, performed a systematic review and meta-analysis to assess: (a) the prevalence of positive results from drug challenge test for children with suspected beta-lactam allergy; (b) the prevalence of serious

[1]Department of Pediatrics, Tsugaruhoken Medical COOP Kensei Hospital, 2-2-2 Ougi-machi, Hirosaki, Aomori 036-8511, Japan. [2]Scientific Research WorkS Peer Support Group (SRWS-PSG), Osaka, Japan. [3]Department of Health Promotion and Human Behavior, Kyoto University Graduate School of Medicine/School of Public Health, Yoshida Konoe cho, Sakyo-ku, Kyoto 606-8501, Japan. [4]Department of Nephrology and Dialysis, Kyoritsu Hospital, Chuo-Cho 16-5, Kawanishi, Hyogo 666-0016, Japan. [5]Department of Psychiatry, Seichiryo Hospital, Tsurumai 4-16-27, Showa-ku, Nagoya 466-0064, Japan. [6]Department of Psychiatry, Nagoya University Graduate School of Medicine, Tsurumai-cho 65, Showa-ku, Nagoya 466-8560, Japan. [7]Division of Rehabilitation, Department of Clinical Practice and Support, Hiroshima University Hospital, 1-2-3, Kasumi, Minami-ku, Hiroshima 734-8551, Japan. [8]Department of Physical Therapy, School of Health Sciences at Fukuoka, International University of Health and Welfare, 137-1 Enokizu, Okawa-shi, Fukuoka 831-8501, Japan. [9]Department of Neurology, Tohoku University Graduate School of Medicine, 2-1 Seiryo-machi, Aoba-ku, Sendai 980-8575, Japan. ✉email: yasutakakuniyoshi@yahoo.co.jp

Received: 13 January 2022 Revised: 16 March 2022 Accepted: 22 March 2022
Published online: 21 April 2022

adverse events in response to a drug challenge test; and (c) the prevalence of positive results from a second drug challenge result among children with negative results from a first drug challenge.

METHODS

Protocol and registration

We conducted a systematic review and meta-analysis in accordance with the Preferred Reporting Item for Systematic Review and Meta-analysis guidelines (PRISMA 2020).[13] We developed a detailed protocol and registered our study with the International Prospective Register of Systematic Reviews (PROSPERO: CRD42021240535).

Article inclusion criteria

We included all cohort studies. We did not restrict publication date and status (full publications, conference abstracts, and unpublished data), or languages. We excluded case reports, case series, case–control studies, and reviews. Study participants were children under the age of 18 with suspected beta-lactam allergy who underwent a drug challenge test. We excluded studies in which the participants were mostly adults, however, included studies involving both adults and children where data on children could be extracted independently. We accepted other definitions of children and differing procedures for administering the drug challenge test. We allowed the determination of results used in the original studies. We also excluded studies with (i) overlapping participants, (ii) unclear information on the suspected drug, or (iii) unclear information on the challenge test.

Primary outcomes

The primary outcomes were (a) the prevalence of positive results from a drug challenge test for children with suspected beta-lactam allergy; (b) the prevalence of serious adverse events in response to a drug challenge test; and (c) the prevalence of positive results from a second drug challenge among children with negative results from a first drug challenge.

We defined the prevalence of positive results in the first challenge as the number of positive first drug challenge results per total number of first drug challenge tests. We accepted the criteria adopted by the original studies for the results of the drug challenge test.

The prevalence of serious adverse events was defined as the number of serious adverse events per total number of first drug challenge tests. We accepted the definition adopted by the original studies for serious adverse events. If the authors did not assess the severity of adverse events, we defined serious adverse events as immediate reactions with anaphylaxis or requiring administration of epinephrine or delayed reactions such as Stevens–Johnson syndrome or toxic epidermal necrolysis.

The prevalence of positive results in the second challenge after a first negative result was defined as the number of positive second drug challenge results per total number of second drug challenge tests. Children who had a negative first drug challenge result underwent a second drug challenge test. Those who were positive were defined as having a positive second drug challenge result. We did not set a criterion for the interval between the first and second challenges. If a second drug challenge test was not performed, children who used the suspect drug and developed allergy-like symptoms after a first negative result were also classified as having a positive second drug challenge result.

Search strategy

We searched PubMed MEDLINE, Dialog EMBASE, the Cochrane Central Register of Controlled Trials, and the World Health Organization International Clinical Trials Platform Search Portal (ICTRP), and ClinicalTrials.gov from inception to March 5, 2021. We did not restrict to language or date of publication. Supplementary Table S1 shows the details of search terms used for each database. We also confirmed the references included in the previous systematic reviews,[2,3] clinical guidelines,[1,14] or position papers,[15,16] and studies included in the present review.

Study selection and data extraction

Two of the three review authors (YK, NT, and HT) screened the titles and abstracts of studies according to the inclusion and exclusion criteria and then obtained and independently assessed the full text for eligibility. If we were not sure whether a study met the inclusion criteria, we emailed the authors of original study and requested additional information. We used Rayyan software[17] for study selection. Two of the three review authors (YK, NT, and HT) independently extracted data from the studies. Any disagreements between the two reviewers were resolved through discussion or with a third reviewer if required.

