Author’s response to reviews

Title: Early screening for Chlamydia trachomatis in young women for primary prevention of pelvic inflammatory disease (i-Predict): study protocol for a randomized controlled trial

Authors:
Jeanne Tamarelle (jeanne.tamarelle@uvsq.fr; jeanne.tamarelle@gmail.com)
Anne Thiébaut (anne.thiebaut@inserm.fr)
Bénédicte Sabin (benedicte.sabin@gmail.com)
Cécile Bébéar (cecile.bebear@u-bordeaux.fr)
Philippe Judlin (p.judlin@chru-nancy.fr)
Arnaud Fauconnier (afauconnier@chi-poissy-st-germain.fr)
Delphine Rahib (delphine.rahib@santepubliquefrance.fr)
Layidé Méaude-Roufai (layide.meaude@aphp.fr)
Jacques Ravel (jravel@som.umaryland.edu)
Servas Morré (samorretravel@yahoo.co.uk)
Bertille de Barbeyrac (bertille.de-barbeyrac@u-bordeaux.fr)
Elisabeth Delarocque-Astagneau (elisabeth.delarocque-astagne@pasteur.fr)

Version: 1 Date: 01 Sep 2017

Author’s response to reviews:

Dear Editor,

We would like to thank the reviewers for their careful and constructive comments. All comments have been addressed, with corresponding changes made directly to the manuscript when appropriate. Changes are highlighted in yellow in the text.

Reviewer #1:

An interesting paper, given the lack of available information in relation to this condition. Generally the paper was well written however I would suggest a few clarifications and some language/grammatical changes:
Clarifications

• What process is followed following analysis of control samples at end of study? Will this group be dealt with as per the intervention arm or referred into routine care for treatment if required?

- All results for Ct will be made available to the gynaecologist at the final visit. The gynaecologist will review the results and decide how the patient should be cared for according to these results. A clarification was added line 255 and line 351.

• What instruments are used to collect data - are any of these standardised instruments?

- Data on sexual practices, gynaecological history, consumptions etc. are obtained from the participant through standardized self-administered questionnaires directly on the dedicated internet platform. These questionnaires were elaborated with a French sociologist expert in sexual health and validated in a pilot study. Clinical data are collected through questionnaires filled in by physicians and elaborated with two gynaecology experts. They are entered by a clinical research assistant on the dedicated internet platform. For self-administered questionnaires, the specification “self-administered” was added, and for clinical data on specific forms, it was specified that data entry is carried out by a trained clinical research assistant.

• How is the randomisation sequence generated - Is this by someone independent of the trial? Is the electronic platform used a central system or site specific?

- Randomization is performed by the electronic platform of the trial (cleanweb® module) in a site-specific manner, as described in the randomization paragraph. A clarification was brought to this paragraph.

• Immunogenetics are a secondary objective however the associated outcomes do not appear to be discussed in the Outcomes section.
- In the outcomes section, only immediate endpoints are presented (number of cases, incidence, prevalence). The results of the immunogenetic analyses will be generated after a specific analysis described briefly in the statistical analyses section, at the end of the study. The related endpoints are cases of Ct infection and cases of PID, since these will be compared to non-Ct cases and non-PID cases. Prevalence and incidence of Ct infection are already presented as endpoints, we added incidence of PID as an endpoint since it will be necessary for immunogenetic analyses.

- Are samples sent straight to -80C or frozen at another temperature first before longer term storage at -80C (Materials Section).

- They are stored directly at -80 degrees. A clarification was brought line 416.

- Secondary analyses require clarification - how will time to develop PID/duration of infection be analysed?

- Time to develop PID is defined as the time between incident Ct infection and incident PID. PID is evaluated continuously during the whole follow-up since participants are encouraged to visit participating GED, we will therefore have the precise date. For Ct infection, since samples are taken every six months, the date of infection is considered to be the middle of the period. A clarification was brought to this paragraph.

- Discussion section should consider the generalisability of sample. I would suggest this is considered under practical/operational issues where the population is discussed.

- A paragraph was added as suggested in the practical/operational issues of the discussion to discuss the issue related to generalizability and how we addressed these issues.

- Inclusion of patient compensation is noted in the discussion. This would better located within the methods section.
- A sentence was added in the methods section (line 360) after the final visit paragraph to make this information known to the reader in a more adequate section.

Language/Grammar
• Tense should be reviewed throughout to convert present tense to future tense (e.g. 'is performed' -Line 252 - should be 'will be performed').

- Corrections were made where necessary.

• Throughout there is reference that the control samples will be 'differed' (e.g. Line 383). This is a different word to that which you intend. The manuscript should be reviewed and 'differed' (unlike or dissimilar) amended to 'deferred' (action to be completed at a later time, delayed).

- There were four occurrences of the word “differed”. It has been replaced by “deferred” at each occurrence.

