BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (http://bmjopen.bmj.com/site/about/resources/checklist.pdf) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

**ARTICLE DETAILS**

<table>
<thead>
<tr>
<th>TITLE (PROVISIONAL)</th>
<th>Adjunctive clindamycin for cellulitis: clinical trial comparing flucloxacillin with or without clindamycin for the treatment of limb cellulitis</th>
</tr>
</thead>
<tbody>
<tr>
<td>AUTHORS</td>
<td>Brindle, Richard; Williams, Owen; Davies, Paul; Harris, Tim; Jarman, Heather; Hay, Alastair; Featherstone, Peter</td>
</tr>
</tbody>
</table>

**VERSION 1 - REVIEW**

<table>
<thead>
<tr>
<th>REVIEWER</th>
<th>Sarah Walker</th>
</tr>
</thead>
<tbody>
<tr>
<td>University of Oxford, UK</td>
<td></td>
</tr>
<tr>
<td>REVIEW RETURNED</td>
<td>31-Jul-2016</td>
</tr>
</tbody>
</table>

**GENERAL COMMENTS**

This generally clearly written report describes a double-blind randomised controlled trial testing the value of two-days adjunctive clindamycin to standard beta-lactam monotherapy for cellulitis. The study was unfortunately troubled by substantial lost to follow-up and initial over-optimism about missing data, and was also sadly terminated early due to inadequate funding, but it still provides an important contribution to the literature.

I note that I agree with the way the authors have presented their primary outcome data, firstly as a percentage of those evaluable and secondly as a percentage of all randomised patients. The latter effectively counts missing data as failures (so-called ITT M=F), which also makes large assumptions which may not be justified in this population where those not returning had less severe disease and therefore plausibly are more likely to have been successes. The key point is that there was no difference between randomised groups in those not attending. Other more complex methods such as multiple imputation would have ideally been specified in a Statistical Analysis Plan before analysis: without this, straightforward presentations of the data as done in Table 2 are warranted.

Major comment

1. Methods, primary outcome p7 line 28ff and Results p9 line 6ff: the primary outcome had three components relating to temperature, limb swelling and erythema (the latter two relative to the unaffected limb). Missing data in one or more components might therefore have been envisaged at the trial design stage, in contrast to protocol section 22.5. However, it is not clear whether, when the amount of missing data was appreciated, a Statistical Analysis Plan was written detailing how it would be dealt with before analysis started. In the Results some rules are presented for classifying outcome data that would be better presented in Methods; and the text should be clear whether these were pre-specified in a Statistical Analysis Plan before data analysis or determined during the data analysis stage.
Further, the rules appear to be contradictory and should be clarified – line 7 suggests that those with temperature but without unaffected limb data were categorised as missing, whereas line 9 implies such patients were categorised as not improved. It is also a bit unclear why a patient with an improved temperature, but missing data on limb circumference and limb temperature, would be classified as not improved? A supplementary table detailing exactly how missing data in each component of the primary outcome was handled might be useful here.

Minor comments

2. Methods p6 line 17: my understanding from the protocol is that the exclusion criterion was reported allergy to penicillin/clindamycin, ie allergy testing was not conducted as part of the trial, but it would be helpful to clarify this.

3. Methods p7 line 42ff: from the protocol it appears that the original sample size calculation did not include any inflation factor for lost-to-follow-up. It would be helpful to clarify this in the main methods.

4. Table 1: rather than presenting the width of the IQR, it would be clearer to give the 25th to the 75th percentile for those variables which are markedly skewed so the reader can understand the overlap in distributions between the two groups. This would also give the reader some idea as to the proportions of patients in the various subgroups described later in the results (at the moment this is completely unclear).

5. Table 1: What % received IV vs oral beta-lactam in each group? Was there any impact of this in subgroup analyses?

6. Results, withdrawals p8 line 50-51: please provide the breakdown of withdrawals for personal reasons and AEs by randomised group.

7. Results, model for non-attendance p8 line 55: I assume that the factors affecting non-attendance are from a multivariable model, but it would be helpful to confirm this in the text. Also at minimum p-values should be provided for each factor, since it is not explicit what “significantly” means here. (Also p9 line 31)

8. Results, p9 line 16ff, Table2: Table 2 presents % to 0dp and p-values to 2dp, as standardly done. In contrast additional decimal places are included in both percentages and p-values in the main Results text for unclear reasons. In particular 3 dp for p-values >0.01 is spurious precision and these would be better presented as in Table 2. Also p9 line 57, p10 line 4 and line 9. Also differs in Table 4.