Quality assessment

Two of the three review authors (YK, NT, and HT) independently assessed the risk of bias for each study by using the Joanna Briggs Institute prevalence critical appraisal tool.[18] This tool consists of nine domains. Each domain is evaluated as Yes, No, Unclear, or Not Applicable. The quality score for each study was defined as the percentage of Yes answers obtained for the total number of domains evaluated. Quality scores were graded as low quality if the percentage was <50%, moderate quality if it was 50–80%, and high quality if it was >80%.[19,20] Any conflict between the two reviewers were resolved through discussion or with a third reviewer if required.

Data synthesis and statistical analysis

All statistical analyses were performed using R software (version 4.0.3).[21] For continuous variables, data were expressed as median (interquartile range (IQR)). To estimate the pooled prevalence, we synthesized the data with the "meta" and "metafor" packages for R and combined proportions with 95% confidence intervals (CIs) and 95% prediction intervals. We used an inverse-variance-weighted random-effects model with the DerSimonian–Laird estimator to estimate the between-study variance and normal approximation intervals based on summary measures to calculate CIs for individual study results. To stabilize variances, we transformed the data with the Freeman–Tukey double arcsine transformation. The random-effects model (DerSimonian–Laird method) rather than fixed-effect model was selected at the time of the study protocol because of the expected heterogeneity of the included studies.

We visually evaluated heterogeneity using forest plots. We also calculated and analyzed I^2 statistics (I^2 values of 0–40% = might not be important; 30–60% = moderate heterogeneity; 50–90% = substantial heterogeneity; 75–100% = considerable heterogeneity). To calculate the I^2 statistics, we applied the Cochrane χ^2 test (Q-test). We also calculated τ^2 statistics for the prediction intervals.[22]

To assess the reporting bias, we searched clinical trial registry systems (ICTRP and ClinicalTrials.gov) to identify completed but unpublished studies. We assessed potential publication bias by visual inspection of a funnel plot and also performed the Egger test. Funnel plots are used to assess the presence of potential publication bias in meta-analyses. However, these are likely to be inaccurate in the meta-analyses of prevalence studies with low proportions of outcomes.[23] Therefore, we did not create funnel plots for meta-analyses that included <20 studies[24,25] or studies with similar sample sizes.

Subgroup analyses

We considered the following subgroups for the prevalence of positive results in the first challenge: children who had undergone the drug challenge with or without a preliminary test, including skin prick test, intradermal test, or specific IgE quantification;

children with a history of an immediate or nonimmediate reaction; children with a suspected penicillin class allergy or cephalosporin allergy; children who underwent a direct drug challenge test with a history of nonimmediate reaction. We also considered the following ad hoc subgroups for the prevalence of positive results in the second challenge after a first negative result: children who were evaluated with a second drug challenge test or with questionnaires.

RESULTS

Literature search results

Figure 1 displays a PRISMA 2020 flow chart detailing the process of study selection, including exclusion criteria. We screened 4112 records, after the removal of duplicates, and assessed 347 full texts. Eighty-four studies reporting a prevalence estimate were synthesized in the meta-analysis. In the clinical trial registry system, 49 records were screened, and 3 records were included. We found one unpublished study.

Study characteristics

Table 1 and Supplemental Table S2 show the characteristics of the 85 studies. Of them, 55 studies were published in or after 2015.

Results of synthesis

Prevalence of positive results in the first challenge. A total of 12,846 participants were included from 78 studies. The methodological quality of the 78 studies was moderate (Supplemental Table S3A). The reason for the moderate quality was inappropriate sampling processes. The prevalence from the included studies was 0.052 (0.028–0.077). The pooled prevalence was 0.049 (95% CI, 0.041–0.057; $I^2 = 71\%$; Fig. 2). Visual inspection of funnel plot asymmetry did not indicate the presence of publication bias

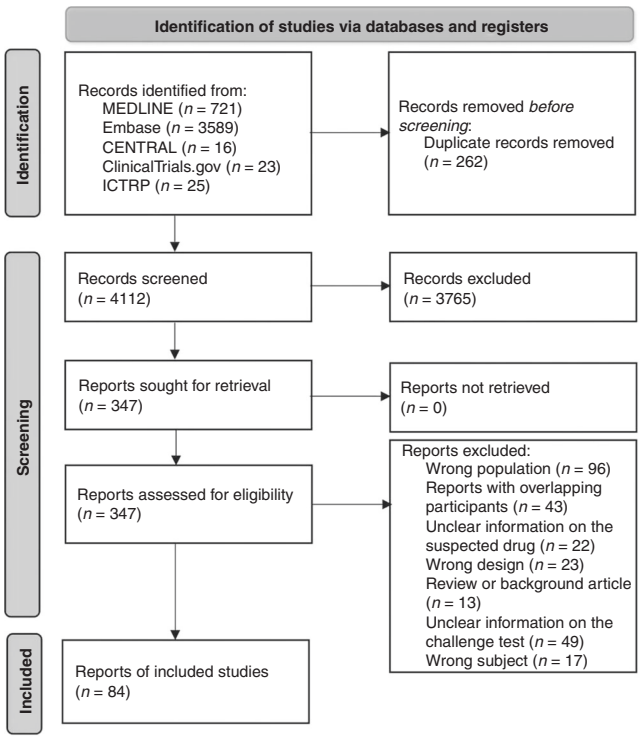

Fig. 1 Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) diagram for literature search and study selection. A list of the excluded studies with reasons is shown in Supplemental Table S4.

