PEER REVIEW HISTORY

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

TITLE (PROVISIONAL)	KANUKA HONEY VERSUS ACICLOVIR FOR THE TOPICAL TREATMENT OF HERPES SIMPLEX LABIALIS: A RANDOMISED CONTROLLED TRIAL
AUTHORS	Semprini, Alex; Singer, Joseph; Braithwaite, Irene; Shortt, Nick; Thayabaran, Darmiga; McConnell, Melanie; Weatherall, Mark; Beasley, Richard

VERSION 1 - REVIEW

REVIEWER	Brent A. Bauer MD Mayo Clinic USA
REVIEW RETURNED	11-Sep-2018

GENERAL COMMENTS	Overall, a straightforward report of a novel treatment approach (and research strategy) that has broad applications. I suspect the utilization of pharmacists as enrollers will generate some discussion as it is both novel and likely cost-effective. Replication to other studies would be enhanced if there could be some deeper documentation (? appendix) that describes this novel approach in greater detail. If it has been detailed in prior publications, linking to those would be helpful.

REVIEWER	Juraj Majtan Institute of Molecular Biology, Slovak Academy of Sciences
REVIEW RETURNED	12-Sep-2018

GENERAL COMMENTS	Submitted manuscript (original article) entitiled “KANUKA HONEY VERSUS ACICLOVIR FOR THE TOPICAL TREATMENT OF HERPES SIMPLEX LABIALIS: A RANDOMISED CONTROLLED TRIAL” has focused on characterisation of kanuka honey clinical efficacy for treatment of herpes simplex infection. Study was conducted as a randomised clinical trial using acyclovir as a standard therapeutic against herpes simplex infection. Overall, authors showed that aciclovir is not superior to honey treatment considering the time to return to normal skin (stage 7). Furthermore, no differences between both therapeutic groups were found in term of pain duration and treatment accessibility. I appreciate high number of subjects participated in the clinical trial (in total 952 adults). However, there is a small previous study (16 subjects), where honey (dark multifloral honey from UAE) was significantly superior to acyclovir regarding time of healing. Authors might be disappointed from the results of robust clinical trial where they

	actually did not prove outcomes from a small clinical study published by Al-Walli in 2004. Comments to authors: 1. Protocol for this clinical trial has already been reviewed and published in BMJ open. This protocol is very well conducted and methodology is at a high standard. Is the product (Honevo) based on kanuka honey standardised? We know that honey and its composition and efficacy can vary from year to year. Is there any internal control of in vitro efficacy of used kanuka honey? 2. Authors did not sufficiently discuss the differences between previous study (Al-Walli, 2004) and present study. Is it possible that other type of honey can be more efficient in treatment of HSV than used kanuka honey? 3. Page 15, line 20-27 vs. line 30-37. On one hand, authors indicated the possible aciclovir resistance and loss of its efficacy and on other hand (line 30-37) they have priority to test the combination of these two therapeutics. Do authors have any idea which compounds in honey are responsible for antiviral effects? It might be useful to combine kanuka honey with some other natural and well characterised antiviral compounds in order to strengthen the antiviral efficacy of honey.

REVIEWER	Seungwon Shin Kyung Hee University, Republic of Korea
REVIEW RETURNED	21-Nov-2018

GENERAL COMMENTS	This review has been solely focused on the statistical methods and analyses used in the primary study, as requested by the editorial office. 1. The author used the expression, or single blind, in the section of trial design, while the different expression, or open-label, was used in the attached protocol. I guess that the authors wanted to emphasize that the biostatistician or the center coordinator was blinded to allocation. However, the words, or single blind, are generally used for study participants or practitioners (investigators). Needs to clarify the meaning of 'single' blind or just address open-label as protocol. 2. The authors said that the data was analyzed with the set of the intention to treat (ITT). Generally, ITT should be pre-defined for the study. However, both the attached protocol and the manuscript are saying ITT principle without specific definition of it. In this case, ITT should be GENERALLY and STRICTLY interpreted as 'once randomized, always analyzed,' I guess. In this study, 952 patients were randomized, while 852 patients were included in the analysis. Please explain this discrepancy in the manuscript. (I can see that some of the patients were dropped out. Then, why were those data excluded from ITT analysis?) 3. Furthermore, the data for 'Time from stage 4 to stage 7' is addressing with N=840, not even N=852. Please explain this discrepancy in the manuscript. 4. Also, the authors planned both ITT and PP analysis in the protocol, however, I can not find out any comments on PP analysis in the manuscript. 5. Please clarify the calculation process of sample size, including the equation (or its' reference) or the used package, null hypothesis, the references or explanation of the effect size (a five-

	day median time and one-day median difference in favor of honey). 6. Authors did not statistically compare baseline characteristics between Aiclovir and honey groups before the outcome analysis? If not, please address why they did not and whether those baseline characteristics would be any confounding factors or not in the outcome analysis.