9. Table 3: if I understand the Methods correctly, all p-values except for pain VAS score at day 5 are from ANCOVA (adjusted for baseline). It would be helpful to state this in a footnote, as currently only the method for the p-value for the mann whitney u-test is clear.

10. Were any patients unblinded during their trial follow-up?

11. The protocol reports an additional outcome of reduction in the number of cellulitis recurrences at one-year follow-up (section 9.2.7). However it is not clear from the protocol how this was to be ascertained since the follow-up stopped at 30 days. Was this
endpoint ever collected?

**REVIEWER**
Josh Davis  
Menzies School of Health Research  
Darwin, NT, Australia

**REVIEW RETURNED**
02-Aug-2016

**GENERAL COMMENTS**

Intro  
Line 13/14 – need a citation to support assertion that microbiological studies are positive in 25% of patients  
Line 42/43 – Macrolide resistant strains of staph and strep are usually resistant to clindamycin, and many labs use erythromycin susceptibility testing as a surrogate for clindamycin. Hence this line seems misleading.  
Line 56/57 – Failure to demonstrate benefit of adjunctive clindamycin in mild cellulitis does not necessarily mean it’s use in much more severe conditions (nec fasc and toxic shock syndrome) need to be re-evaluated.

Methods  
Duration of adjunctive clindamycin was only 2 days – what evidence is there to support the potential efficacy of this duration? Why not use 5 days since the primary endpoint was assessed at day 5?  
Eligibility criteria – why were patients with abscesses/carbuncles/wounds not excluded? These patients may present with surrounding cellulitis, but are different from uncomplicated cellulitis in that they are more likely to have S.aureus and other pathogens in addition or alternative to Streptococci. Also they tend to resolve with drainage alone and the role of antibiotics is unclear.  
If such patients were indeed eligible (it is not entirely clear from the paper but should be made so), then can the authors provide the number in each group who had such presentations?  
Blinding. Clindamycin has a characteristic odour/taste. Were the placebo and clindamycin capsules identical in odour, taste and appearance?  
Line 54 – Was “core temperature” really measured? (this means using a rectal or oesophageal probe).  
Primary outcome – if there was an improvement in limb swelling but not temperature, how was this interpreted? And vice-versa? This needs to be made clearer. Was the 0.2 SD applicable to the skin surface temperature only or the swelling as well? Also please explain more clearly exactly how these numbers were determined (was the difference between the 2 limbs at day 5 compared with the difference at baseline? Or was the difference at day 5 the measurement in question? (i.e. was it change in the difference between the 2 limbs, or was it just the difference between the 2 limbs)?

Results  
Note the placebo group had higher CRP and thus potentially more marked inflammation – this should be commented on  
Primary outcome reporting – the way the primary outcome has been summarised (87% vs 81%) is really a per protocol analysis (not intention to treat as stated). The next row, reporting all those randomised (67% vs 68%) is the intention to treat analysis and should be the main reported outcome in the abstract, table and results.  
What number/proportion of patients were treated with IV versus oral
flucloxacillin in each group? What was the duration of IV flucloxacillin in those who received it?
How was IV flucloxacillin given (i.e. was this only possible for hospital inpatients?). What proportion were admitted to hospital and for how long? Was there any outpatient parenteral antibiotic therapy?

Discussion
Lines 9-10. As stated above I disagree it is reasonable to extrapolate these results to more severe GAS infections. This should be discussed in a more balanced way.
Discussion lines 12-32 – Nicely made points. However, lines 25-32 could also be an explanation of why clindamycin did not work in this trial: because it was started too late.

VERSION 1 – AUTHOR RESPONSE

Reviewer One
1. Methods, primary outcome p7 line 28ff and Results p9 line 6ff: the primary outcome had three components relating to temperature, limb swelling and erythema (the latter two relative to the unaffected limb). Missing data in one or more components might therefore have been envisaged at the trial design stage, in contrast to protocol section 22.5. However, it is not clear whether, when the amount of missing data was appreciated, a Statistical Analysis Plan was written detailing how it would be dealt with before analysis started. In the Results some rules are presented for classifying outcome data that would be better presented in Methods; and the text should be clear whether these were pre-specified in a Statistical Analysis Plan before data analysis or determined during the data analysis stage.

