The Impact of Dosing Frequency on the Efficacy of 10-Day Penicillin or Amoxicillin Therapy for Streptococcal Tonsillopharyngitis: A Meta-analysis

Andrew J. Lan, MPH*, and John M. Colford, Jr, MD, PhD‡§

ABSTRACT. Objective. The recommended dosing frequency of oral penicillin for the treatment of acute streptococcal tonsillopharyngitis has long been 3 to 4 times daily. In 1994, treatment guidelines included twice-daily (BID) dosing for the first time, a recommendation that could significantly increase the ease of compliance. This meta-analysis was performed to determine whether overall cure rates differed between BID or once-daily (QD) versus more frequent dosing schedules in the treatment of streptococcal tonsillopharyngitis.

Data Sources. Candidate studies for this meta-analysis included all clinical trials of therapy for streptococcal tonsillopharyngitis published through August 1998 and identified using Medline, Dissertation Abstracts, conference proceedings, and bibliographies of all retrieved articles.

Study Selection. A study was eligible for inclusion if it was a randomized clinical trial that compared the efficacies of different dosing frequencies of 10-day penicillin or amoxicillin in the treatment of streptococcal tonsillopharyngitis. Of the 30 articles initially identified, 6 studies met eligibility criteria.

Outcome Measure. The measure of interest was the difference in proportion cured between the BID or QD dosing group and the comparison group with more frequent dosing.

Results. The results of this analysis suggest that BID dosing of 10-day penicillin is as efficacious as more frequent dosing regimens in the treatment of streptococcal tonsillopharyngitis. This result also holds true in a subgroup analysis confined to pediatric cases and does not vary with total daily dose of the regimen. QD dosing of penicillin is associated with a cure rate that is 12 percentage points lower than more frequent dosing (95% confidence interval: 3–21). In contrast, this decreased efficacy is not found with QD dosing of amoxicillin.

Conclusions. This meta-analysis supports current recommendations for BID dosing of penicillin in treating streptococcal tonsillopharyngitis. QD penicillin is associated with decreased efficacy and should not be used. Simplified regimens of amoxicillin of shorter duration or of less frequent dosing should be further investigated.

PEDIATRICS 2000;105(2). URL: http://www.pediatrics.org/cgi/content/full/105/2/e19; amoxicillin, group A streptococcal, penicillin, pharyngitis, streptococcal infections, Streptococcus pyogenes.

ABBREVIATIONS. GABHS, group A β-hemolytic streptococcus; BID, twice-daily; QD, once-daily; TID, 3-times-daily; QID, 4-times-daily; CI, confidence interval.

Group A β-hemolytic streptococcus (GABHS) is the most common bacterial cause of acute tonsillopharyngitis in the pediatric population.1 An estimated 11% of all school-aged children in the United States visit a doctor every year for tonsillopharyngitis.2 Diagnosis and prompt treatment of streptococcal tonsillopharyngitis is essential to prevent acute rheumatic fever, which can lead to permanent heart valve damage and the subsequent need for lifelong antibiotics, and to prevent supplicative complications, such as peritonsillar abscess and cervical adenitis.3

The earliest clinical trials on treatment of streptococcal tonsillopharyngitis led to a recommendation first issued in 1953 by the American Heart Association for oral treatment with a 10-day regimen of penicillin taken 3 to 4 times daily.4 Despite many clinical trials evaluating this topic in subsequent years, the recommended antibiotic type and duration of treatment have not changed even to the present day.