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1

Overall, a straightforward report of a novel treatment approach (and research strategy) that has broad applications. I suspect the utilization of pharmacists as enrollers will generate some discussion as it is both novel and likely cost-effective. Replication to other studies would be enhanced if there could be some deeper documentation (? appendix) that describes this novel approach in greater detail. If it has been detailed in prior publications, linking to those would be helpful.

Thank you. We agree that the novel methodology and trials network that has been established offers a unique capacity to conduct research in an under evidenced sector and we have three additional RCTs this year. As such we are working on a specific publication detailing the network's evolution to the current fully digital interface. We feel that a manuscript in it's on right will provide a more in depth and accessible description for best use in future study considerations.

Reviewer: 2

Submitted manuscript (original article) entitled "KANUKA HONEY VERSUS ACICLOVIR FOR THE TOPICAL TREATMENT OF HERPES SIMPLEX LABIALIS: A RANDOMISED CONTROLLED TRIAL" has focused on characterisation of kanuka honey clinical efficacy for treatment of herpes simplex infection. Study was conducted as a randomised clinical trial using acyclovir as a standard therapeutic against herpes simplex infection.

Overall, authors showed that aciclovir is not superior to honey treatment considering the time to return to normal skin (stage 7). Furthermore, no differences between both therapeutic groups were found in term of pain duration and treatment accessibility. I appreciate high number of subjects participated in the clinical trial (in total 952 adults).

However, there is a small previous study (16 subjects), where honey (dark multifloral honey from UAE) was significantly superior to acyclovir regarding time of healing. Authors might be disappointed from the results of robust clinical trial where they actually did not prove outcomes from a small clinical study published by Al-Walli in 2004.

Comments to authors:

1. Protocol for this clinical trial has already been reviewed and published in BMJ open. This protocol is very well conducted and methodology is at a high standard. Is the product (Honevo) based on kanuka honey standardised? We know that honey and its composition and efficacy can vary from year to year. Is there any internal control of in vitro efficacy of used kanuka honey?

In terms of assessing and controlling variability from batch to batch and from season to season, the in vitro quality assurance measure used for Honevo, the kanuka honey formulation used in the study, is called Total Activity (TA). This is a laboratory test which determines the antibacterial activity of honey with the result being the % of phenol that it is equivalent to. It is the standard in vitro test used to determine the antibacterial strength and quality of honey. For Honevo, the minimum TA of batches must be 20 ie. equivalent to 20% phenol.

In addition, Honevo is tested in vitro with proprietary chemical "fingerprinting" technology to show that the honey is in fact kanuka. Honevo is filtered to 50 microns which removes all visible and most invisible impurities, including most of the pollen, and then fast thermalized (quickly heated then cooled). These 2 steps, ultrafiltration and thermalization are what makes the honey pharmaceutical-grade (also known as medical-grade), which is safe to apply to the skin. These processes further standardize the product and reduce the inter-batch variability. Hydroxymethylfurfural levels are tested to ensure they are <40 mg/kg. This test shows that the honey has not been excessively heated, which theoretically could reduce the efficacy by denaturing important components.

2. Authors did not sufficiently discuss the differences between previous study (Al-Walli, 2004) and present study. Is it possible that other type of honey can be more efficient in treatment of HSV than used kanuka honey?

Thank you for this suggestion. We have added a discussion around the potential differences in outcome to include the two honey varieties used and potential bioactive composition leading to differences in efficacy. We have also commented on the differences in sample size and physician vs patient reported outcomes between studies.

3. Page 15, line 20-27 vs. line 30-37. On one hand, authors indicated the possible aciclovir resistance and loss of its efficacy and on other hand (line 30-37) they have priority to test the combination of these two therapeutics. Do authors have any idea which compounds in honey are responsible for antiviral effects? It might be useful to combine kanuka honey with some other natural and well characterised antiviral compounds in order to strengthen the antiviral efficacy of honey.

