



Review

Efficacy and Safety of a Single Dose Versus Three Weekly Doses of Benzathine Penicillin G for Early Syphilis: A Systematic Review and Meta-Analysis

Thanyarat Phumthian ¹, Sudapree Sorasuchart ²  and Samadhi Patamatamkul ^{3,*} 

¹ College of Medicine and Public Health, Ubon Ratchathani University, Ubon Ratchathani 34190, Thailand; thanyarat.p@ubu.ac.th

² Department of Anesthesiology, King Chulalongkorn Memorial Hospital, The Thai Red Cross Society, Bangkok 10330, Thailand; sudapree.s@chula.ac.th

³ Department of Medicine, Suddhavej Hospital, Faculty of Medicine, Mahasarakham University, Maha Sarakham 44000, Thailand

* Correspondence: samadhi.p@msu.ac.th; Tel.: +66-815441374

Abstract

Background/Objectives: To compare the efficacy and safety of a single dose versus three weekly doses of benzathine penicillin G (BPG) for the treatment of early syphilis. **Methods:** We searched Embase, PubMed, and Scopus for randomized controlled trials (RCTs) and prospective comparative studies published in English up to September 2025. The primary outcome was serological cure (≥ 4 -fold decline in non-treponemal titers) at 6–12 months, analyzed on an intention-to-treat (ITT) basis. Risk of bias was assessed using RoB 2 for RCTs and ROBINS-I for non-randomized studies; certainty of evidence was rated with GRADE. Data were synthesized using random-effects meta-analysis and trial sequential analysis (TSA). **Results:** Three studies ($n = 886$) met the inclusion criteria. The primary ITT meta-analysis showed no significant difference in serological cure between the three-dose and single-dose regimens (risk ratio [RR] 1.06; 95% CI 0.92–1.21; $p = 0.42$), with substantial heterogeneity. In an exploratory pre-specified subgroup of patients with high baseline RPR ($\geq 1:32$), the three-dose regimen was associated with higher cure rates (RR 1.12; 95% CI 1.02–1.23; $p = 0.02$; $I^2 = 0\%$), but no significant benefit was seen for people with HIV (PWH) (RR 1.08; $p = 0.13$) or for those with CD4 counts < 350 cells/mm³. Risk of bias was serious across all studies, and GRADE certainty was very low for efficacy and low for safety. In a post hoc per-protocol analysis restricted to early latent syphilis, the three-dose regimen yielded higher serological cure (pooled risk difference 12%, 95% CI 1–23; $p = 0.03$). TSA indicated that the evidence is inconclusive and underpowered. **Conclusions:** On the basis of the ITT analysis, a single dose of BPG appears to remain effective and safe for most cases of early syphilis, including in PWH, supporting current CDC and WHO recommendations. The apparent advantages of three doses in high-titer and early latent subgroups derive from exploratory and post hoc analyses of studies at high risk of bias with very low certainty of evidence and should be regarded as hypothesis-generating rather than practice-changing. Adequately powered, well-designed trials are needed before any dose stratification can be recommended.



Academic Editor: Haim Levy

Received: 14 May 2026

Revised: 13 June 2026

Accepted: 24 June 2026

Published: 26 June 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

Keywords: syphilis; benzathine penicillin G; meta-analysis; HIV; antibiotic dosing

1. Introduction

Syphilis, caused by *Treponema pallidum* subspecies *pallidum*, remains a significant global public health concern with rising incidence rates, particularly among men who have sex

with men (MSM) and people with HIV (PWH). Early syphilis (primary, secondary, and early latent stages) is highly infectious and is the primary target for interrupting transmission. Benzathine penicillin G (BPG) is the cornerstone of therapy. Guidelines from the U.S. Centers for Disease Control and Prevention (CDC) and the World Health Organization (WHO) recommend a single intramuscular dose of 2.4 million units (MU) for all stages of early syphilis [1,2]. However, clinical equipoise persists regarding the optimal dosing schedule. Observational data contradicts; one study found significantly lower serological response in PWH using single-dose BPG [3] whereas the other study did not [4]. Therefore, skepticism remains regarding whether a single dose maintains treponemicidal levels long enough to eradicate persistent organisms in high-risk groups, such as PWH. Consequently, some clinicians advocate for three weekly doses, the regimen for late latent syphilis, to prevent serological failure [5].

This systematic review and meta-analysis aims to synthesize current evidence comparing single versus three weekly doses of BPG for early syphilis, utilizing rigorous risk of bias assessment and trial sequential analysis (TSA) to determine if intensified treatment is warranted.

2. Materials and Methods

2.1. Protocol and Registration

This systematic review was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement [6]. The study protocol was registered prospectively with the International Prospective Register of Systematic Reviews (PROSPERO) on 16 September 2025 (Registration number: CRD420251148355).

2.2. Data Sources and Search Strategy

We systematically searched the following electronic databases up to 12 September 2025: Embase, PubMed, and Scopus. The search was conducted without restrictions on publication date (for Embase since 1996) but was limited to studies published in the English language. The search strategy utilized a combination of medical subject headings (MeSH) and free-text keywords related to “Syphilis,” “Benzathine Penicillin G,” “Early Syphilis,” and relevant study designs (Table S1). To ensure a comprehensive review, we also searched clinical trial registries for ongoing or unpublished trials and contacted experts and study authors to identify additional relevant data where necessary.

2.3. Eligibility Criteria

We included both randomized controlled trials (RCTs) and prospective non-randomized comparative studies that enrolled adults (≥ 18 years) with a confirmed diagnosis of early syphilis, encompassing primary, secondary, and early latent stages. Early latent syphilis was strictly defined as an infection acquired within the preceding 12 months, supported by patient history or serological evidence (e.g., a documented non-reactive non-treponemal test within the prior year). Eligible studies compared an intervention group receiving three weekly intramuscular doses of benzathine penicillin G (BPG) (2.4 MU on days 1, 8, and 15) against a control group receiving a single intramuscular dose of 2.4 MU BPG. The primary outcome was serological cure, defined as a four-fold (or two-dilution) decline in non-treponemal antibody titers (e.g., RPR or VDRL) assessed at 6 and/or 12 months post-treatment. Intention-to-treat (ITT) analysis population from the included studies was used to analyze the primary outcome. Secondary outcomes included the incidence of progression to neurosyphilis and treatment-related adverse events, such as the Jarisch–Herxheimer reaction.

We excluded studies involving patients with neurosyphilis, ocular syphilis, or otosyphilis; late latent or tertiary syphilis; pregnant women; and patients with a known penicillin allergy. We planned a priori subgroup analyses according to the stage of early syphilis (primary, secondary, and early latent) to explore potential effect modification. Additional subgroup analyses were planned according to RPR titer ($\geq 1:32$), history of prior syphilis infection, people with HIV (PWH), and, among PWH, antiretroviral therapy (ART) status and CD4 count < 350 cells/mm³.

Definitions of reinfection and treatment failure differed across studies and were extracted verbatim. Hook et al. [7] defined treatment failure as a ≥ 4 -fold RPR increase, with presumed re-exposure adjudicated by a blinded committee and excluded, post-retreatment data omitted from all analyses, and serologic response counted whenever achieved before any exclusion event. Andrade et al. [8] defined treatment success as a ≥ 4 -fold RPR decline within 12 months and did not separately define, adjudicate, or censor reinfection; because the success endpoint was uni-directional, a later ≥ 4 -fold increase was not counted as treatment failure. Yang et al. [9] used a composite treatment-failure endpoint (failure to achieve a ≥ 4 -fold decline at 12 months, a ≥ 4 -fold rise after an initial response, or any retreatment) and recorded but did not separate clinically documented reinfections (new primary or secondary syphilis).