Table 1. Study characteristics.

Characteristics	N
Prevalence of positive results in the first challenge (n = 78)	
Publication type	
Article	55
Conference abstract	23
Publication year	
<2015	27
≥2015	51
Culprit drug	
Beta-lactam	46
Penicillin class only	30
Cephalosporin only	2
History of reaction	
Immediate only	4
Nonimmediate only	12
Both—Separated	9
Both—Not separated	53
Preliminary testing	
Performed only	38
Not performed only	24
Both—Separated	6
Not applicable	10
Prevalence of serious adverse events (n = 62)	
Publication type	
Article	47
Conference abstract	15
Publication year	
<2015	22
≥2015	40
Culprit drug	
Beta-lactam	33
Penicillin class only	27
Cephalosporin only	2
History of reaction	
Immediate only	2
Nonimmediate only	10
Both—Separated	10
Both—Not separated	40
Preliminary testing	
Performed only	31
Not performed only	19
Both—Separated	5
Not applicable	7
Prevalence of positive results in the second challenge after the first negative result (n = 18)	
Publication type	
Article	15
Conference abstract	3
Publication year	
<2015	6
≥2015	12
Culprit drug	
Beta-lactam	13
Penicillin class only	5
Confirmation method	
Drug challenge test	3
Questionnaire	15

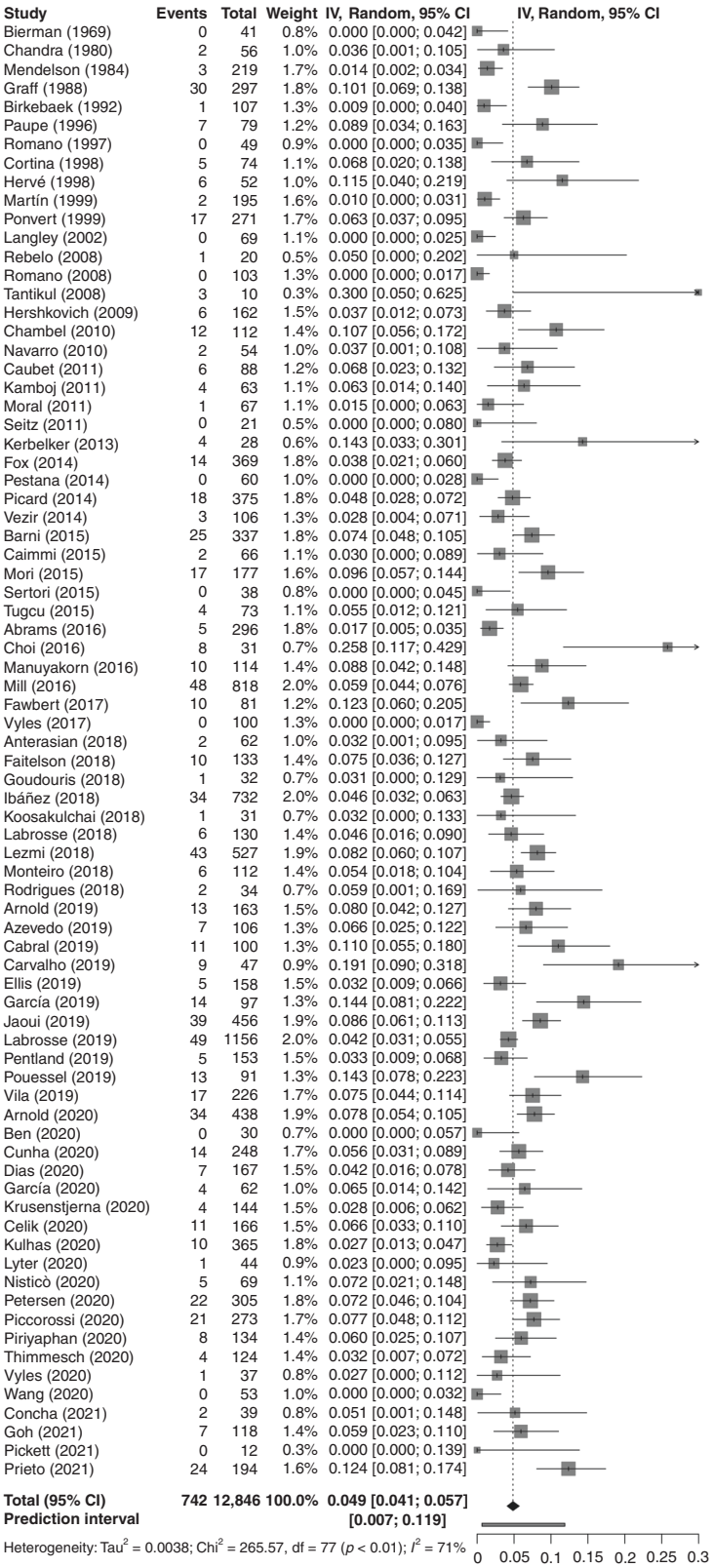

Fig. 2 Pooled prevalence of positive results for the first challenge. An inverse-variance-weighted random-effect meta-analysis was used. The size of the point estimate (square data markers) is proportional to the weight of the study. Horizontal lines indicate 95% confidence intervals (CIs) of the estimate. The diamond data marker represents the overall estimator based on included studies.