Penicillin remains the antibiotic of choice, because its effectiveness in treating streptococcal tonsillopharyngitis is undiminished, despite many claims to the contrary.5 Isolation of a penicillin-resistant strain of GABHS has yet to occur.6 Newer antibiotics such as cephalosporins and macrolides, although possessing at least equal efficacy and more convenient dosing, have been associated with higher costs as well as the worrisome problem of increased antibiotic resistance attributable to their wider spectrum of activity than that of penicillin.1 Indeed, in Finland and Japan, increased macrolide use coincided with increased macrolide resistance in local bacterial strains, a trend that was reversed in Finland when national guidelines were changed to decrease the use of erythromycin for such infections as streptococcal tonsillopharyngitis.1

The recommendation that the penicillin course last 10 days also has not changed, because studies have shown that shortening the course of treatment is associated with lower cure rates.7–10 However, in recognition of the fact that compliance with a 10-day...
regimen taken 3 to 4 times daily is difficult for most patients, clinical trials continue to explore the possibility of simplifying the treatment of streptococcal tonsillopharyngitis without compromising efficacy. In particular, reduction of the required regimen to twice-daily (BID) dosing would significantly increase the ease of compliance, because children would not have to receive a dose of antibiotic in the middle of the day while in day care or school.

In 1995, the American Heart Association for the first time included BID dosing in its guidelines for treating streptococcal tonsillopharyngitis. BID dosing also appeared for the first time in the Red Book of the American Academy of Pediatrics in its 1994 edition. In light of these changes, it is surprising that the American Academy of Pediatrics in its 1994 edition did not provide serologic response. During certain seasons, as individuals documented to have a throat culture the existence of streptococcal carriers, ie, those in higher strata was performed. Also, because the total daily dose of penicillin regimens. Thus, there is the distinct possibility that differences in efficacy do exist between the various regimens that smaller individual trials are not powerful enough to detect. Notably, such references as Physicians’ Desk Reference and Drug Evaluations of the American Medical Association do not list once-daily (QD) or BID dosing in their guidelines for penicillin use.

This meta-analysis examines all clinical trials that directly compare QD or BID dosing of penicillin or amoxicillin with 3-times-daily (TID) or 4-times-daily (QID) dosing in the treatment of GABHS tonsillopharyngitis. The goal of this study is to determine whether there is any overall evidence that QD or BID dosing is less efficacious than more frequent dosing regimens. Studies of amoxicillin have been included in this meta-analysis because amoxicillin is often used both in clinical trials as well as in clinical practice as an alternative to penicillin, because it compares favorably with respect to gastrointestinal tolerability, serum and tonsillar half-life, cost, absorption that is less affected by food ingestion, and spectrum of bacterial coverage that is only slightly wider than that of penicillin.

This meta-analysis also includes 2 subanalyses to determine whether the results are robust for specific groups. Because streptococcal tonsillopharyngitis leading to rheumatic fever is of particular significance in children, sensitivity analysis of the studies that enrolled predominantly children was performed. Also, because the total daily dose of antibiotic may influence the relationship between dosing frequency and cure rate, stratified analysis separating the studies into higher and lower total daily dose strata was performed.

An important concept in this field of research is the existence of streptococcal carriers, ie, those individuals documented to have a throat culture result positive for GABHS but no detectable illness or serologic response. During certain seasons, as many as 20% of schoolchildren may be carriers. Carriers are not at risk for acute rheumatic fever, but if treatment with a full course of penicillin is nevertheless given, the carrier state often persists. The existence of streptococcal carriers underlines the importance of using strict clinical criteria when evaluating patients for enrollment in trials. For example, inadvertent inclusion of streptococcal carriers who happen to have an acute viral pharyngitis in a randomized trial would lead to non-differential misclassification and result in type II error (that is, concluding that a difference in efficacy does not exist when in fact does). To evaluate objectively the design quality of the studies included in this meta-analysis, a quality scoring scale is presented by which all the trials have been scored.

METHODS

Study Identification and Inclusion Criteria

All candidate studies published through August 1998 were gathered using computer searches of Medline and Dissertation Abstracts. The Medline search used the command “FIND (keyword tonsillopharyngitis or keyword pharyngitis) and (keyword penicillin or keyword amoxicillin) and publication type clinical trial.” The Dissertation Abstracts database was searched using the following word combinations: streptococcus and penicillin, streptococcus and amoxicillin, streptococcal and penicillin, and streptococcal and amoxicillin. Relevant unpublished papers were searched for by perusing abstracts of the annual Interscience Conference on Antimicrobial Agents and Chemotherapy from 1983 to 1997 for the keywords “pharyngitis” and “tonsillopharyngitis” (keyword indexing of abstracts was not available before this time period).