Thank you for this constructive comment. We have altered the sentence by suggesting that honey is an alternative therapeutic choice for HSL treatment by clarifying that it may be of use to particular groups including those susceptible to aciclovir resistance. We have expanded the introduction to describe the preclinical data for honeys' antiviral effects on HSV, rubella and influenza viruses and the currently unidentified anti-viral bioactives present in honey but promise of flavonoids. In addition, we adjusted the final discussion paragraph to include a call for investigation of honey in combination with aciclovir and other natural compounds with demonstrable antiviral effects, as detailed within Lin et al 2014.

Reviewer: 3
This review has been solely focused on the statistical methods and analyses used in the primary study, as requested by the editorial office.

1. The author used the expression, or single blind, in the section of trial design, while the different expression, or open-label, was used in the attached protocol. I guess that the authors wanted to emphasize that the biostatistician or the center coordinator was blinded to allocation. However, the words, or single blind, are generally used for study participants or practitioners (investigators). Needs to clarify the meaning of 'single' blind or just address open-label as protocol.

Thank you for highlighting this error, we have changed single blind to open label as per the protocol manuscript. We have kept the clarification around the pharmacists and participants being party to allocation vs the central investigators being unaware.

2. The authors said that the data was analyzed with the set of the intention to treat (ITT). Generally, ITT should be pre-defined for the study. However, both the attached protocol and the manuscript are saying ITT principle without specific definition of it. In this case, ITT should be GENERALLY and STRICTLY interpreted as 'once randomized, always analyzed,' I guess. In this study, 952 patients were randomized, while 852 patients were included in the analysis. Please explain this discrepancy in the manuscript. (I can see that some of the patients were dropped out. Then, why were those data excluded from ITT analysis?)

We apologise for omitting this definition from the protocol and hope the following explanation clarifies the figures. The 100 participants were defined as those that did not provide any time to event data and therefore unable to be included in the survival analysis of the primary outcome variable, however all participants that provided data were analysed according to allocation in line with ITT principles. This is detailed under the analysis box within figure 2.

3. Furthermore, the data for 'Time from stage 4 to stage 7' is addressing with N=840, not even N=852. Please explain this discrepancy in the manuscript.

The table detailing the survival analyses contains reference to both the censored (contributing data but no event at last protocolised observation) and uncensored participants (provided complete survival data). For time from stage 4 to 7 this is detailed as 'uncensored (censored)': Control 418 (3) and Honey 422 (9). Therefore the total analysis N for each arm is the sum of the censored and uncensored: (418 + 3) + (422 + 9) = 852.

4. Also, the authors planned both ITT and PP analysis in the protocol, however, I cannot find out any comments on PP analysis in the manuscript.

We thank the reviewer for identifying that we haven't been explicit about the pre-specified per-protocol analysis. In the event only five participants (1 in the Kanuka honey arm and 4 in the acyclovir arm) took the other treatment. With such a small number of protocol violations and the similar time to recovery for the two treatment arms, we felt a per-protocol analysis would add nothing to the assessment of possible bias of the ITT analysis. We have added this to the discussion.

5. Please clarify the calculation process of sample size, including the equation (or its' reference) or the used package, null hypothesis, the references or explanation of the effect size (a five-day median time and one-day median difference in favour of honey).

We thank the reviewer for this request for clarification. We had stated that SAS version 9.4 was used for the analyses. The link for the details of the SAS calculation is:

https://support.sas.com/documentation/cdl/en/statug/63962/HTML/default/viewer.htm#statug_power_a0000001003.htm

[accessed 28/1/19].

The details of the SAS code using SAS PROC POWER are:

```
proc power;

  twosamplesurvival test=gehan
```

groupmedsurvtimes = 5 | 4

accrualtime = 180

totaltime = 210

npergroup = .

power = 0.8 ;

run;

The output from this procedure is:

The SAS System 13:24 Friday, January 25, 2019 3

The POWER Procedure
Gehan Rank Test for Two Survival Curves

Fixed Scenario Elements

Method	Lakatos normal approximation
Form of Survival Curve 1	Exponential
Form of Survival Curve 2	Exponential
Accrual Time	180
Total Time	210
Group 1 Median Survival Time	5
Group 2 Median Survival Time	4
Nominal Power	0.8
Number of Sides	2
Number of Time Sub-Intervals	12
Group 1 Loss Exponential Hazard	0
Group 2 Loss Exponential Hazard	0
Alpha	0.05

Computed N Per Group

Actual Power	N Per Group
0.800	423

Please note that the date of output is the date of re-running the code and not the date of the original sample size calculation.