(Supplemental Fig. S1A), and no significant publication bias was found by the Egger test ($p = 0.85$). The results were consistent across all the pre-defined subgroups (Supplemental Figs. S2–S4). The pooled prevalence among children with drug challenge test with or without preliminary test was 0.054 (95% CI, 0.040–0.070; $I^2 = 77\%$) from 28 studies and 0.043 (95% CI, 0.032–0.054; $I^2 = 70\%$) from 44 studies, respectively (Supplemental Fig. S2). The pooled prevalence among children with a history of immediate or nonimmediate reactions was 0.043 (95% CI, 0.014–0.081; $I^2 = 51\%$) from 12 studies and 0.062 (95% CI, 0.048–0.079; $I^2 = 73\%$) from 23 studies, respectively (Supplemental Fig. S3). Of the 12 studies on children with a history of immediate reactions, 8 had a preliminary test, 2 did not, and 2 were unclear. In contrast, on the 23 studies of children with a history of nonimmediate reactions, 10 had preliminary testing, 11 did not, and 2 were unclear. Nine studies assessed children with a history of nonimmediate reactions using a direct drug challenge. Of these, eight studies had been published in or after 2015. The results were consistent, with a pooled prevalence of 0.073 (95% CI, 0.047–0.104; $I^2 = 83\%$; Fig. 3).

Prevalence of serious adverse events. A total of 11,083 participants were included from 62 studies. The methodological quality of the 62 studies was moderate (Supplemental Table S3B). The reason for the moderate quality was inappropriate sampling processes. Of the 11,083 participants, only 13 experienced serious adverse events. All cases developed anaphylaxis and/or required adrenaline administration. The prevalence from the included studies was 0.000 (0.000–0.000). The pooled prevalence was 0.00 (95% CI, 0.00–0.00; $I^2 = 0.0\%$; Fig. 4). Of the nine studies that assessed children with a history of nonimmediate reactions using a direct drug challenge, seven reported prior serious adverse reactions. However, of these, including 2598 participants, none had serious adverse reactions to the drug challenge. Visual inspection of the funnel plot asymmetry indicated the presence of publication bias (Supplemental Fig. S1B), and this was confirmed by significant results on the Egger test ($p < 0.001$).

Prevalence of positive results in the second challenge after a negative result in the first challenge. A total of 1361 participants were included from 18 studies. The methodological quality of the 18 studies was low (Supplemental Table S3C). Three studies assessed the results of drug challenge tests, and the remaining 15 studies confirmed the outcomes of re-exposure to the suspected drug with questionnaires. The prevalence from the included studies was 0.033 (0.019–0.062). Overall, the pooled prevalence was 0.028 (95% CI, 0.016–0.043; $I^2 = 38\%$; Fig. 5). The pooled prevalence from the three studies with the drug challenge test was 0.017 (95% CI, 0.005–0.035; $I^2 = 0.0\%$). The pooled prevalence from the 15 studies that used questionnaires was 0.033 (95% CI, 0.018–0.052; $I^2 = 41\%$; Supplemental Fig. S5).

DISCUSSION

In this systematic review and meta-analysis, approximately 95% of children with a suspected beta-lactam allergy were not truly allergic to the drug based on the drug challenge test result. Few of the children who underwent the drug challenge test had serious adverse events. The false-negative rate for the drug challenge test was also very low.

The prevalence of beta-lactam allergy among children appears to be overestimated. The tolerance rate found was comparable to that found by a recent meta-analysis of adult populations.[26] Overlabeling of beta-lactam allergy can be attributed to the infection process caused by interaction between the virus and drug-induced immune activation in susceptible individuals.[27] A study screening for viral infections by polymerase chain reaction suggested that the viral infection was suspected to be the trigger for the initial rash in most children with negative drug challenge results.[7] Other possible causes include misdiagnosis of a side effect and the disappearance of a long-term IgE-mediated allergy caused by refraining from the culprit drug use.[28] Therefore, if there are no contraindications, a drug challenge should be routinely performed for children with suspected beta-lactam allergies. This is important for future treatments because inpatients suspected with a beta-lactam allergy are likely to receive broad-spectrum antibiotics for the treatment of acute illness to avoid recommended beta-lactams.[4]

The prevalence of positive direct drug challenge results for children with a history of mild nonimmediate reactions was about 7%, which is roughly the same as the overall percentage. Additionally, none had serious adverse reactions. Because beta-lactam allergy in children is more often a mild nonimmediate reaction,[29] establishing a testing protocol for re-diagnosing children with a history of mild nonimmediate reactions is an important challenge. Furthermore, skin tests are commonly conducted before a drug challenge.[30] However, since 2015, there has been an increasing number of reports suggesting that direct drug challenge tests can be safely performed without preliminary skin tests in children with nonimmediate and/or mild symptoms of beta-lactam allergy.[10–12,31] Additionally, there remain concerns regarding the low diagnostic accuracy of skin tests performed for children with suspected nonimmediate allergic reactions.[32,33] Recently published clinical guidelines[1,14] and position papers[15,16] have recommended direct challenge testing for children with a history of mild nonimmediate reactions, and the results of the present study support this strategy.