A study was included in the meta-analysis if it possessed all the following characteristics: 1) it was a randomized clinical trial comparing the efficacies of different dosing frequencies of 10-day oral penicillin or amoxicillin in the treatment of acute GABHS tonsillopharyngitis; 2) the diagnosis was made based on both symptoms and positive throat culture results or positive antigen detection test results; 3) antibiotic treatment was initiated on diagnosis; and 4) the results of the study were not published elsewhere. If multiple publications arose from the same study population, only the more recent publication was included in the meta-analysis.

Unless a study title or abstract clearly indicated that it did not meet the above inclusion criteria, all articles resulting from the Medline search were retrieved. The Dissertation Abstracts search described above yielded no articles. The bibliographies of all retrieved studies, review articles, and meta-analyses were examined for any additional studies that might be candidates for inclusion in the meta-analysis.

Data Abstraction and Study Quality Assessment

Data were abstracted in an unblinded fashion from studies using a standardized data collection form. A list of desirable study characteristics was assembled from various discussions about study design quality in the context of both clinical trials evaluating antibiotic treatment of streptococcal tonsillopharyngitis as well as clinical trials in general. The presence or absence of the following characteristics was noted for each study. Steps taken to improve diagnostic accuracy: 1) Study performed between December and May; 2) Subjects under 3 years of age excluded from study; 3) Diagnosis established based on clinical symptoms consistent with streptococcal tonsillopharyngitis (pharyngeal pain, dysphagia, fever, malaise, headache, abdominal pain, vomiting, erythematous pharynx, patchy exudate on posterior pharynx or tonsils, palatal petechiae, and enlarged and tender anterior cervical lymph nodes) and signs and symptoms consistent with viral pharyngitis not included (rhinorrhea, cough, hoarseness, conjunctivitis, and diarrhea). Steps to decrease and distinguish subsequent reinfection after cure from persistent infection: 4) Data available on follow-up cultures performed fewer than 14 days after completing treatment; 5) Comparison of pretreatment and posttreatment GABHS serotypes for subjects with positive follow-up culture results. General elements of study design: 6) Investigators responsible for follow-up assessment were blinded to antibiotic regimen; 7) Compliance with antibiotic regimen assessed; 8) Subject withdrawals and reasons for withdrawal discussed; and 9) Adequacy of randomization evaluated for subject demographic variables.
A quality score was then calculated for each study in the meta-analysis, with a maximum possible score of 12 points. Two points were assigned for fulfillment of criteria 3, 4, and 5, and 1 point was assigned for the remaining criteria. Because the reliability and validity of quality scoring has yet to be formally evaluated for use in meta-analysis, scoring was intended only to provide the reader with a summary indication of the quality of each clinical trial being examined in this meta-analysis.

Statistical Analysis

The measure of effect of interest for this meta-analysis was difference in the proportion cured between the 2 treatment arms under comparison in each trial. The proportion cured was calculated as the number of culture-negative subjects at follow-up divided by total number of subjects in the treatment arm. In studies with more than 2 treatment arms, meaningful 2-arm comparisons were constructed with the following in mind: the primary objective of the meta-analysis, which was to determine whether there were cure rate differences between TID or QID and QD or BID dosing; and the preference for comparing treatment arms that were as similar as possible with respect to total daily dose (to avoid any confounding effect that total daily dose might have on cure rate).