Reviewer #2:
• The i-Share study is first referred in page 9 (design/setting) and comes a bit as a surprise. It is then later described succinctly in page 11. Please move to background to give context to the reader, and add further details (e.g. what is the overall aim of the i-Share cohort study?).

- Most of the paragraph page 11 was moved to the end of the introduction section, and details were added on the objective of the i-Share cohort.

• Please add section "Status of trial", including when it was open to recruitment, expected end of recruitment and timing for primary analysis.

- A section “status of the trial” was added after the discussion with the required information.

• The control arm is sometimes referred to as "current guidelines" of screening. This can be a bit misleading, as regular samples are being taken on these patients; despite the analysis of them
being deferred, the control arm could not be considered "the current screening guideline". This could be inferred in pg. 14 Primary outcome and from the discussion in page 19. A group following 'current recommendations' would not be taking samples every 6 months; they would undoubtedly be more aware of the issue. Undoubtedly, the challenge with the design is the choice of control group and the ethical challenge of deferring the testing - the team explain the arguments in favour in the discussion, but maybe could elaborate as to whether this was discussed and challenged in the process of set-up and approval of the trial.

- Indeed, the control group is not a perfect “following current screening guidelines” group since participants of this group are given additional information and are more likely to attend an STI clinic than female student not participating in the study. They are thus more likely to follow the current recommendations than women of the same age not participating to the trial. Despite a very careful review of the protocol by an ethical board including physicians, nurses and patients, this design did not jeopardize the approval. The only ethical issue would have been to test these samples without giving the result of the test to the participant, but it was considered acceptable to defer testing.

The choice of collecting samples in this control group is now justified in the discussion as suggested, as well as the issues resulting from this choice in terms of ethics and statistical power.

• Secondary objectives, page 9: 2nd bullet point text between brackets is a bit unclear "at the beginning and end of infection, throughout the course of infection" -- what does this refer to? progression to PID relative to Ct infection start? Not clear.

- Phrasing in bracket was indeed adding confusion, the part between brackets was removed.

• How is Ct clearance defined? (page 9) This does not appear in Endpoints (page 15).

- The definition of Ct clearance was added page 17 in the secondary endpoints.

• Page 10, exclusion criteria: the POPI trial excluded patients that had been tested for Chlamydia in the past 3 months, was this considered in this trial? Is pregnancy the only exclusion criteria?
- Excluding patients who had been tested for Ct in the past three months was discussed but not adopted since it would change the estimation of prevalence. Besides, this information is collected in the self-administered questionnaire and we will be able to discuss this fact and take it into account in our analysis and interpretation. Known pregnancy is the only exclusion criteria.

- One sentence was added in the paragraph describing the control group to explain why sampling at the same frequency but without testing is necessary.

Trial assessments in pages 10 to 14 could be organised in a sequential order for better flow, i.e. please explain first Recruitment, treatment allocation (with details of randomisation), baseline assessments and follow-up (as in Figure 2), with the following considerations:

- How patients on the iShare cohort study are contacted? I.e. search on patients fulfilling inclusion criteria are contacted directly through the system? Is this what is intended with the description of the information retrieved from the system (e.g. type of training, registration etc. at the top of page 12)? It is unclear if that is used for screening patients or as baseline information for the patients consenting to the trial. I understand the trial also accepts volunteers from i-Share or UHS who would proactively would contact the team to participate in the study.

- Female students can be contacted or informed about i-Predict through the i-Share cohort, but they can also be informed and asked to participate to i-Predict without knowing of i-Share (e.g through university health services). They will then have to participate to both, since we need data from the i-Share baseline questionnaire (type of training, registration etc.). Some details on how students are recruited (through i-Share or not) were added. To clarify what use is made of i-Share data (type of training, registration etc.), this sentence was moved to the initial visit paragraph.
• The initial visit does not necessarily need to occur after randomisation - would be even better to occur before, after consent and collection of M0 sample and completion of questionnaires.

- Randomization occurs during the initial visit, since group assignment is generated at the same time as participants’ ID number on the Cleanweb platform. It occurs just after consent, but before M0 sample and questionnaire. For clarification, it was moved after the initial visit paragraph.

• Randomisation section: Last sentence (rando after recruitment, obvious) and aim of it unclear. This section needs to explain how the allocation sequence is concealed (I understand the web system generates new codes and access to any pre-specified list is restricted to personnel not directly related to contact with participants). I think the correct term is "block randomisation" or "randomisation by blocks", not "blocked".

- Last sentence was removed and replaced by a sentence explaining who has access to group assignment. “Blocked” was corrected and replaced by “block”.