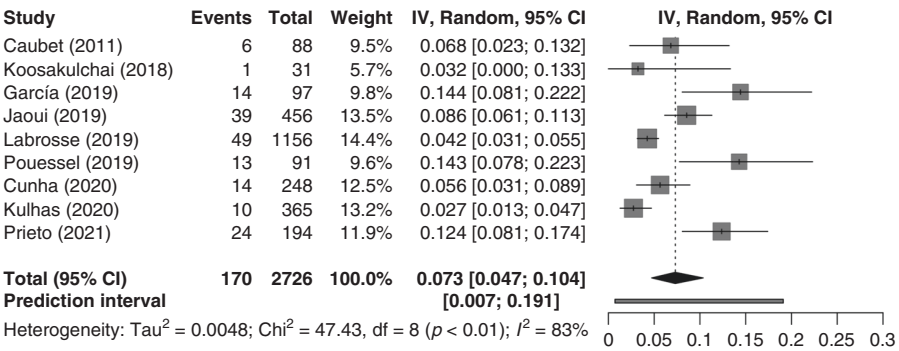

Fig. 3 Pooled prevalence of positive results for the first challenge (direct challenge for children with a history of nonimmediate reaction). An inverse-variance-weighted random-effect meta-analysis was used. The size of the point estimate (square data markers) is proportional to the weight of the study. Horizontal lines indicate 95% confidence intervals (CIs) of the estimate. The diamond data marker represents the overall estimator based on included studies.

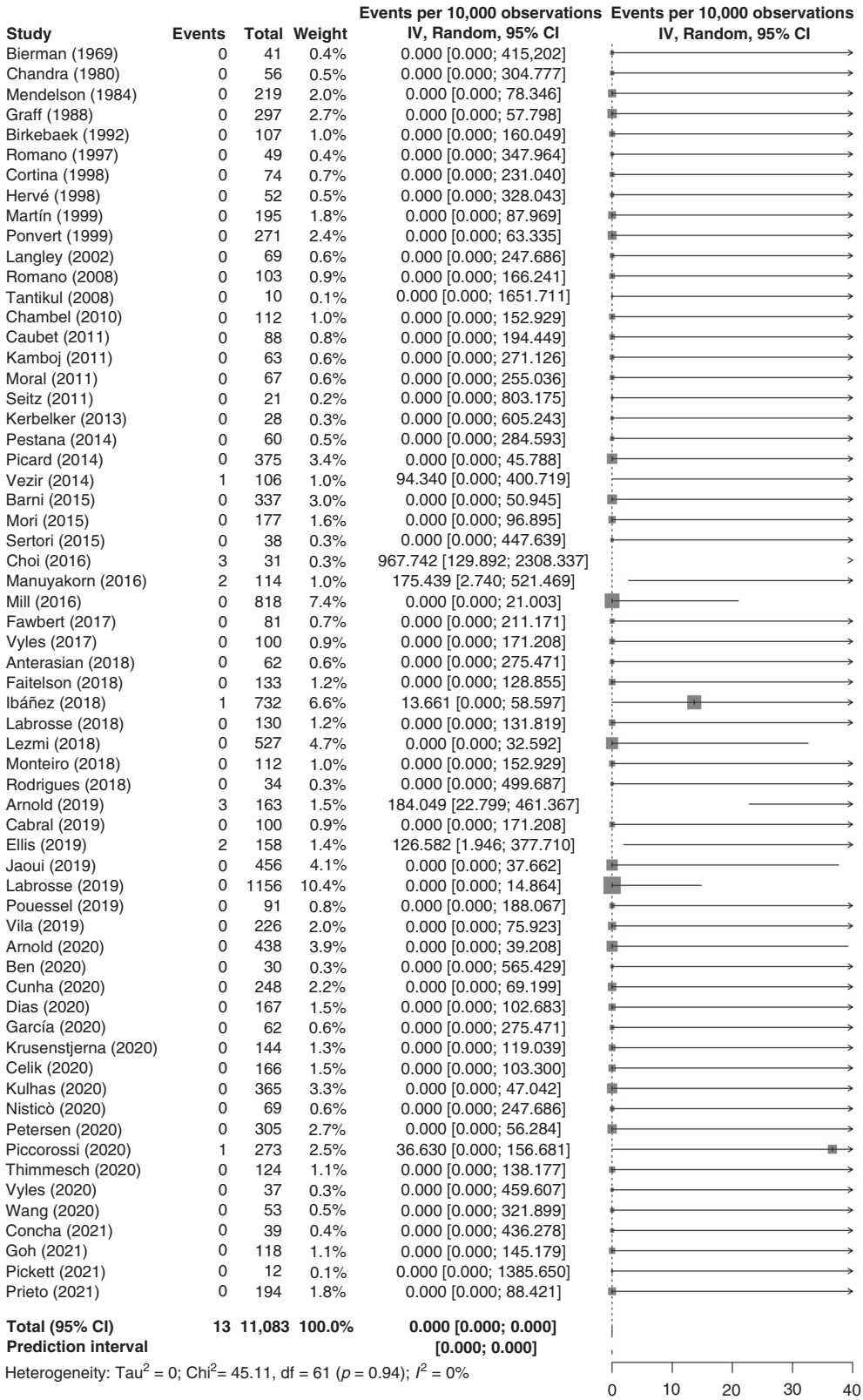

Fig. 4 Pooled prevalence of serious adverse events. An inverse-variance-weighted random-effect meta-analysis was used. The size of the point estimate (square data markers) is proportional to the weight of the study. Horizontal lines indicate 95% confidence intervals (CIs) of the estimate. The diamond data marker represents the overall estimator based on included studies.