Because the measure of effect under analysis was a rate difference, the general variance-based method was used for calculation of summary measures of effect and for the test of homogeneity.26 If the test of homogeneity indicated the possibility of significant heterogeneity ($P < .10$) such that presentation of a summary measure of effect would not be meaningful, the studies being compared were investigated to determine the likely source of heterogeneity and appropriate additional analyses were performed.

Summary measures of effect were calculated for studies comparing TID or QID dosing with BID dosing and for studies comparing TID or QID dosing with QD dosing. Because GABHS tonsillopharyngitis leading to rheumatic fever is of particular significance in children, sensitivity analysis of the studies predominately enrolling children was conducted. Finally, because total daily dose may influence the relationship between dosing frequency and cure rate, stratified analysis was performed separating the studies into higher and lower total daily dose strata to determine whether any statistical interaction or confounding was present.

Publication bias is a concern in the interpretation of any meta-analysis. This bias is a distortion in the calculated summary measure of effect arising from the fact that not all research pertaining to a particular topic may have been published. Because this meta-analysis includes only published papers, the results of unpublished papers, if any, will be missed. The presence of publication bias was assessed by searching abstracts of the annual Interscience Conference on Antimicrobial Agents and Chemotherapy during the period from 1983 to 1997 for pertinent unpublished papers. Another technique used in the evaluation of publication bias, the funnel plot, did not contribute meaningfully to the assessment because of the existence of few studies qualifying for this meta-analysis.26

RESULTS

Summary of Search Process and Studies Included in the Meta-analysis

Thirty articles that potentially fulfilled inclusion criteria were retrieved. Fifteen of these studies were identified through Medline, and the remaining 15 were identified through manual searching of bibliographies. Six of the 30 studies were retained in the meta-analysis. The excluded studies and reasons for exclusion are shown in Table 1. The most common reason for exclusion was the presence of only TID penicillin arms, seen in 9 studies.

A summary of the 6 studies included in the meta-analysis appears in Table 2.27-32 All but 2 of the studies enrolled chiefly pediatric subjects.27-30 Only 1 study contained an amoxicillin treatment arm.31 Although all studies used penicillin V, only the study by Kaufhold et al29 specified the particular salt form of penicillin V used, and this study used 2 different salts, potassium and benzathine, for the TID and BID regimens, respectively. There is evidence that these 2 salts possess different pharmacokinetic half-lives.

All studies compared TID or QID dosing with QD or BID dosing. The study by Krober et al30 contained 3 treatment arms comparing QID with BID with QD dosing; for the analysis, this study was split into 2 comparisons, the QID arm with the BID arm, and the

* M indicates Medline search; B, bibliography scan.