Further, the rules appear to be contradictory and should be clarified – line 7 suggests that those with temperature but without unaffected limb data were categorised as missing, whereas line 9 implies such patients were categorised as not improved. It is also a bit unclear why a patient with an improved temperature, but missing data on limb circumference and limb temperature, would be classified as not improved?

Answer
No formal analysis of the trial data was undertaken before the end of the trial but the exact determination of the primary outcome was established by the trial steering group during the trial. The decision to allocate a patient as a treatment failure or ‘missing’ was because the protocol had defined improved as being afebrile plus one of the local features having improved so a patient who had no measurement of body temperature could not be scored either way. The key point is that this definition was applied to both study allocations so it is not a source of bias.

We do not think an extra table would be very useful as most of the information is in the text and the whole dataset will be released.

Action
Clarify the definition as requested by Reviewer 2 and add a line of text reminding the reader of the protocol definition of the primary outcome within the results section.

2. Methods p6 line 17: my understanding from the protocol is that the exclusion criterion was reported allergy to penicillin/clindamycin, ie allergy testing was not conducted as part of the trial, but it would be helpful to clarify this.

Answer
Allergy was based on the patients’ statement of allergy or their medical records
Action
Add text on patients’ allergy

3. Methods p7 line 42ff: from the protocol it appears that the original sample size calculation did not include any inflation factor for lost-to-follow-up. It would be helpful to clarify this in the main methods.

Answer
The Feasibility section of the protocol discusses and estimates the recruitment rate and likely losses to follow up. Unfortunately there was no source of data to make secure estimates as the only UK trial of cellulitis in an emergency department (the primary source of patients for this study) used parenteral antibiotics. (Leman 2005). Most of our patients had oral therapy. Our original target was to recruit 450 patients but to adjust the duration of recruitment on losses, but delays in getting new sites (often >6 months from their approach to us) recruiting, meant we were unable to do this within the funded time frame.

Action
Add a statement about estimates of loss to follow up and the uncertainty associated with this.

4. Table 1: rather than presenting the width of the IQR, it would be clearer to give the 25th to the 75th percentile for those variables which are markedly skewed so the reader can understand the overlap in distributions between the two groups. This would also give the reader some idea as to the proportions of patients in the various subgroups described later in the results (at the moment this is completely unclear).

Answer
The use of interquartile range when presenting medians appears to be well established. We do not think the addition to or replacement of the IQR with the 25th to the 75th percentile (a measure of the range) for those variables which are markedly skewed (mostly blood results) would help the intended audience (general practitioners, emergency physicians and infection specialist) understand the overlap any better. A histogram presentation of the variables would show this well, but would mean a very large increase in the size of the paper. There are small differences between the allocations for most variables but these differences generally remain during follow-up suggesting that they were not influenced by differences in treatment allocation.

5. Table 1: What % received IV vs oral beta-lactam in each group? Was there any impact of this in subgroup analyses?

Answer
The quality of the data on duration and route of antibiotic was poor. Also, we had not collected data on dose which is why this data is not presented here because we assumed a standard dose of oral flucloxacillin but we underestimated the variations of dose, route and duration that were possible. We know from previous trials that the evidence does not support the use of IV therapy for cellulitis and this trial was not designed to determine differences in outcome associated with antibiotic route. We did not do a subgroup analysis by route of administration.

Action
Add a section of text stating why we did not do any analysis based on route of administration of flucloxacillin.

6. Results, withdrawals p8 line 50-51: please provide the breakdown of withdrawals for personal reasons and AEs by randomised group.
Action
Revise the section
Expanded the text and listed all withdrawals not just up to Day 5

7. Results, model for non-attendance p8 line 55: I assume that the factors affecting non-attendance are from a multivariable model, but it would be helpful to confirm this in the text. Also at minimum p-values should be provided for each factor, since it is not explicit what “significantly” means here. (Also p9 line 31)

Action
The numbers with percentages and actual p-values for each factor quoted have been added. Add text stating this was a univariate analysis.