• Follow-up: it is unclear how data from clinical episodes will be collected: when Ct is diagnosed in the intervention group, patient is referred to their physician or nearest STI, and the physician upon examination would fill-out a specific form on clinical findings. So, do you contact with (any) physician or STI clinic the patient decides to go? Or she is referred to the trial sites only and they know about trial procedures?

- It can be any physician or STI clinic. The participant brings the clinical form that needs to be completed by the physician, with a prepaid envelope. If she forgets, she can give the name and contact of the physician she went to through the cleanweb platform or directly to the UHS, in order for the clinical research assistant to contact the physician and retrieve the information needed. Some details about this procedure were added in the follow-up section.
• Same for when patient refers pelvic pain and is contacted to attend a GED, the contact to physician to fill in the information is performed?

- Physicians in the GED are informed of the existence of i-Predict (several information sessions on the study and the data collected. We checked that the data we need are systematically recorded in the medical files), but they don’t have to fill in an additional specific questionnaire. The clinical research assistants have been granted access to medical files to retrieve the data. They then complete a standardized clinical research form, and enter the data on Cleanweb. Some details about this procedure were added in the follow-up section.

• Page 14 Outcomes: lines 344-346 could go in background section. If there is controversy around definition of PID, you could say "For the purposes of this trial, PID will be defined as follows".

- The definition used is not really controverted, it is the definition given by the national college of gynaecologists and obstetricians, but as it is mainly clinical, the definition used in other papers differs by a bit. The phrasing suggested by the reviewer was adopted.

• Page 15 Secondary endpoints:
  ○ Prevalence is listed as 2ary objective but not as 2ary endpoint. Please check objectives and endpoints are aligned.

- Prevalence of Ct infection is indeed an important endpoint and has been added.

  ○ Can you clearly define duration of Ct infection, proportion of infections associated with diagnosis of PID (is it Ct diagnosed at PID diagnosed?), time to develop PID (since baseline, Ct infection?), spontaneous Ct infection resolution (e.g. Ct positive at 6 months, but Ct Negative at 12 months? Is that even possible? How do you investigate this, can it be patient taken antibiotics for another reason? Error of Ct diagnostic test?).
- Details were added in these secondary endpoints for clarification. Regarding antibiotics taken for another reason, this information is asked in baseline and follow-up questionnaires, as well as additional testing outside the study, as explained in the follow-up section, and will be taken into account (a first analysis on all samples from control group, then restricted to participants who didn’t take antibiotics for example).

Concerning Ct testing, the methods used by the CNR are highly sensitive so this possibility has been ruled out. For Ct-positive samples, they are re-run for confirmation.

○ Calculation of re-incidence refers to ref. 21 in page 15, and to refs. 25-26 in page 16: do they refer to the same procedure? If so make it clear or add ref. 21 to page 16; so it's not mistaken by 2 different methods.

- Ref 25-26 are for the typing method. Ref 21 is for the algorithm to discriminate between treatment failure, reinfection or persistent infection; this algorithm uses typing information (see ref 25-26) but also other exposure data. A sentence in brackets was added to make it clearer.

• Statistical analysis:
○ I could not replicate exactly the sample size: could you please provide further details?

- The table summarizing the hypothesis tested to calculate sample size was added as supplementary material for reviewers’s use only. We rounded the resulting number to 4000, so this could be the source of the confusion.

○ The primary endpoint is cumulative incidence (at 18-24 months/last assessment, binary), and the primary analysis will be on an Intention to Treat basis. How will you consider dropouts without PID diagnose (i.e. pts dropping out before final assessment): assume no PID?

- We don’t expect much differences in terms of dropouts between the two groups, therefore it should not impact the primary analysis. However, we plan to do a sensitivity analysis to evaluate the impact of dropouts on our results, including assuming no PID for all dropouts.
○ If using Cumulative incidence, a Cox model is not adequate (rather it should be logistic regression, as primary endpoint is PID yes/no. Cox model would be adequate to model time to PID.

- Primary endpoint is indeed binary, and not time-varying. Cumulative incidence will be used to compare both arms. We will obtain a relative risk. But for further analysis, we use a cox regression model for two reasons: PID is evaluated continuously during the trial (we will have the precise date), and we want to explore the role of some explanatory variables that can have different values at M0 M6 M12 and M18 (Ct infection, sexual practices etc.). Therefore, a logistic regression model can’t be used.

• Section Stakeholders’ contributions could go within the text, in a section Trial Governance. Please remove contact details of the monitor of the study, not needed for the manuscript -rather explain the level and extent of monitoring for this trial. Remove "Independent experts" from this section as this has been already explained in the text.

- A section “Trial governance” was added in Methods.

• Order of figures do not correspond to legends and reference in text, please amend.

- Figure 1 is the SPIRIT figure and corresponds to the design of the trial; it is required by the journal and must be included in the manuscript directly, whereas figure 2 and 3 must be submitted as separate files. Maybe this is the source of the confusion?