Post hoc sensitivity analysis with per-protocol population was conducted for early latent syphilis given theoretical benefit use of three-dose BPG may be beneficial. A formal pooled sensitivity analysis subtracting probable reinfections was not feasible, because the studies handle reinfection asymmetrically: Hook et al. [7] removed it by adjudication and Andrade et al. [8] by a uni-directional success endpoint that never counts a later ≥ 4 -fold rise as failure (neither leaving a separable per-arm count); whereas, only Yang et al. [9] retained reinfection within its failure endpoint.

2.4. Study Selection and Data Extraction

Two reviewers (S.P. and T.P.) independently screened titles and abstracts against the eligibility criteria. Full-text articles of potentially relevant studies were subsequently retrieved and assessed for inclusion. Any discrepancies regarding study selection were resolved through discussion and arbitration by a third reviewer (S.S.).

Data extraction was performed independently by two reviewers (S.P. and T.P.) using a standardized data extraction form. Collected data included: study characteristics (author, year, country, design), participant demographics (age, sex, HIV status), baseline serological titers, intervention details, follow-up duration, and outcomes of interest. Missing data were requested from study authors when possible.

2.5. Risk of Bias Assessment

Two reviewers (S.P. and T.P.) independently assessed the risk of bias for each included study. For randomized controlled trials, we used the revised Cochrane risk-of-bias tool for randomized trials (RoB 2, version 2019), evaluating bias across five domains: randomization process, deviations from intended interventions, missing outcome data, measurement of the outcome, and selection of the reported result. For prospective non-randomized studies, we utilized the risk of bias in non-randomized studies—of interventions, version 2 (ROBINS-I V2, November 2024) assessment tool for follow-up studies. This tool evaluates bias across seven domains, including confounding, selection of participants, classification of interventions, deviations from intended interventions, missing data, measurement of outcomes, and selection of reported results.

We examined publication bias visually with a funnel plot. Because only a few studies were available ($k < 10$), we did not apply formal asymmetry tests such as Egger's, in

keeping with Cochrane guidance that such tests are underpowered when the number of studies is small.

2.6. Data Synthesis, Analysis and Statistical Analysis

Covidence was used for literature searching, reference importation, title and abstract screening, full-text review, and data extraction. Data synthesis was performed using R software (version 4.4.2) for Mac (The R Foundation for Statistical Computing), utilizing the meta and metafor packages. For dichotomous outcomes, study-level event counts were extracted and converted into effect estimates for meta-analysis. Subgroup-specific data were extracted separately when reported. In the post hoc per-protocol sensitivity analysis for early latent syphilis, results were additionally presented as risk differences for clinical interpretability. Missing or unclear data were sought from study authors when possible. No major data transformation was required; effect estimates were calculated directly from reported event counts.

Findings are displayed as a PRISMA flow diagram, descriptive tables of study features and adverse events, forest plots for the overall and subgroup analyses, supplementary risk-of-bias tables, a funnel plot, and trial sequential analysis figures.

For the primary outcome, pooled risk ratios (RRs) with 95% confidence intervals (CIs) were calculated using an inverse-variance random-effects meta-analysis. A random-effects model was selected a priori because of the expected clinical and methodological heterogeneity across the included studies. Between-study variance (τ^2) was estimated using the restricted maximum-likelihood (REML) method. Statistical heterogeneity was assessed using Cochran's Q test and quantified with the I^2 statistic; τ^2 and τ were also estimated, with 95% CIs for τ^2 and τ calculated using the Q-profile method. An I^2 value greater than 50% was considered indicative of substantial heterogeneity.

We graded the certainty of evidence for each outcome using GRADE, weighing risk of bias, inconsistency, indirectness, imprecision, and publication bias, and assigning a final rating of high, moderate, low, or very low.

All analyses were performed in R, and two-sided p values < 0.05 were considered statistically significant.

2.7. Trial Sequential Analysis (TSA)

To gauge the influence of random error (type I and type II) arising from limited data and repeated significance testing of accumulating evidence, we ran a trial sequential analysis in TSA software v0.9.5.10 Beta (Copenhagen Trial Unit). We derived the required information size (RIS) for a plausible treatment effect and adjusted the confidence limits and efficacy thresholds accordingly. Plotting the cumulative Z-curve against the O'Brien–Fleming monitoring boundaries indicated whether the accumulated evidence was conclusive or whether further trials would be futile.

3. Results and Discussions

3.1. Study Selection and Characteristics

The literature search, conducted in accordance with PRISMA guidelines, is summarized in Figure 1. An initial search of Scopus, PubMed, and Embase yielded a total of 1235 records. Specifically, the search identified 581 records from Scopus, 351 from PubMed, and 303 from Embase. No additional records were identified through other sources or citation searching. After the removal of 302 duplicate records (301 identified by automation tools and one identified manually), 933 articles were screened based on their titles and abstracts. Following this initial screening, 923 records were excluded, and 10 full-text records were

sought for retrieval (Figure 1). All 10 records were successfully retrieved and assessed for eligibility. Search strategies were shown in the Supplementary Table S1.