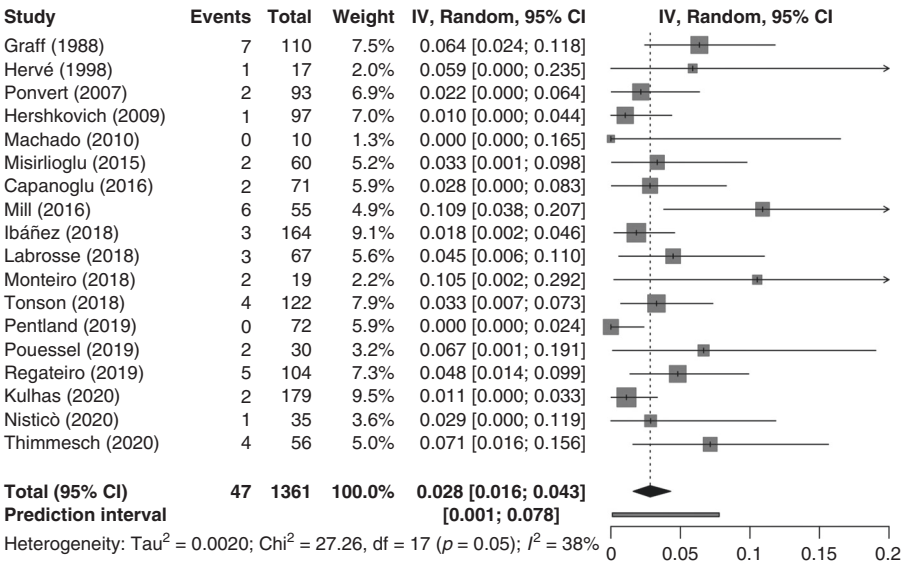

Fig. 5 Pooled prevalence of positive results for the second challenge after a first negative result. An inverse-variance-weighted random-effect meta-analysis was used. The size of the point estimate (square data markers) is proportional to the weight of the study. Horizontal lines indicate 95% confidence intervals (CIs) of the estimate. The diamond data marker represents the overall estimator based on included studies.

In our subgroup analyses, caution should be exercised in comparing the results of studies on children with a history of immediate reactions with those on children with a history of nonimmediate reactions. In the present study, both had comparable percentages of positive results, but a higher percentage of studies for children with a history of immediate reactions was conducted for those with a negative preliminary test. There were two studies in which direct challenge tests were performed on children with a history of immediate reactions.[8,34] In both studies, those with a history of severe immediate reactions were excluded, and the positive challenge test rates were reported to be 0% (0/38)[8] and 7.2% (5/69),[34] with no children receiving adrenaline or steroids. For children with a history of immediate reactions, a drug challenge test should be performed cautiously, with a thorough evaluation of the severity of past allergic reactions and additional preliminary testing if necessary.

The present study indicated that the drug challenge test seems to be a very safe procedure in children, with serious adverse events being rare. The incidence of anaphylaxis, fatal or otherwise, in patients receiving beta-lactams for the treatment of infections is also rare, with estimated rates of 0.015–0.004% and 0.0015–0.002%, respectively.[1,35] Clinicians should not refrain from the drug challenge test for fear of serious adverse events. Of necessity, challenge testing should be avoided in children with a history of delayed reaction with organ dysfunction or delayed severe cutaneous reaction including skin desquamation, purpura, mucosal lesions, drug reactions with eosinophilia and systemic symptoms, Stevens–Johnson syndrome/toxic epidermal necrolysis, or acute generalized exanthematous pustulosis.[14]

Based on the low false-negative rate, repeat testing for confirmation may not be required following a negative drug challenge result.[36] Because many of the original studies included in this meta-analysis used questionnaires to determine whether allergy-like symptoms occurred in children who were re-administered a culprit drug, the true false-negative rate may be even lower than the point estimate obtained in this meta-analysis. A previous study reported that many children with a negative result from a first drug challenge test who developed symptoms on re-administration tested negative on a second drug challenge.[37] If the drug challenge test is negative, IgE-mediated allergy can be largely ruled out. However, the children will experience other benign skin rashes with the same incidence as that of the general population.[38] Even if mild skin symptoms appear when re-administering a culprit drug after confirming a negative result in the drug challenge test, it does not necessarily denote an allergic reaction, and the drug may not need to be uniformly restricted.

To our knowledge, this is the most recent and comprehensive systematic review and meta-analysis to examine the true-positive rate in children with suspected beta-lactam allergy based on the results of drug challenge tests. We included 20 times more original studies than the previous meta-analysis.[3] The large number of original studies allowed investigation of subgroup differences and reporting biases, increasing the robustness of our results.

There were some limitations to this study. First, the background of participants who underwent the drug challenge varied between studies. Although our subgroup analyses showed that true-positive rates for almost all subgroups were comparable, the certainty that children truly have the allergy may also vary. Second, the drug challenge testing protocols followed in studies differed in dose step, dose interval, and test length. The optimal drug challenge protocol for children with suspected beta-lactam allergy is still controversial.[39,40] Third, the research settings differed between studies: they included an emergency department,[41] outpatient pediatric allergy clinics,[42,43] and a drug allergy center.[44] These differences in challenge testing protocols and research settings might have affected the results. Fourth, there were no studies examining false-positives resulting from the drug challenge test. However, if there is a history of suspected beta-lactam allergy and the drug challenge test is positive, false-positives are less likely to occur and will have less impact on the results.