<table>
<thead>
<tr>
<th>Authors, Year</th>
<th>Search Method*</th>
<th>Reason for Exclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Breese, 1953</td>
<td>B</td>
<td>Not confined to streptococcal pharyngitis</td>
</tr>
<tr>
<td>Breese and Disney, 1958</td>
<td>B</td>
<td>All treatment arms TID</td>
</tr>
<tr>
<td>Breese et al, 1977</td>
<td>M</td>
<td>All treatment arms TID</td>
</tr>
<tr>
<td>Breese et al, 1965</td>
<td>B</td>
<td>Not confined to streptococcal pharyngitis</td>
</tr>
<tr>
<td>Breese et al, 1974</td>
<td>M</td>
<td>All treatment arms TID</td>
</tr>
<tr>
<td>Breese et al, 1974</td>
<td>B</td>
<td>Comparison of penicillin with other antibiotics</td>
</tr>
<tr>
<td>Colcher and Bass, 1972</td>
<td>M</td>
<td>Comparison of oral vs intramuscular penicillin</td>
</tr>
<tr>
<td>Edmond et al, 1966</td>
<td>M</td>
<td>7-day regimens, all treatment arms TID</td>
</tr>
<tr>
<td>Ginsburg et al, 1982</td>
<td>M</td>
<td>Comparison of oral vs intramuscular penicillin</td>
</tr>
<tr>
<td>Gopichand et al, 1998</td>
<td>M</td>
<td>All treatment arms TID</td>
</tr>
<tr>
<td>Howie et al, 1971</td>
<td>B</td>
<td>Comparison of oral vs intramuscular penicillin</td>
</tr>
<tr>
<td>Huang and High, 1953</td>
<td>B</td>
<td>Not confined to streptococcal pharyngitis</td>
</tr>
<tr>
<td>Matsen et al, 1974</td>
<td>B</td>
<td>Comparison of oral vs intramuscular penicillin</td>
</tr>
<tr>
<td>McLinn, 1983</td>
<td>B</td>
<td>Antibiotic tested was cephalaxin</td>
</tr>
<tr>
<td>Miller, 1958</td>
<td>B</td>
<td>Subjects all had history of rheumatic fever</td>
</tr>
<tr>
<td>Pankey et al, 1981</td>
<td>M</td>
<td>Treatment not randomized</td>
</tr>
<tr>
<td>Rabinowitch et al, 1973</td>
<td>B</td>
<td>Comparison of penicillin with other antibiotics</td>
</tr>
<tr>
<td>Schalet et al, 1958</td>
<td>B</td>
<td>All treatment arms TID</td>
</tr>
<tr>
<td>Shapera et al, 1973</td>
<td>M</td>
<td>Penicillin was given intramuscularly</td>
</tr>
<tr>
<td>Stillerman et al, 1960</td>
<td>B</td>
<td>All treatment arms TID</td>
</tr>
<tr>
<td>Stillerman et al, 1964</td>
<td>B</td>
<td>All treatment arms TID, not randomized</td>
</tr>
<tr>
<td>Stillerman et al, 1973</td>
<td>B</td>
<td>Treatment not randomized</td>
</tr>
<tr>
<td>Stillerman et al, 1974</td>
<td>M</td>
<td>All treatment arms TID</td>
</tr>
<tr>
<td>Vann et al, 1972</td>
<td>B</td>
<td>Only single penicillin regimen was tested</td>
</tr>
</tbody>
</table>

http://www.pediatrics.org/cgi/content/full/105/2/e19 3 of 8
QID arm with the QD arm. Statistical dependence of observations was not problematic because analyses were always stratified into BID and QD dosing; thus, in no analysis was it necessary to use artificially created pooled comparisons.

Total daily dose was comparable between treatment arms within each study with the exception of the study of Kaufhold et al., in which 1 of the treatment regimens contained double the total daily dose of the other 2 regimens. Therefore, this arm was excluded from the analysis to minimize possible confounding by total daily dose.

Cure was uniformly defined as culture-negativity at every follow-up culture. If serotyping was performed, subjects who were infected with a new serotype on follow-up were considered cured. A bacteriologic failure was uniformly defined as culture-positivity on any follow-up throat swabbing.

However, the timing and number of follow-up throat cultures were not uniform among the various studies. Studies scheduled anywhere from 1 to 3 follow-up cultures, ranging from 1 to 24 days after completion of therapy; therefore, an attempt was made to increase the comparability of results for this meta-analysis. It has been recommended that subsequent cultures be obtained no longer than 14 days after completion of treatment to minimize reinfection with the same GABHS strain. Thus, specifically for the study by Gerber et al., information concerning only this 2-week follow-up period was noted and used for the analysis, to the exclusion of longer-term data; however, because of the way in which data were presented in this particular study, serotyping information was then rendered unusable. Restriction to a 14-day follow-up period could not be achieved for those studies in which follow-up occurred exclusively after 14 days and in which data for all follow-up periods were pooled together.