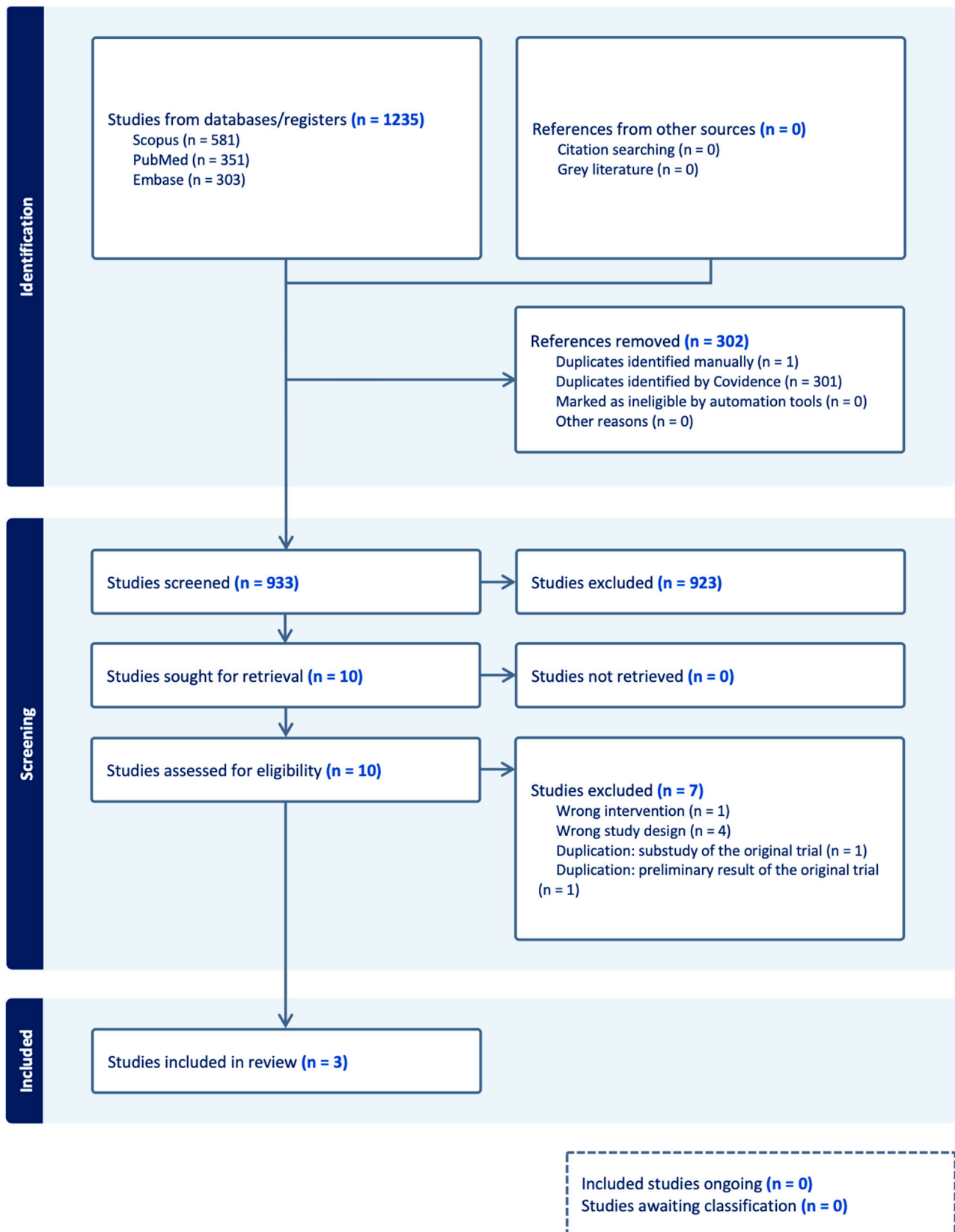


Figure 1. PRISMA flow diagram of study selection.

A total of seven studies were deemed ineligible and were excluded. The reasons for exclusion were wrong study design ($n = 4$), wrong intervention ($n = 1$), duplication as a substudy of the original trial ($n = 1$), and duplication as a preliminary result of the original trial ($n = 1$). Three studies were included in the final review (two randomized controlled trials and one prospective comparative study). The characteristics for each study are shown in Table 1.

Table 1. Characteristics of the included studies. Abbreviations: BPG, benzathine penicillin G; HIV, human immunodeficiency virus; RCT, randomized controlled trial; RPR, rapid plasma reagin.

Study	Year	Country	Design	No. of Patients Receiving Single-Dose BPG	No. of Patients Receiving Three-Dose BPG	Population	Primary Outcome	Follow-Up Schedule	Longest Follow-Up
Andrade et al. [8]	2017	USA	Prospective, randomized Open-label, parallel-group RCT	35	29	Adults aged ≥ 18 years with HIV Early syphilis (primary, secondary, or early latent)	≥ 2 -dilution (≥ 4 -fold) decline in RPR titer or seroreversion at 6 months (noninferiority margin 10 percentage points)	3, 6, 9, and 12 months	12 months
Hook et al. [7]	2025	USA	Multicenter, randomized, controlled Open-label, noninferiority RCT	124	125	Adults aged ≥ 18 years Early syphilis with or without HIV infection	≥ 2 -dilution (≥ 4 -fold) decline in RPR titer or seroreversion at 6 months (noninferiority margin 10 percentage points)	1 week, 2 weeks, 1 month, 3 months, 6 months, and 12 months	12 months
Yang et al. [9]	2014	Taiwan	Multicenter, prospective observational study	295	278	Adults aged ≥ 20 years with HIV Early syphilis	≥ 4 -fold decline in RPR titer at 12 months	Every 3 to 6 months	12 months

3.2. Quality Assessment of the Included Studies

Overall, the risk-of-bias assessment suggested that the internal validity of the included evidence was limited. Among the two randomized trials assessed using RoB 2, Hook et al. [7] was judged to be at overall high risk of bias, driven mainly by concerns in domain 2, as imbalance was concentrated in the three-dose BPG group: 17% (21/125) did not receive all three assigned doses, without appropriate handling through a prespecified adherence estimand or alternative methods such as a causal per-protocol analysis. Further concerns arose in domain 3 due to missing outcome data. Additionally, Andrade et al. [8] was judged as high risk of bias overall, with concerns in the randomization process and deviations from intended interventions, and a high-risk judgment for missing outcome data. In contrast, domains related to outcome measurement and selection of the reported result were generally low risk in both randomized studies. The non-randomized study by Yang et al. [9], was judged as having serious risk of bias overall, mainly due to confounding and selection of participants into the study, while classification of interventions and measurement of outcomes were considered low risk; the remaining domains showed moderate concerns. The details for each study's quality assessment were shown in the Supplementary Tables S2 and S3.

3.3. Primary Outcome: Serological Response to Treatment

A total of three studies (Andrade et al. [8], Yang et al. [9], and Hook et al. [7]) comprising 886 participants were included in the primary meta-analysis comparing the efficacy of three weekly doses of benzathine penicillin G (BPG) versus a single dose for early syphilis [7–9]. The analysis included 432 participants in the three-dose group and 454 participants in the single-dose group, with a total of 642 observed serological cure events. Variation was noted in the timing of the serological cure assessment: Andrade et al. [8] and

Yang et al. [9] measured outcomes at 12 months, compared to a 6-month follow-up in the study by Hook et al. [7].

The random-effects meta-analysis yielded a pooled risk ratio (RR) of 1.06 (95% CI: 0.92 to 1.21) ($p = 0.42$), indicating that there was no statistically significant difference in serological cure rates between the three-dose and single-dose regimens. Substantial statistical heterogeneity was observed across the included studies ($I^2 = 61.3\%$; $\tau^2 = 0.0090$). The test for heterogeneity approached statistical significance ($Q = 5.17$, $df = 2$, $p = 0.08$), suggesting variation in treatment effects among the studies (Figure 2).

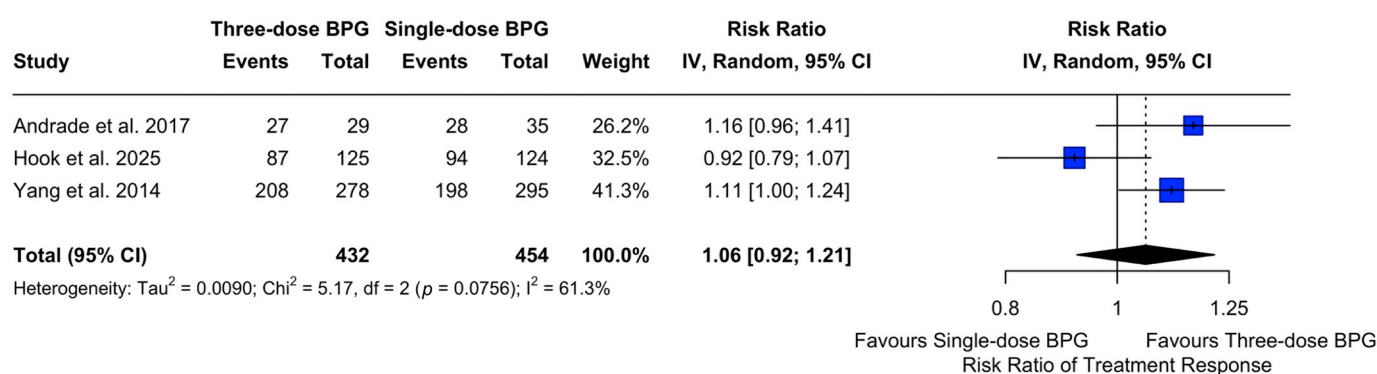


Figure 2. Forest plot of treatment response comparing three weekly doses versus a single dose of benzathine penicillin G (BPG) for early syphilis. Yang et al. [9], which carried the highest weight in the analysis (41.3%), favored the three-dose regimen with an RR of 1.11 (95% CI: 1.00 to 1.24). In contrast, Hook et al. [7], accounting for 32.5% of the weight, showed a trend toward the single-dose regimen (RR 0.92; 95% CI: 0.79 to 1.07), while Andrade et al. [8] showed no significant difference (RR 1.16; 95% CI: 0.96 to 1.41).