8. Results, p9 line 16ff, Table2: Table 2 presents % to 0dp and p-values to 2dp, as standardly done. In contrast additional decimal places are included in both percentages and p-values in the main Results text for unclear reasons. In particular 3 dp for p-values >0.01 is spurious precision and these would be better presented as in Table 2. Also p9 line 57, p10 line 4 and line 9. Also differs in Table 4.

Action
Standardise presentation of p values as suggested.

9. Table 3: if I understand the Methods correctly, all p-values except for pain VAS score at day 5 are from ANCOVA (adjusted for baseline). It would be helpful to state this in a footnote, as currently only the method for the p-value for the mann whitney u-test is clear.

Action
Add ANCOVA statement in footnote

10. Were any patients unblinded during their trial follow-up?

Answer
Yes. We have a record of only one. We do not know what their allocation was but they were withdrawn and given open-label clindamycin

Action
Add a statement about unblinding

11. The protocol reports an additional outcome of reduction in the number of cellulitis recurrences at one-year follow-up (section 9.2.7). However it is not clear from the protocol how this was to be ascertained since the follow-up stopped at 30 days. Was this endpoint ever collected?

Answer
A proposal was put to RfPB, during the trial, for RfPB to fund clerical time for a one year follow up. This was refused as it was regarded as new application and would have extended the period of study for another year (beyond the usual RfPB term). We inserted this into the protocol at the point of application to RfPB (it was not in the original) but did not pursue it or another application.

Reviewer Two

1. Line 13/14 – need a citation to support assertion that microbiological studies are positive in 25% of patients
Answer
The statement is taken from the 2010 Cochrane Review of cellulitis and is an average of rates from different sources. The actual positivity rate varies from study to study, type of sample, type of test and whether one considers the bacterium isolated as causative. Reference 8 (Toleman et al) which studied patients from this trial includes the statement:
‘Despite utilising a range of diagnostic methods, a bacteriological diagnosis was only achieved in 43 % of patients with a clinical diagnosis of cellulitis. Seventeen per cent of patients tested positive for GAS by any method but only 4 % were positive by PCR, whilst S. aureus was detected in 34% of samples.’

Action
Revise the text to state ‘in a variable proportion’. The rest of the paragraph references different methodologies and rates.

2. Line 42/43 – Macrolide resistant strains of staph and strep are usually resistant to clindamycin, and many labs use erythromycin susceptibility testing as a surrogate for clindamycin. Hence this line seems misleading.

Answer
We do not think this is so. Many laboratories will report S.aureus and haemolytic streptococci as resistant to clindamycin but this is because they test for macrolide inducible resistance with a D test or similar. Actual clindamycin resistance and subsequent treatment failure is much rarer. The reporting of inducible resistance as opposed to actual resistance is of uncertain clinical value.

Action
Text checked and not changed.

3. Line 56/57 – Failure to demonstrate benefit of adjunctive clindamycin in mild cellulitis does not necessarily mean it’s use in much more severe conditions (nec fasc and toxic shock syndrome) need to be re-evaluated.

Answer
Thank you for raising this important point; this is a key part of this study and one of the reasons for doing it. Our view is that nearly everybody thought clindamycin was (‘is’ until this trial is published) of value in cellulitis. We have shown that we cannot demonstrate any benefit. We would like to see infection specialists undertaking a trial to establish the benefit of clindamycin in invasive group A streptococcal disease rather than continue to use it without trial data to support its benefit. Hence the statement for its use in these conditions needs (re-)evaluation.

4. Duration of adjunctive clindamycin was only 2 days – what evidence is there to support the potential efficacy of this duration? Why not use 5 days since the primary endpoint was assessed at day 5?

Answer
We could not find any supportive data for clindamycin to help us. The reason for 2 days was that we knew clindamycin was associated with diarrhoea and that, in general short antibiotic courses are as good as long, so we tried to keep the potential benefit and reduce the potential harm. If the study had shown benefit then further studies would have been necessary to clarify duration, route and dose.

Action
The reason for two days is already in the text (Methods).
Eligibility criteria – why were patients with abscesses/carbuncles/wounds not excluded? These patients may present with surrounding cellulitis, but are different from uncomplicated cellulitis in that they are more likely to have S.aureus and other pathogens in addition or alternative to Streptococci. Also they tend to resolve with drainage alone and the role of antibiotics is unclear. If such patients were indeed eligible (it is not entirely clear from the paper but should be made so), then can the authors provide the number in each group who had such presentations?