In conclusion, the prevalence of children with suspected beta-lactam allergy who are truly positive on the drug challenge test was very low. Serious adverse events in response to the drug challenge were rare. The false-negative rate for drug challenge tests was also very low. Physicians should consider performing a drug challenge test after or without skin tests according to the severity of past allergic reactions, if not contraindicated, such as a suspected history of severe allergic reactions, to prevent the misdiagnosis of beta-lactam allergy.

DATA AVAILABILITY

All data generated or analyzed during this study are included in this published article and its supplementary information files.

REFERENCES

1. Mirakian, R. et al. Management of allergy to penicillins and other beta-lactams. *Clin. Exp. Allergy* **45**, 300–327 (2015).
2. Marrs, T., Fox, A. T., Lack, G. & du Toit, G. The diagnosis and management of antibiotic allergy in children: systematic review to inform a contemporary approach. *Arch. Dis. Child.* **100**, 583–588 (2015).
3. 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.* **128**, 557–562 (2016).
4. Mancini, C. M. et al. Association of penicillin or cephalosporin allergy documentation and antibiotic use in hospitalized patients with pneumonia. *J. Allergy Clin. Immunol. Pract.* **9**, 3060–3068.e1 (2021).
5. Norton, A. E., Konvinse, K., Phillips, E. J. & Broyles, A. D. Antibiotic allergy in pediatrics. *Pediatrics* **141**, e20172497 (2018).
6. Sousa-Pinto, B. et al. Accuracy of penicillin allergy diagnostic tests: a systematic review and meta-analysis. *J. Allergy Clin. Immunol.* **147**, 296–308 (2021).
7. Caubet, J. C. et al. The role of penicillin in benign skin rashes in childhood: a prospective study based on drug rechallenge. *J. Allergy Clin. Immunol.* **127**, 218–222 (2011).
8. Ibáñez, M. D. et al. Prospective assessment of diagnostic tests for pediatric penicillin allergy: from clinical history to challenge tests. *Ann. Allergy Asthma Immunol.* **121**, 235–244.e3 (2018).
9. Piccorossi, A. et al. Epidemiology and drug allergy results in children investigated in allergy unit of a tertiary-care paediatric hospital setting. *Ital. J. Pediatr.* **46**, 5 (2020).
10. Mill, C. 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.* **170**, 1–8 (2016).
11. Confino-Cohen, R. et al. Oral challenge without skin testing safely excludes clinically significant delayed-onset penicillin hypersensitivity. *J. Allergy Clin. Immunol. Pract.* **5**, 669–675 (2017).
12. Prieto, A. et al. Single dose prolonged drug provocation test, without previous skin testing, is safe for diagnosing children with mild non-immediate reactions to beta-lactams. *Allergy* **76**, 2544–2554 (2021).
13. Page, M. J. et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* **372**, n71 (2021).
14. Jeimy, S. et al. Practical guide for evaluation and management of beta-lactam allergy: position statement from the Canadian Society of Allergy and Clinical Immunology. *Allergy Asthma Clin. Immunol.* **16**, 95 (2020).
15. Gomes, E. R. et al. Drug hypersensitivity in children: report from the pediatric task force of the EAACI Drug Allergy Interest Group. *Allergy* **71**, 149–161 (2016).
16. Romano, A. et al. Towards a more precise diagnosis of hypersensitivity to beta-lactams—an EAACI position paper. *Allergy* **75**, 1300–1315 (2020).
17. Ouzzani, M., Hammady, H., Fedorowicz, Z. & Elmagarmid, A. Rayyan—a web and mobile app for systematic reviews. *Syst. Rev.* **5**, 210 (2016).
18. Munn, Z., Moola, S., Lisy, K., Riitano, D. & Tufanaru, C. Methodological guidance for systematic reviews of observational epidemiological studies reporting prevalence and cumulative incidence data. *Int. J. Evid. Based Healthc.* **13**, 147–153 (2015).
19. Li, C. et al. Global prevalence and incidence estimates of oral Lichen Planus: a systematic review and meta-analysis. *JAMA Dermatol* **156**, 172–181 (2020).
20. De Sola, H., Dueñas, M., Salazar, A., Ortega-Jiménez, P. & Failde, I. Prevalence of therapeutic use of opioids in chronic non-cancer pain patients and associated factors: a systematic review and meta-analysis. *Front. Pharmacol.* **11**, 564412 (2020).
21. R Core Team. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2021 URL https://www.R-project.org/.
22. IntHout, J., Ioannidis, J. P. A., Rovers, M. M. & Goeman, J. J. Plea for routinely presenting prediction intervals in meta-analysis. *BMJ Open* **6**, e010247 (2016).
23. Hunter, J. P. et al. In meta-analyses of proportion studies, funnel plots were found to be an inaccurate method of assessing publication bias. *J. Clin. Epidemiol.* **67**, 897–903 (2014).
24. Simmonds, M. Quantifying the risk of error when interpreting funnel plots. *Syst. Rev.* **4**, 24 (2015).