3.4. Subgroup Analyses

3.4.1. Phase of Syphilis

Stratification by disease stage ($n = 873$) revealed no significant subgroup differences ($p_{\text{interaction}} = 0.27$). Participants with primary syphilis conferred no increased risk in serological failure with single dose BPG vs. three-dose BPG: RR 0.91 (95% CI: 0.73 to 1.13; $I^2 = 0\%$). The results were similar across secondary syphilis and early latent syphilis: RR 1.11 (95% CI: 0.99 to 1.25; $I^2 = 7.7\%$), and RR 1.11 (95% CI: 0.84 to 1.47; $I^2 = 57.7\%$), respectively (Figure 3). p -values for these subgroups are not shown.

3.4.2. People with HIV (PWH)

In the subgroup of PWH ($n = 790$), the analysis revealed no significant difference in serological cure rates between the two regimens (RR 1.08; 95% CI: 0.98 to 1.20; $p = 0.13$). Heterogeneity in this subgroup was moderate ($I^2 = 32.6\%$) (Figure S1).

Further stratification within the PWH population showed consistent results with no significant differences detected with PWH and CD4 count < 350 cells/ mm^3 : RR 1.12 (95% CI: 0.96 to 1.30; $p = 0.16$; $I^2 = 0.0\%$) (Figure S2); while PWH on antiretroviral therapy (HAART): RR 1.07 (95% CI: 0.96 to 1.19; $p = 0.25$; $I^2 = 0.0\%$) (Figure S3).

3.4.3. Baseline RPR Titer $\geq 1:32$

A pre-specified subgroup analysis was conducted for participants with high baseline non-treponemal titers (RPR $\geq 1:32$), incorporating data from Andrade et al. [8] and Yang et al. [9]. In this high-load subgroup ($n = 533$), the three-dose regimen was associated with a statistically significant improvement in serological cure compared to the single-dose regimen (RR 1.12; 95% CI: 1.02 to 1.23; $p = 0.02$). Notably, heterogeneity was absent in this comparison ($I^2 = 0.0\%$). Subgroup definitions varied slightly, with Andrade et al. [8]

including participants with an RPR $\geq 1:32$, and Yang et al. [9] including those with an RPR $> 1:32$ (Figure S4). Because this analysis is based on only two studies at high or serious risk of bias, is subject to multiplicity, and is not corroborated by the overall ITT analysis, it should be interpreted as exploratory and hypothesis-generating.

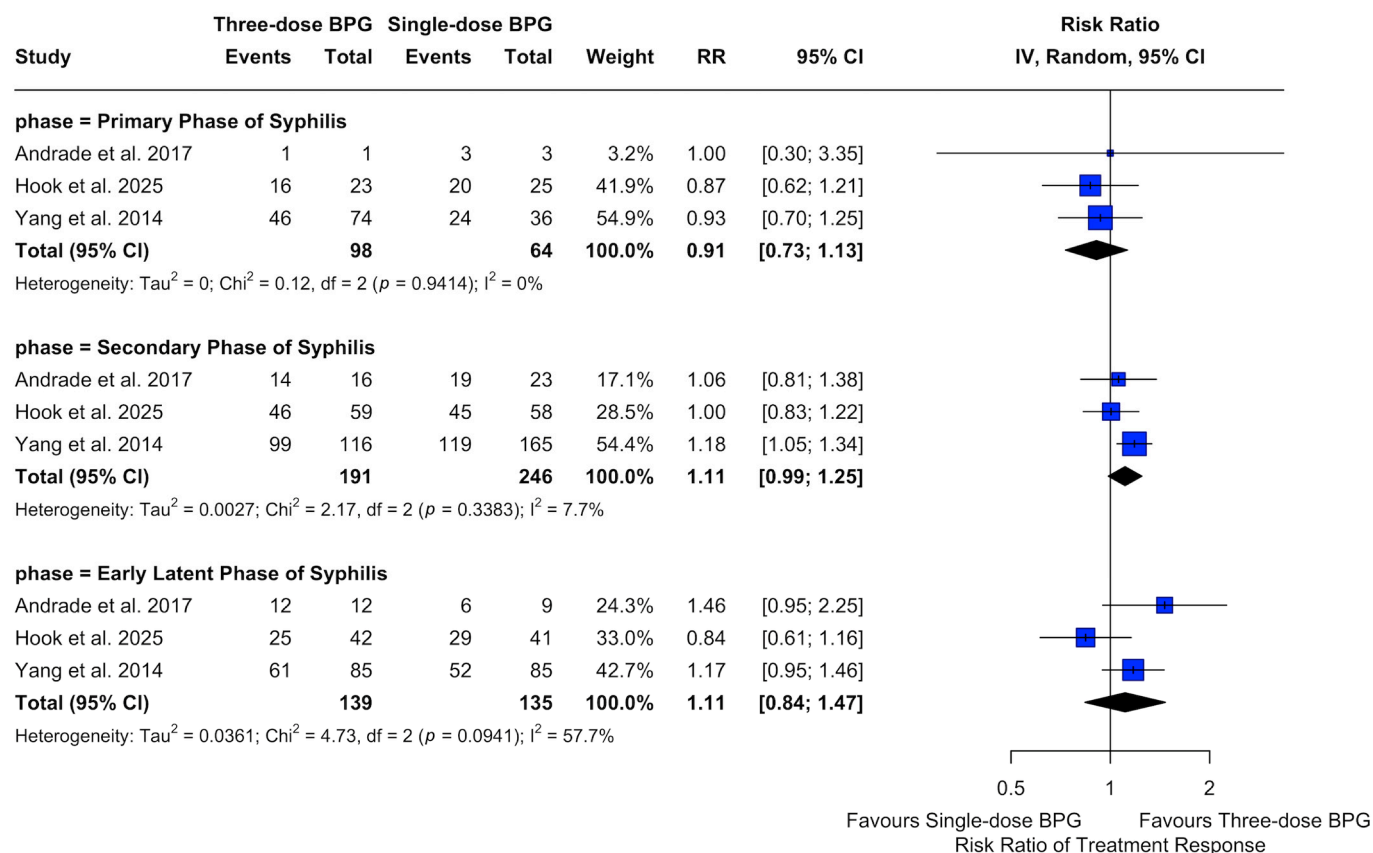


Figure 3. Forest plot of subgroup analyses by syphilis stage comparing three weekly doses versus a single dose of benzathine penicillin G (BPG). The included studies were Andrade et al. (2017) [8], Hook et al. (2025) [7], and Yang et al. (2014) [9].