Finally, studies varied in the management of cultures taken in the midst of treatment that were found to be positive. Of note, 2 studies, both testing a QD dosing regimen, cultured subjects 24 to 48 hours after beginning antibiotics and removed any subjects from the QD arm whose culture result remained positive at that time, likely out of safety concerns. However, Gerber et al. chose to exclude 4 subjects from the analysis completely, whereas Shvartzman et al. counted their 3 subjects as failures. For this meta-analysis, the 4 subjects from the study by Gerber et al. were excluded rather than counted as failures because the 18- to 24-hour follow-up period of the study seemed too short to evaluate bacteriologic resolution. In fact, the American Academy of Pediatrics recommends that children with streptococcal tonsillopharyngitis remain at home for at least 24 hours because of the evidence that this is the typical amount of time required before throat culture results become negative. In contrast, it was decided that the 3 subjects in the study by Shvartzman et al. would be counted as failures. In this study, cultures were performed 1 to 2 and 14 to 21 days after starting therapy, respectively.
treatment, and all positive culture results were counted as failures, even if from the same subject on 2 occasions. To avoid double counting in the meta-analysis, only culture data on day 2 were used, and consistency was maintained by counting all positive culture results as failures, regardless of treatment arm membership.

Difference in the proportion cured was calculated as the cure rate of the more frequent dosing arm minus the cure rate of the less frequent dosing arm. These differences are presented graphically with 95% confidence intervals (CIs) in the upper half of Fig 1. In this figure, studies were grouped by the presence of BID or QD dosing arms and were ordered within the 2 groups by quality score so that trends in effect measure based on quality score, if any, could be easily noted. No trends were evident.

---

**Table 1.** Differences in proportion cured (more frequent dosing minus less frequent dosing)

<table>
<thead>
<tr>
<th>Study Description</th>
<th>Effect Size (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Gerber, Randolph, et al., 1989</td>
<td>0.083 (0.064-0.103)</td>
</tr>
<tr>
<td>2. Krober, Weir, et al., 1990</td>
<td>0.204 (0.173-0.235)</td>
</tr>
<tr>
<td>3. Shvartzman, Tabenkin, et al., 1993</td>
<td>-0.033 (-0.056-0.011)</td>
</tr>
</tbody>
</table>

**Figure 1.** Differences in proportion cured for individual studies and summary measures of effect with 95% CIs.
The quality scoring of studies is detailed in Table 3. Scores showed much variation, with a range of 1 to 9 of a maximum of 12. The average score was 5 and the median score was 4.

**Principal Analyses and Subanalyses**

All summary proportion differences are shown with 95% CIs in the lower half of Fig 1. There was significant evidence of heterogeneity for the analysis of TID or QID versus QD dosing ($P = .014$). On examination of Fig 1, the source of this heterogeneity seemed to be the study by Shvartzman et al, which firmly demonstrated no difference in efficacy between TID or QID versus QD dosing, whereas the other 2 studies showed TID and QID dosing to possess significantly better efficacy. Interestingly, the study by Shvartzman et al was unique in that it examined amoxicillin. Although amoxicillin is considered to be pharmacologically equivalent to penicillin, there is evidence that amoxicillin is better absorbed and produces higher and longer serum and tissue concentrations. In additional support of this finding, recent studies published in this journal found that amoxicillin given TID for 6 days or QD for 10 days was as efficacious as standard 10-day penicillin. Therefore, it seemed reasonable to repeat the analysis excluding the study that used amoxicillin. Doing so resulted in the significantly better efficacy of TID and QID over QD dosing, by 12 percentage points.

A similar analysis of TID and QID versus BID dosing demonstrated no significant difference in cure rate. Because streptococcal tonsillitis leading to rheumatic fever occurs predominantly in children, this analysis was repeated for studies including only pediatric subjects (younger than 21 years of age). This subgroup analysis also demonstrated no difference in cure rate. Finally, to evaluate for possible interaction effects among total daily dose, dosing frequency, and differences in cure efficacy, stratified analysis was performed by splitting the BID dosing studies into studies with higher total daily doses and studies with lower total daily doses. One might hypothesize that with lower total daily doses, less frequent dosing regimens are more likely to show poorer efficacy, compared with more frequent dosing. However, in both the higher and lower dose strata, the summary cure rate for BID dosing was no different from that of TID or QID dosing.