6. Authors did not statistically compare baseline characteristics between Aciclovir and honey groups before the outcome analysis? If not, please address why they did not and whether those baseline characteristics would be any confounding factors or not in the outcome analysis.

We thank the reviewer for this comment. In general we consider it is poor statistical practice to formally compare the distribution of baseline variables in a randomised trial for the following reasons:

the study is not powered to detect differences in baseline variables, it is almost always unclear what constitutes a clinically meaningful difference in baseline variables, and it adds to the experiment-wide type I error rate.

We agree with the reviewer that it is useful to include in sensitivity models important co-variates that otherwise predict the outcome of interest however we have been unable to identify a systematic review or other research that robustly reports prognostic variables for recovery time for herpes labialis. There is a body of literature that refers to risk factors for recurrence of herpes labialis (which is not the primary outcome variable of the clinical trial) which may conceivably also be risk factors for time to healing. These are older age (less likely to have recurrent herpes labialis) and menstruation (and by extension female sex), during which it is more likely to have recurrent herpes labialis.

Although not a pre-specified analysis we have run a sensitivity analysis including age and sex added to the primary outcome variable analysis. The HR (95% CI) for age (per decade older) was 0.96 (0.91 to 1.00), P=0.058; and for female (compared to male) sex; 0.87 (0.74 to 1.02), P=0.084. The (now) adjusted HR for Honey versus Control was more or less identical to the unadjusted estimate and confidence interval: 1.06 (0.92 to 1.22), P=0.41 [unadjusted 1.06 (0.92 to 1.22), P=0.56].

In the absence of robust literature to support possible confounding variables for time to healing of herpes labialis we would prefer not to add a post-hoc analysis to the main text, even though the estimates from an analysis adjusted for age and sex are no different, but this material will be accessible to readers as part of the open review process.

VERSION 2 – REVIEW

REVIEWER	Juraj Majtan Institute of Molecular Biology, Slovakia
REVIEW RETURNED	05-Feb-2019

GENERAL COMMENTS	Authors successfully answered all raised issues.

REVIEWER	Seungwon Shin Kyung Here University, Republic of Korea
REVIEW RETURNED	21-Feb-2019

GENERAL COMMENTS	This review has been focused on statistics, followed by the previous revision. Most issues in the previous review have been revised or reasonably explained by the authors. Still, I recommend the followings; 1. “This is detailed under the analysis box within figure 2.” → It would be better that the author explains this in the results part, too, for the readers. 2. About the the sample size calculation, I understand that the authors used SAS version 9.4. Still, the references or explanation of the adopted effect size (a five-day median time and one-day median difference in favour of honey) should be clarified in the method part, however.

VERSION 2 – AUTHOR RESPONSE

Many thanks for the opportunity to further respond and to the reviewers for recommending publication.

We actioned the requests of reviewer 3 as follows:

1. 1. “This is detailed under the analysis box within figure 2.” → It would be better that the author explains this in the results part, too, for the readers.

Within the manuscript we have adapted the following paragraph with the final two sentences.

‘The flow of participants is shown in figure 2. Four participants in the aciclovir group and one in the honey group, were dispensed the incorrect treatment. There were 91 participants lost to follow up (49 aciclovir, 42 honey) and 9 withdrew from the study due to adverse events (3 aciclovir, 6 honey). In the final intention to treat analysis 852 participants provided data. The 100 participants excluded were defined as those that did not provide any time to event data and therefore unable to be included in the survival analysis of the primary outcome variable. All participants that provided data were analysed according to allocation in line with ITT principles.’

2. About the sample size calculation, I understand that the authors used SAS version 9.4. Still, the references or explanation of the adopted effect size (a five-day median time and one-day median difference in favour of honey) should be clarified in the method part, however.

We have included the justification as follows within the sample size calculation paragraph which now reads:

‘The sample size calculation was based on previously reported Hazard Ratios of 1.23 and 1.24 for a median five day duration of symptoms which implies a one day median reduction to four days. [22] Using this assumption of a five-day median time to healing and clinically significant one-day median difference in favour of honey, with an associated HR of 1.25, 80% power and 5% type I error rate, a total of 423 participants were required per arm. 950 in total were to be randomised, to take in to account an assumed attrition rate of 10%.’