3.4.4. History of Previous Syphilis

Among participants with a documented history of previous syphilis ($n = 225$), there was a trend favoring the three-dose regimen, but this did not reach statistical significance (RR 1.17; 95% CI: 0.98 to 1.40; $p = 0.08$). No heterogeneity was observed ($I^2 = 0.0\%$) (Figure S5).

3.5. Adverse Reactions and Secondary Outcomes of Progression to Neurosyphilis, Ootosyphilis, or Ocular Syphilis

Jarisch–Herxheimer reactions or adverse events were not available in the Yang et al. [9] study. In Hook et al. [7], one case of neurosyphilis was diagnosed in the single-dose BPG group, compared with none in the three-dose BPG group; however, the investigators judged that the event would have occurred regardless of the treatment regimen administered [7] (Table 2). No participants developed neurosyphilis, otosyphilis, or ocular syphilis during the follow-up period in the studies by Yang et al. [9] and Andrade et al. [8] (Table 2).

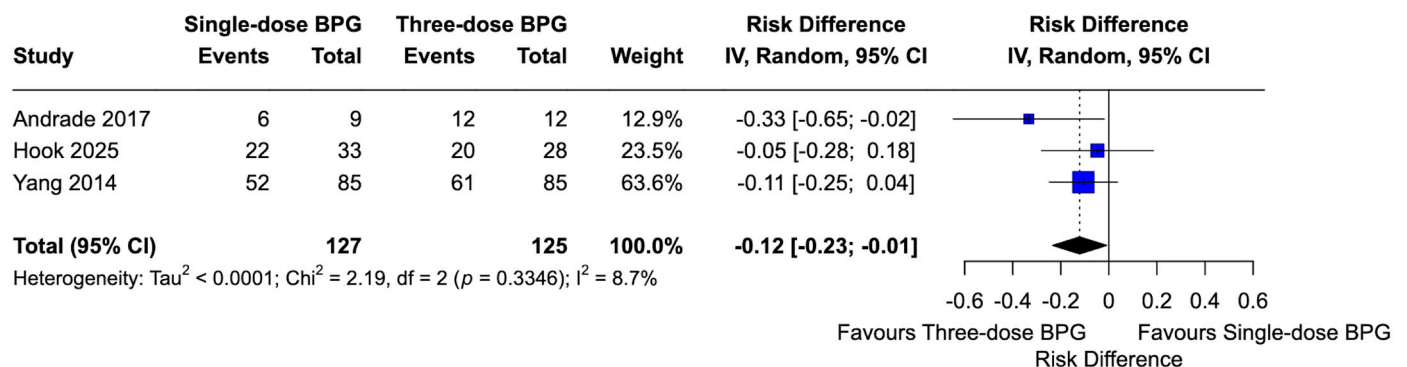
Table 2. Adverse events reported in the included studies. * There were no reports of local adverse reactions in Andrade et al. 2017 [8].

Study	Intervention	Events	Number of Subjects with Events	Incidence
Andrade et al., 2017 [8] *	single-dose BPG vs. three-dose BPG	Jarisch–Herxheimer reactions	0 vs. 0	0 vs. 0
		Death	0 vs. 0	0 vs. 0
Hook et al., 2025 [7]		Local injection site pain, redness, or tenderness	95 vs. 106	76.6% vs. 84.8%
		Related serious adverse events	0 vs. 0	0 vs. 0
		Jarisch–Herxheimer reactions	27 vs. 32	21.8% vs. 25.6%

3.6. Sensitivity Analysis of Per-Protocol Population for Early Syphilis

A total of three studies comprising 252 participants (127 in the single-dose BPG group and 125 in the three-dose BPG group) were included in this meta-analysis. This analysis included two randomized controlled trials (RCTs) utilizing per-protocol populations and one prospective observational study. Because Yang et al. [9] excluded participants with missing outcome data, their analyzed cohort closely approximates a per-protocol population. Results are presented as both risk ratios and risk differences to provide clearer clinical context, given that both included RCTs employed a 10% non-inferiority margin.

Using an inverse variance random-effects model, the overall pooled risk (proportion) difference was -12% (95% CI: -0.23 to -0.01) ($p = 0.034$) (Figure 4). This indicates a statistically significant decrease of 12 percentage points in the serological cure rates among patients receiving single-dose BPG compared to those receiving three-dose BPG. Heterogeneity across the included studies was found to be low and not statistically significant ($I^2 = 8.7\%$; $\tau^2 < 0.0001$; $\text{Chi}^2 = 2.19$, $p = 0.33$).

**Figure 4.** Forest plot of treatment response expressed as pooled proportion differences comparing three weekly doses versus a single dose of benzathine penicillin G (BPG) for early latent syphilis, restricted to the per-protocol analysis. The included studies were Andrade et al. (2017) [8], Hook et al. (2025) [7], and Yang et al. (2014) [9].

For the risk ratio analysis, the random-effects model yielded a pooled RR of 0.84 (95% CI: 0.714 to 0.997; $p = 0.046$), demonstrating a statistically significant advantage favoring the three-dose BPG regimen. There was no evidence of statistical heterogeneity across the included studies ($I^2 = 0.0\%$; $\tau^2 = 0.00$; Cochran's $Q = 1.26$, $p = 0.53$).

3.7. Publication Bias and GRADE

Visual inspection of the funnel plot for the primary outcome (serological response) was limited by the small number of included trials ($k = 3$) (Figure S6). While the plot showed studies distributed across both sides of the pooled effect, a definitive assessment of symmetry was not possible. A comprehensive manual search of abstracts and conference proceedings identified one potentially relevant study (abstract showed study methodology

and results similar to Yang et al. [9]); however, despite repeated inquiries, the authors did not respond, and sufficient data for inclusion could not be obtained. Therefore, the risk of publication bias cannot be entirely excluded. Using the GRADE approach, the certainty of evidence was rated very low for serological efficacy and low for safety (Table S4). The efficacy outcome was downgraded two levels for very serious risk of bias (both randomized trials at high risk and the non-randomized study at serious risk), one level for inconsistency (substantial unexplained heterogeneity, $I^2 = 61.3\%$), and one level for imprecision (few events and a pooled estimate compatible with both no effect and a modest benefit); safety outcomes were downgraded mainly for risk of bias and imprecision. Crude absolute risk differences yielded an approximate NNT of 454 for progression to symptomatic neurosyphilis, ocular syphilis, or otosyphilis, and NNH of 12 for local injection site pain, redness, or tenderness, and 26 for Jarisch–Herxheimer reaction with three-dose versus single-dose BPG. These estimates are approximate, non-pooled, and should be interpreted cautiously, particularly for progression outcomes because only one event occurred overall.