25. Sterne, J. A. C., Gavaghan, D. & Egger, M. Publication and related bias in meta-analysis: power of statistical tests and prevalence in the literature. *J. Clin. Epidemiol.* **53**, 1119–1129 (2000).
26. DesBiens, M. et al. A closer look at penicillin allergy history: systematic review and meta-analysis of tolerance to drug challenge. *Am. J. Med.* **133**, 452–462.e4 (2020).
27. Mori, F. et al. Amoxicillin allergy in children: five-day drug provocation test in the diagnosis of nonimmediate reactions. *J. Allergy Clin. Immunol. Pract.* **3**, 375–380.e1 (2015).
28. Lang, D. M., Castells, M. C., Khan, D. A., Macy, E. M. & Murphy, A. W. Penicillin allergy testing should be performed routinely in patients with self-reported penicillin allergy. *J. Allergy Clin. Immunol. Pract.* **5**, 333–334 (2017).
29. Zambonino, M. A. et al. Diagnostic evaluation of hypersensitivity reactions to beta-lactam antibiotics in a large population of children. *Pediatr. Allergy Immunol.* **25**, 80–87 (2014).
30. Joint Task Force on Practice Parameters, American Academy of Allergy, Asthma and Immunology, American College of Allergy, Asthma and Immunology, Joint Council of Allergy, Asthma and Immunology. Drug allergy: an updated practice parameter. *Ann. Allergy Asthma Immunol.* **105**, 259–273 (2010).
31. Wang, L. A., Patel, K., Kuruvilla, M. E. & Shih, J. Direct amoxicillin challenge without preliminary skin testing for pediatric patients with penicillin allergy labels. *Ann. Allergy Asthma Immunol.* **125**, 226–228 (2020).
32. Ponvert, C. 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.* **22**, 411–418 (2011).
33. Barni, S. et al. Utility of skin testing in children with a history of non-immediate reactions to amoxicillin. *Clin. Exp. Allergy* **45**, 1472–1474 (2015).
34. Nisticò, D. et al. Direct drug provocation test for the diagnosis of self-reported, mild and immediate drug hypersensitivity reaction in children and adolescents: our real-life experience. *Minerva Pediatr.* **73**, 209–214 (2021).
35. Pouessel, G., Winter, N., Lejeune, S., Thumerelle, C. & Deschildre, A. Oral challenge without skin testing in children with suspected non-severe betalactam hypersensitivity. *Pediatr. Allergy Immunol.* **30**, 488–490 (2019).
36. Capanoglu, M. et al. Additional provocation testing in patients with negative provocation test results with β-lactam antibiotics. *Ann. Allergy Asthma Immunol.* **116**, 82–83 (2016).
37. Ponvert, C. et al. Allergy to betalactam antibiotics in children: a prospective follow-up study in retreated children after negative responses in skin and challenge tests. *Allergy* **62**, 42–46 (2007).
38. Shenoy, E. S., Macy, E., Rowe, T. & Blumenthal, K. G. Evaluation and management of penicillin allergy: a review. *JAMA* **321**, 188–199 (2019).
39. Torres, M. J. et al. Controversies in drug allergy: beta-lactam hypersensitivity testing. *J. Allergy Clin. Immunol. Pract.* **7**, 40–45 (2019).
40. García Rodríguez, R. et al. Provocation tests in nonimmediate hypersensitivity reactions to β-lactam antibiotics in children: are extended challenges needed? *J. Allergy Clin. Immunol. Pract.* **7**, 265–269 (2019).
41. Vyles, D. et al. Allergy testing in children with low-risk penicillin allergy symptoms. *Pediatrics* **140**, S224.1–S224 (2017).
42. Lyter-Reed, L., Gupta, M. & Anagnostou, A. Results from 46 clinic-based antibiotic oral challenges in a pediatric population. *J. Allergy Clin. Immunol.* **145**, AB97 (2020).
43. Anterasian, C. M. & Geng, B. Penicillin skin testing in the management of penicillin allergy in an outpatient pediatric population. *Allergy Asthma Proc.* **39**, 305–310 (2018).
44. Azevedo, J. et al. Anaphylaxis to beta-lactam antibiotics at pediatric age: six-year survey. *Allergol. Immunopathol.* **47**, 128–132 (2019).

ACKNOWLEDGEMENTS

We are very grateful to Dr. Stefano Passanisi, Gaetano Barresi Department of Human Pathology in Adult and Developmental Age, University of Messina; Dr. Arnon Broides, Faculty of Health Sciences, Ben-Gurion University of the Negev; and Dr. Niels H. Birkebæk, Department of Pediatrics, Aarhus University Hospital for providing additional data. We are also very grateful to Dr. Yuki Kataoka, Department of Internal Medicine, Kyoto Min-Iren Asukai Hospital, and Dr. Dawid Pieper, Institute for Research in Operative Medicine (IFOM)—Department for Evidence-based Health Services Research, University Witten/Herdecke for providing support to our research. The authors would like to thank Enago (www.enago.jp) for their English language review.

AUTHOR CONTRIBUTIONS

All authors designed the study and wrote the manuscript. Y.K., N.T., and H.T. contributed to data collection. Y.K. performed the statistical analysis and interpretation of the results. All authors read and approved the final manuscript.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41390-022-02076-x.

Correspondence and requests for materials should be addressed to Yasutaka Kuniyoshi.