Answer
There was a section on the screening form for other potential exclusions and those patients with drainable abscess were excluded (most were never referred as they would have been surgical patients). Any that came to light were then removed from further follow up. Some were not discovered to have abscesses until they had completed the study drug (see the revised text on withdrawals). We did include patients with minor wounds who also had cellulitis.

Action
Added text. Patients with obvious abscesses were not eligible.

5. Line 54 – Was “core temperature” really measured? (this means using a rectal or oesophageal probe).

Answer
We did not use a rectal probe. The term ‘core’ was used to clearly differentiate this measurement from the skin surface temperature. The measurements were tympanic.

Action
Remove ‘core’

6. Primary outcome – if there was an improvement in limb swelling but not temperature, how was this interpreted? And vice-versa? This needs to be made clearer. Was the 0.2 SD applicable to the skin surface temperature only or the swelling as well? Also please explain more clearly exactly how these numbers were determined (was the difference between the 2 limbs at day 5 compared with the difference at baseline? Or was the difference at day 5 the measurement in question? (i.e. was it change in the difference between the 2 limbs, or was it just the difference between the 2 limbs)?

Answer
The first part is dealt with in the response to Reviewer 1. The 0.2 SD was applicable to both limb measurements. The difference between the two limbs at Day 5 determined improvement or not

Action
Clarify this section of text with additions as suggested.

7. Note the placebo group had higher CRP and thus potentially more marked inflammation – this should be commented on. Primary outcome reporting – the way the primary outcome has been summarised (87% vs 81%) is really a per protocol analysis (not intention to treat as stated). The next row, reporting all those randomised (67% vs 68%) is the intention to treat analysis and should be the main reported outcome in the abstract, table and results.

The two allocations differed in small ways in every parameter measured and these differences remained throughout the follow up.
Papers on trials vary in the way that results are presented: The (87% vs 81%) includes all those that were randomised and were evaluable (the denominator is all those evaluable) and the (67% vs 68%) uses the total randomised population as denominator. Reviewer 1 thinks this is an acceptable and clear presentation of the results.

8. What number/proportion of patients were treated with IV versus oral flucloxacillin in each group? 
What was the duration of IV flucloxacillin in those who received it? How was IV flucloxacillin given (i.e. was this only possible for hospital inpatients?). What proportion were admitted to hospital and for how long? Was there any outpatient parenteral antibiotic therapy?

Answer
As for Reviewer 1., Point 5.

Action
Already described

9. Lines 9-10. As stated above I disagree it is reasonable to extrapolate these results to more severe GAS infections. This should be discussed in a more balanced way.

Answer
Answered earlier on

10. Discussion lines 12-32 – Nicely made points.

11. However, lines 25-32 could also be an explanation of why clindamycin did not work in this trial: because it was started too late.

Clindamycin may not have worked because it was given too late but we could not find evidence to support this statement.

### VERSION 2 – REVIEW

| REVIEWER               | A/Prof Joshua Davis  
|------------------------|----------------------|
|                        | Menzies School of Health Research  
|                        | Australia              |
| REVIEW RETURNED        | 24-Nov-2016          |

| GENERAL COMMENTS | All my key concerns/queries have been addressed by the authors |
Adjunctive clindamycin for cellulitis: a clinical trial comparing flucloxacillin with or without clindamycin for the treatment of limb cellulitis

Richard Brindle, O Martin Williams, Paul Davies, Tim Harris, Heather Jarman, Alastair D Hay and Peter Featherstone

_BMJ Open_ 2017 7:
doi: 10.1136/bmjopen-2016-013260

Updated information and services can be found at:
http://bmjopen.bmj.com/content/7/3/e013260

These include:

**References**
This article cites 25 articles, 5 of which you can access for free at:
http://bmjopen.bmj.com/content/7/3/e013260#BIBL

**Open Access**
This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See:
http://creativecommons.org/licenses/by/4.0/

**Email alerting service**
Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

**Topic Collections**
Articles on similar topics can be found in the following collections

- Dermatology (60)
- Infectious diseases (557)

**Notes**

To request permissions go to:
http://group.bmj.com/group/rights-licensing/permissions

To order reprints go to:
http://journals.bmj.com/cgi/reprintform

To subscribe to BMJ go to:
http://group.bmj.com/subscribe/