3.8. Trial Sequential Analysis

We performed a trial sequential analysis (TSA) using a random-effects model (Sidik-Jonkman estimator) to adjust for potential heterogeneity. Trial sequential analysis for the comparison of single-dose versus 3-weekly dose benzathine penicillin G (BPG) showed that the conventional random-effects meta-analysis was not statistically significant (RR 1.06, 95% CI 0.90–1.25; $p = 0.50$), with moderate-to-substantial heterogeneity (Cochran's $Q = 5.17$, $p = 0.08$; $I^2 = 61.3\%$; $D^2 = 77\%$). Using a two-sided O'Brien–Fleming alpha-spending approach with a type I error of 5%, power of 80%, and a variance-based heterogeneity correction, the required information size (RIS) was estimated at 2595 participants (Figure S7). The anticipated intervention effect was set to a serological-cure proportion of 78.8% in the three-dose arm versus 68.8% in the single-dose arm, corresponding to a relative risk of approximately 1.15. The cumulative Z-curve, based on the currently accrued sample size of 886 participants, remained within the trial sequential monitoring boundaries and did not cross either the conventional threshold for statistical significance or the TSA-adjusted boundary for benefit or harm, although it trended slightly toward favoring three-dose BPG. Thus only 34.1% (886/2595) of the required information size has been accrued; the evidence is underpowered and inconclusive and can neither confirm nor exclude a modest benefit of three doses, so no firm conclusion of benefit or futility is currently warranted.

3.9. Discussions

This systematic review challenges the utility of intensified BPG dosing for early syphilis, particularly among PWH, who represent 89% (790/886) of the meta-analysis population. Our primary finding found that three doses offer no overall benefit over a single dose, supporting current CDC and WHO guidelines [1,2]. However, significant heterogeneity was observed in the primary analysis, in particular, with Yang et al. [9] and Andrade et al. [8] weighing towards favoring three-dose BPG [8,9]. The possible beneficial effects of three doses observed in the earlier literature appears largely attributable to methodological biases. The study by Yang et al. [9] was subject to confounding by indication, where physicians likely assigned the three-dose regimen to higher-risk patients, and the analysis was complicated by definitions of failure that included reinfection. Similarly, the Andrade et al. [8] trial was biased by a “missing = failure” imputation strategy that disproportionately penalized the single-dose arm due to differential dropout rates. Nevertheless, the most recent and more methodologically robust trial [7] was still judged to have high risk of bias because of non-adherence in the three-dose BPG arm without proper statistical adherence estimand use and missing outcome data attributable to attrition bias,

with similar loss to follow-up in the single-dose and three-dose BPG groups (21/124 and 23/125, respectively). The trial demonstrated no benefit of the three-dose regimen, and the balanced attrition between groups was unlikely to have meaningfully influenced the conclusion that single-dose BPG was non-inferior. However, given the substantial imbalance in non-adherence to the three-dose BPG regimen, the findings may have been biased toward favoring the non-inferiority of single-dose BPG. In conclusion, as all included studies were judged to be at high risk of bias, the overall evidence remains inconclusive but signals no clear benefit of three-dose BPG over single-dose BPG. Because the ITT analysis is less susceptible to deviations from the intended intervention and to selective attrition, we treat it as the primary basis for inference, with per-protocol and post hoc analyses regarded as supportive and exploratory only.

Despite these limitations, our analysis identified a possible signal that baseline RPR titers may matter. In patients with RPR titers $\geq 1:32$, the three-dose regimen conferred a statistically significant benefit (RR 1.12), with zero heterogeneity ($I^2 = 0\%$); however, this estimate derives from only two studies at high risk of bias and must be considered exploratory. A retrospective observational study suggests that high baseline RPR titers, particularly with PWH, are associated with a poorer serological response to single-dose BPG [10]; whereas, a previous meta-analysis including both PWH and HIV-negative persons reported the opposite, with higher RPR titers associated with a greater likelihood of serological cure [11]. Given these conflicting data, our findings can at most generate the hypothesis that high baseline RPR titer, rather than HIV status alone, may identify patients who could benefit from intensified treatment. We found no significant benefit of three doses in PWH analyzed as a distinct group.

The post hoc, per-protocol sensitivity analysis restricted to early latent syphilis suggested a statistically significant advantage of three-dose over single-dose BPG (risk difference -0.12 ; 95% CI -0.23 to -0.01 ; $I^2 = 8.7\%$). Because this analysis was post hoc, confined to the per-protocol population, and not supported by the ITT analysis, it should be interpreted as hypothesis-generating rather than confirmatory. A biological rationale is nonetheless plausible: early latent syphilis is associated with a longer time to, and lower rates of, serological cure than primary or secondary disease [11–14], and three doses may be required to achieve sustained treponemicidal concentrations at sanctuary sites such as the central nervous system. Up to 30% of patients with latent syphilis after untreated primary or secondary syphilis develop CSF abnormalities, and confirmed asymptomatic neurosyphilis has been reported in approximately 5% of those with early latent syphilis [15]. Serum penicillin concentrations on day 3 after a single 2.4-MU dose of BPG range from 0.0042 to 0.0258 $\mu\text{g/mL}$ [16]; whereas, 0.32 $\mu\text{g/mL}$ is achieved on day 7 after the third dose [17]; using the quantified MIC of 0.0005 $\mu\text{g/mL}$ for *T. pallidum* [18] and assuming 1–5% CSF penetration [19], a single dose would yield estimated CSF concentrations generally below the MIC; whereas, three weekly doses would be expected to maintain CSF concentrations consistently above it.

The previously proposed conservative treponemicidal threshold of 0.03 IU/mL (0.018 $\mu\text{g/mL}$) [20] may have been overstated; the original in vivo rabbit study estimated a treponemicidal penicillin concentration of 0.005–0.01 $\mu\text{g/mL}$ [21], a threshold that the three-dose schedule would still exceed in inflamed meninges. Consistent with partial efficacy of a single dose, treatment of asymptomatic neurosyphilis with a single 2.4- or 2.5-MU dose of BPG was associated with a cumulative CSF relapse rate of about 20% at 18 months, implying that nearly 80% of patients were cured even with one dose [22]. Taken together, these data raise the possibility that a small proportion of patients with latent syphilis—particularly PWH—harbor unrecognized asymptomatic neurosyphilis in whom

three doses may improve serological cure, an effect that a trial enrolling a heterogeneous population would be unlikely to detect.

These mechanistic considerations remain speculative. Critically, the results of this post hoc analysis are influenced by lower-quality studies and multiple testing, and the primary ITT analysis demonstrated no significant benefit. Because the ITT result is more robust to non-adherence and missing data, it takes precedence; the per-protocol/post hoc signal therefore provides, at best, supportive and exploratory evidence and cannot by itself justify a change in practice.

Another possible explanation for the observed beneficial effect of three-dose BPG for early latent syphilis is the high proportion PWH among the early latent syphilis cases. Although the exact number of PWH specifically within the early latent subgroups is not detailed in the included studies, PWH typically exhibits a delayed serological response to treatment [12]. A single dose of BPG may eventually achieve serological cure, but a three-dose regimen might simply accelerate the rate of response. This is only hypothetical, the precise biological mechanism underlying this finding remains to be elucidated.

Consequently, the routine clinical application of a three-dose BPG regimen for early latent syphilis remains a subject of debate. Despite the signal toward therapeutic benefit, several practical factors must be heavily weighed, including patient adherence, the discomfort associated with multiple intramuscular injections, and ongoing shortages of BPG [23,24].

Despite signals of benefit in some subgroup populations, the trial sequential analysis serves as a crucial caveat. While we found no overall benefit in the primary analyses, the data are statistically insufficient to declare the three-dose regimen “futile”. We have only accrued 34.1% of the necessary sample size to definitively rule out a modest effect. It is unlikely that a further large randomized controlled trial comparing the efficacy of these two regimens will be undertaken, given the challenges of retaining this highly mobile population through the end of follow-up, the high rate of reinfection, and the possibility that any incremental benefit of three-dose BPG may be small and confined to selected populations. In an era in which effective oral alternatives such as amoxicillin or doxycycline are available, these regimens may represent more attractive options. Therefore, while a universal policy shift to three doses is unsupported, the regimen cannot be completely discarded for selected high-risk cases.