**Assessment of Publication Bias**

Unpublished work on this topic was searched for as described in “Methods.” No relevant work was found from among all available abstracts of a major annual infectious disease conference (1983–1997).

**DISCUSSION**

This meta-analysis provides evidence that penicillin administered BID is as efficacious as TID or QID dosing in the treatment of streptococcal tonsillitis. This finding holds true for the pediatric population and is independent of total daily dose. This study also demonstrates that QD penicillin is less...
efficacious than TID or QID dosing by 12 percentage points.

Evaluation of the relative advantage of amoxicillin over penicillin in the treatment of streptococcal tonsillopharyngitis was not a goal of this meta-analysis. The possibility that amoxicillin may be superior to penicillin, perhaps attributable to pharmacokinetic or microbicidal properties, was raised as the most likely explanation for why the QD study by Shvartzman et al.31 arrived at a different conclusion from the other 2 studies employing QD dosing. As discussed above, recent studies demonstrating the equivalency of amoxicillin given TID for 6 days or QD for 10 days and standard penicillin therapy34,35 should spur increased interest in establishing whether simplified amoxicillin courses can replace 10-day penicillin as a treatment that is as efficacious, more convenient and tolerable for the patient, and whose bacterial spectrum is reasonably close to that of penicillin.

The lack of abstracts in conference proceedings of unpublished papers that would have otherwise qualified for this meta-analysis provides some evidence against the presence of publication bias. Additional evidence is the fact that a primary emphasis of this area of research is the demonstration of clinical equivalency between standard and lightened dosing regimens in the treatment of streptococcal tonsillopharyngitis. Because demonstrating treatment equivalency has been as important an objective as demonstrating treatment superiority or inferiority, the interest to publish clinical trials in this field would be expected to be similar, regardless of the direction of the findings.

Finally, meta-analyses are limited by the quality of the trials that they seek to summarize. Using a quality scoring scale with a maximum of 12 points, only 2 studies scored above 6 points. Although not used in any weighted analyses, the quality scoring was performed simply to illustrate the dearth of well-designed clinical trials in the field of streptococcal tonsillopharyngitis treatment. Even in the study with the highest quality score, that by Gerber et al.,27 which was published with statistically significant results, the results became insignificant (P < .12) after removal of the 14- to 21-day follow-up data (done to reduce misclassification bias). In the same study, 4 subjects were removed from the QD treatment arm per study protocol because their 18- to 24-hour follow-up culture results remained positive. One is left to wonder about the possibility that the 4 refractory cases may have been especially difficult to treat and, if continued on the QD regimen, would have eventually been classified as therapy failures.

Variation in study design can also make meta-analysis difficult. Certainly, the variation in the numbers of follow-up cultures taken at different times was a major consideration. This particular problem has been acknowledged in previous reviews,26 and even proved to influence the results of at least 1 study in this meta-analysis; as with the study by Gerber et al.,26 the results of the study by Shvartzman et al.31 had to be modified for the meta-analysis, in this case because of the possibility of double-counting treatment failures. Again, results that were significant as published were no longer significant on reanalysis. Certainly, any effort to standardize study design regarding the timing and number of throat cultures as well as the handling and counting of potential treatment failures would greatly enhance the comparability of future clinical trial results.

Acknowledging these limitations, the results of this meta-analysis must be carefully interpreted. The randomization process of clinical trials minimizes bias and confounding, and significant differences are more likely to be attributable to differences in treatment.37 In contrast, the low quality scores of the trials included in this meta-analysis indicate that frank evaluation of the state of the research on treatment of streptococcal tonsillopharyngitis continues to be necessary. Suggestions for improving on existing study designs should be made and incorporated into future clinical trials to enhance the quality and comparability of their findings.

REFERENCES


http://www.pediatrics.org/cgi/content/full/105/2/e19 7 of 8