Our meta-analysis has several limitations. The principal limitation is the small number of eligible studies ($n = 3$) and the dominance of studies at high or serious risk of bias which, together with substantial heterogeneity and imprecision, led GRADE to rate the certainty of evidence as very low for efficacy and low for safety. While our search strategy was exhaustive and included gray literature, the inability to retrieve data from one identified conference abstract may lead to an overestimation or underestimation of the true treatment effect. The significant findings in the high-RPR subgroup and the early latent post hoc analysis rest on these lower-quality studies and require cautious, exploratory interpretation. Residual confounding by indication cannot be excluded, as clinicians may have preferentially administered the three-dose regimen to patients judged at higher baseline risk of failure. Definitions and adjudication of presumed reinfection versus true treatment failure also varied across studies and were not reported at the individual-patient level; in particular, counting reinfections as failures (as in part of the Yang et al. [9] cohort) may have inflated the apparent benefit of three doses, and a sensitivity analysis formally excluding probable reinfections was not feasible. Additionally, the classification of early versus late latent syphilis relied on each study’s verification methods (clinical history and prior serology) and on an essentially arbitrary 12-month cut-off; misclassification of late latent cases as early latent, or biological variability in serological response, could therefore

have contributed to the apparent three-dose benefit observed in the early latent subgroup. Neurosyphilis screening protocols were not standardized across studies, so subclinical or asymptomatic neurosyphilis, including otosyphilis and ocular syphilis, cannot be excluded; however, routine lumbar puncture for asymptomatic disease is labor-intensive and raises cost-effectiveness and ethical concerns. Finally, Jarisch–Herxheimer reactions were not actively screened or monitored by Andrade et al. [8], likely underestimating their true incidence.

4. Conclusions

On the basis of the intention-to-treat evidence, a single dose of BPG appears to remain effective for most cases of early syphilis, including in PWH, and our findings support current CDC and WHO recommendations for single-dose treatment. The historical signal favoring three doses is most plausibly explained by bias in older studies. The apparent advantages of three doses in PWH with early latent syphilis and/or high baseline non-treponemal titers ($\geq 1:32$) emerged only from exploratory subgroup and post hoc per-protocol analyses of studies at high risk of bias, with very low certainty of evidence and an underpowered, inconclusive trial sequential analysis. These observations are hypothesis-generating and are not sufficient to recommend routine three-dose therapy for any subgroup; any proposed dose stratification remains speculative. Adequately powered, well-designed trials are needed to determine whether specific high-risk subgroups genuinely benefit from intensified dosing.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/ijtm6030026/s1>. Table S1: search strategy; Table S2: risk of bias assessment for 2 included randomized trials; Table S3: risk of bias assessment for 1 included prospective study; Figure S1: Forest plot of treatment response comparing three weekly doses versus a single dose of benzathine penicillin G (BPG) for people with HIV (PWH); Figure S2: Forest plot of treatment response comparing three weekly doses versus a single dose of benzathine penicillin G (BPG) for PWH with CD4 < 350 cells/mm³; Figure S3: Forest plot of treatment response comparing three weekly doses versus a single dose of benzathine penicillin G (BPG) for PWH on ART; Figure S4: Forest plot of treatment response comparing three weekly doses versus a single dose of benzathine penicillin G (BPG) for participants with baseline RPR $\geq 1:32$; Figure S5: Forest plot of treatment response comparing three weekly doses versus a single dose of benzathine penicillin G (BPG) for participants with previous syphilis; Figure S6: Funnel plot of 3 included studies; Figure S7: Trial sequential analysis (TSA) of treatment response comparing three weekly doses of benzathine penicillin G (BPG) with a single dose of BPG in early syphilis; Table S4: GRADE summary of serological response and adverse reactions.

Author Contributions: T.P.: Investigation, data curation, writing—original draft. S.S.: Validation, formal analysis, Supervision, writing—review and editing. S.P.: Conceptualization, methodology, investigation, data curation, formal analysis, project administration, writing—original draft, writing—review and editing. All authors have read and agreed to the published version of the manuscript.

Funding: This research project was supported by the Mahasarakham University (6-7/2569).

Institutional Review Board Statement: Ethical approval was obtained from the Mahasarakham University Ethics Committee for Research Involving Human Subjects on 28 November 2025 (Approval number: 824-824/2025).

Informed Consent Statement: This systematic review and meta-analysis used only aggregate data from previously published studies and involved no direct participant recruitment or identifiable individual-level data, the requirement for informed consent was waived.

Data Availability Statement: All data analyzed during this study are included in the published article and its Supplementary Information Files. The analytical datasets and R codes used for the meta-analysis are available from the corresponding author upon reasonable request.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

Abbreviation	Full Term
ART	Antiretroviral therapy
BPG	Benzathine penicillin G
CDC	Centers for Disease Control and Prevention
CD4	Cluster of differentiation 4
CI	Confidence interval
CSF	Cerebrospinal fluid
GRADE	Grading of Recommendations Assessment, Development and Evaluation
HAART	Highly active antiretroviral therapy
HIV	Human immunodeficiency virus
I ²	I-squared statistic
ITT	Intention-to-treat
MIC	Minimum inhibitory concentration
MSM	Men who have sex with men
MU	Million units
NNH	Number needed to harm
NNT	Number needed to treat
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PROSPERO	International Prospective Register of Systematic Reviews
PWH	People with HIV
RCT	Randomized controlled trial
REML	Restricted maximum likelihood
RIS	Required information size
RoB 2	Revised Cochrane risk-of-bias tool for randomized trials
ROBINS-I	Risk Of Bias In Non-randomized Studies—of Interventions
RPR	Rapid plasma reagin
RR	Risk ratio
TSA	Trial sequential analysis
VDRL	Venereal Disease Research Laboratory
WHO	World Health Organization
df	Degrees of freedom
IU/mL	International units per milliliter
k	Number of studies
p	Probability value
Q	Cochran's Q statistic
τ^2	Between-study variance
μg/mL	Micrograms per milliliter

References

1. Workowski, K.A.; Bachmann, L.H.; Chan, P.A.; Johnston, C.M.; Muzny, C.A.; Park, I.; Reno, H.; Zenilman, J.M.; Bolan, G.A. Sexually Transmitted Infections Treatment Guidelines, 2021. *MMWR. Recomm. Rep.* **2021**, *70*, 1–187. [CrossRef] [PubMed]
2. World Health Organization. *WHO Consolidated Operational Handbook on Sexually Transmitted Infections*; World Health Organization: Geneva, Switzerland, 2026; Available online: <https://www.who.int/publications/i/item/9789240120365> (accessed on 10 January 2026).

3. Andrade, R.A.; Villareal-Williams, E.; Risser, J. Co-infection with *Treponema pallidum* and HIV: Effect of the CD4 cell count and penicillin dosage in the response of syphilis to therapy. In Proceedings of the 4th IAS Conference on HIV Pathogenesis, Treatment and Prevention, MOPEB104, Sydney, Australia, 22–25 July 2007.
4. Tittes, J.; Aichelburg, M.C.; Antoniewicz, L.; Geusau, A. Enhanced therapy for primary and secondary syphilis: A longitudinal retrospective analysis of cure rates and associated factors. *Int. J. STD AIDS* **2013**, *24*, 703–711. [[CrossRef](#)] [[PubMed](#)]
5. Dowell, D.; Polgreen, P.M.; Beekmann, S.E.; Workowski, K.A.; Berman, S.M.; Peterman, T.A. Dilemmas in the management of syphilis: A survey of infectious diseases experts. *Clin. Infect. Dis.* **2009**, *49*, 1526–1529. [[CrossRef](#)] [[PubMed](#)]
6. Page, M.J.; McKenzie, J.E.; Bossuyt, P.M.; Boutron, I.; Hoffmann, T.C.; Mulrow, C.D.; Shamseer, L.; Tetzlaff, J.M.; Akl, E.A.; Brennan, S.E.; et al. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ* **2021**, *372*, n71. [[CrossRef](#)] [[PubMed](#)]
7. Hook, E.W., 3rd; Dionne, J.A.; Workowski, K.; McNeil, C.J.; Taylor, S.N.; Batteiger, T.A.; Dombrowski, J.C.; Mayer, K.H.; Seña, A.C.; Hamill, M.M.; et al. One dose versus three doses of benzathine penicillin G in early syphilis. *N. Engl. J. Med.* **2025**, *393*, 869–878. [[CrossRef](#)] [[PubMed](#)]
8. Andrade, R.; Rodriguez-Barradas, M.C.; Yasukawa, K.; Villarreal, E.; Ross, M.; Serpa, J.A. Single dose versus 3 doses of intramuscular benzathine penicillin for early syphilis in HIV: A randomized clinical trial. *Clin. Infect. Dis.* **2017**, *64*, 759–764. [[CrossRef](#)] [[PubMed](#)]
9. Yang, C.J.; Lee, N.Y.; Chen, T.C.; Lin, Y.H.; Liang, S.H.; Lu, P.L.; Huang, W.C.; Tang, H.J.; Lee, C.H.; Lin, H.H.; et al. One dose versus three weekly doses of benzathine penicillin G for patients co-infected with HIV and early syphilis: A multicenter, prospective observational study. *PLoS ONE* **2014**, *9*, e109667. [[CrossRef](#)] [[PubMed](#)]
10. Marchese, V.; Tiecco, G.; Storti, S.; Degli Antoni, M.; Calza, S.; Gulletta, M.; Viola, F.; Focà, E.; Matteelli, A.; Castelli, F.; et al. Syphilis infections, reinfections and serological response in a large Italian sexually transmitted disease centre: A monocentric retrospective study. *J. Clin. Med.* **2022**, *11*, 7499. [[CrossRef](#)] [[PubMed](#)]
11. Seña, A.C.; Zhang, X.H.; Li, T.; Zheng, H.P.; Yang, B.; Yang, L.G.; Salazar, J.C.; Cohen, M.S.; Moody, M.A.; Radolf, J.D.; et al. A systematic review of syphilis serological treatment outcomes in HIV-infected and HIV-uninfected persons: Rethinking the significance of serological non-responsiveness and the serofast state after therapy. *BMC Infect. Dis.* **2015**, *15*, 479. [[CrossRef](#)] [[PubMed](#)]
12. González-López, J.J.; Guerrero, M.L.; Luján, R.; Tostado, S.F.; de Górgolas, M.; Requena, L. Factors determining serologic response to treatment in patients with syphilis. *Clin. Infect. Dis.* **2009**, *49*, 1505–1511. [[CrossRef](#)] [[PubMed](#)]
13. Spagnuolo, V.; Poli, A.; Galli, L.; Nozza, S.; Bossolasco, S.; Cernuschi, M.; Maillard, M.; Hasson, H.; Gianotti, N.; Guffanti, M.; et al. Incidence and predictors of serological treatment response in early and late syphilis among people living with HIV. *Open Forum. Infect. Dis.* **2018**, *6*, ofy324. [[CrossRef](#)] [[PubMed](#)]
14. Tong, M.L.; Lin, L.R.; Liu, G.L.; Zhang, H.L.; Zeng, Y.L.; Zheng, W.H.; Liu, L.L.; Yang, T.C. Factors associated with serological cure and the serofast state of HIV-negative patients with primary, secondary, latent, and tertiary syphilis. *PLoS ONE* **2013**, *8*, e70102. [[CrossRef](#)] [[PubMed](#)]
15. Lukehart, S.A.; Hook, E.W., 3rd; Baker-Zander, S.A.; Collier, A.C.; Critchlow, C.W.; Handsfield, H.H. Invasion of the central nervous system by *Treponema pallidum*: Implications for diagnosis and treatment. *Ann. Intern. Med.* **1988**, *109*, 855–862. [[CrossRef](#)] [[PubMed](#)]
16. Polnikorn, N.; Witoonpanich, R.; Vorachit, M.; Vejajiva, S.; Vejajiva, A. Penicillin concentrations in cerebrospinal fluid after different treatment regimens for syphilis. *Br. J. Vener. Dis.* **1980**, *56*, 363–367. [[CrossRef](#)] [[PubMed](#)]
17. Ducas, J.; Robson, H.G. Cerebrospinal fluid penicillin levels during therapy for latent syphilis. *JAMA* **1981**, *246*, 2583–2584. [[CrossRef](#)] [[PubMed](#)]
18. Norris, S.J.; Edmondson, D.G. In vitro culture system to determine MICs and MBCs of antimicrobial agents against *Treponema pallidum* subsp. *pallidum* (Nichols strain). *Antimicrob. Agents Chemother.* **1988**, *32*, 68–74. [[CrossRef](#)] [[PubMed](#)]
19. Barza, M. Antimicrobial spectrum, pharmacology and therapeutic use of antibiotics. Part 2: Penicillins. *Am. J. Hosp. Pharm.* **1977**, *34*, 57–67. [[CrossRef](#)] [[PubMed](#)]
20. Idsoe, O.; Guthe, T.; Willcox, R.R. Penicillin in the treatment of syphilis. The experience of three decades. *Bull. World Health Organ.* **1972**, *47*, 1–68. [[PubMed](#)]
21. Eagle, H.; Fleischman, R.; Musselman, A.D. The effective concentrations of penicillin in vitro and in vivo for streptococci, pneumococci, and *Treponema pallidum*. *J. Bacteriol.* **1950**, *59*, 625–643. [[CrossRef](#)] [[PubMed](#)]
22. Smith, C.A.; Kamp, M.; Olansky, S.; Price, E.V. Benzathine penicillin G in the treatment of syphilis. *Bull. World Health Organ.* **1956**, *15*, 1087–1096. [[PubMed](#)]

23. Mangone, E.; Bell, J.; Khurana, R.; Taylor, M.M. Treatment completion with three-dose series of benzathine penicillin among people diagnosed with late latent and unknown duration syphilis, Maricopa County, Arizona. *Sex. Transm. Dis.* **2023**, *50*, 298–303. [[CrossRef](#)] [[PubMed](#)]
24. Nurse-Findlay, S.; Taylor, M.M.; Savage, M.; Mello, M.B.; Saliyou, S.; Lavayen, M.; Seghers, F.; Campbell, M.L.; Birgirimana, F.; Ouedraogo, L.; et al. Shortages of benzathine penicillin for prevention of mother-to-child transmission of syphilis: An evaluation from multi-country surveys and stakeholder interviews. *PLoS Med.* **2017**, *14*, e1002473. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.