PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Impact of legal status change on undocumented migrants' health and well-being (Parchemins): protocol of a 4-year prospective mixed-methods study.
AUTHORS	Jackson, Yves; Courvoisier, Delphine; Duvoisin, Aline; Ferro-Luzzi, Giovanni; Bodenmann, Patrick; Chauvin, Pierre; Guessous, Idris; Wolf, Hans; Cullati, Stéphane; Burton-Jeangros, Claudine

VERSION 1 - REVIEW

REVIEWER	Hannah Bradby Uppsala University Sweden
REVIEW RETURNED	29-Dec-2018

GENERAL COMMENTS	The very particular circumstances of the reforms that this study proposes to examine make this a potentially fascinating 'natural experiment'. Some further details could be given on what will constitute an adequate 'ability to communicate' for inclusion in the study, since 'communication' could cover a very basic level of ability. (page 8, line 15) Is it appropriate to allocate those who are refused legal status to the control group? (page 8, line 28). Might those who wanted but were refused legal status be rather different from those who did not want to apply in the first place? Are there any implications of getting legal status for healthcare for pregnancy and birth? So, is it equally possible to have ante-natal and post-natal care without and with legal status? If not, then does there need to be special attention paid to women (and perhaps men too?) of child-bearing age? (page 9) How is the likely drop out rate calculated and is it high enough? (page 9 ,line 42) Minor matters Remove UM as an unnecessary and unattractive acronym that does not aid comprehension. page 3, line 38 restrict rather than restricts page 3, line 51, impact rather than impacts

	page 5, line 60, hypothesise rather than hypothesis page 9, line 43 - secure, rather than a secured server

REVIEWER	Päivikki Koponen National Institute for Health and Welfare, Finland
REVIEW RETURNED	31-Dec-2018

GENERAL COMMENTS	The study is unique and focuses on important and timely topics for the health and wellbeing of undocumented migrants. The manuscript shows that the study has a well-defined purpose and background. However, some key facts on the protocol need more details and specification, while some parts of the manuscript could be abbreviated. The abstract is quite long but lacks some key details on the protocols: target group, age and selection/eligibility criteria for study and control groups, as well as sample size. • The dates of the study are not reported: when will the data collection begin (or has begun) and when it's planned to be finalized? • Some details on the target group are missing. Based on the inclusion criteria the study will include migrants aged 18 and over, who have stayed in the country for at least 3 years and who plan to stay at least 3 years, have no criminal record and are not registered as asylum seekers. How will the researchers be able to check the criminal records and asylum seeker status? It's not clear if persons with a previous asylum seeker status are included. The target group should be better defined in the abstract and all parts of the manuscript. How will the control group be selected? • Parts of the manuscript refer to children (eg. under the Introduction, as statistics and previous literature). In my opinion it would be logical to focus on adults throughout the manuscript. • The statistics presented in the introduction are quite old: the situation in Europe changed a lot between 2014-2016 and this should be taken into account. Also several references are quite old and the authors should update the manuscript with newer studies and literature. As the situation and regularization policies are so different in USA than in Europe, it might be better to focus on the situation and policies in European countries. The introduction, theoretical perspectives and references to previous literature could be summarized (now over 3 pages, could be max 2 pages) to focus more on the protocols of the Parchemins study. • The description of theoretical perspectives could be abbreviated and specified in a figure of the conceptual framework of the study to give better background for the primary and secondary outcome measures and additional variables. Now it's not clear why satisfaction with current life is the secondary outcome while several items on quality of life are listed under the additional variables. It's not clear if the Table 1 presents only these additional variables. It would also be more informative to include references to chosen standard instruments in Table 1 to link the variables to the instruments. • The validity of the instruments in different cultures of the potential participants should be addressed. It is not clear if the questionnaires in the quantitative data collection will be translated to the languages mentioned under exclusion criteria (languages spoken by investigators). The authors should foresee how will the limitations in used languages, and different translations, as well

	cultural issues, affect the data collection and the results. This will also have an impact on the study ethics if it will also be possible to give the informed consent in different languages and if the information material will be available in different languages. It is not clear, if all questionnaires will be administered in face-to-face interviews or which questionnaires/instruments will be self-administered. • Under “Intervention” it’s mentioned that “in the unlikely occurrence of refusal, participants will be allocated to the control group”. What does this mean? Why does the study group assume that refusals will be unlikely? Is it ethically acceptable to allocate refused persons to the control group? How is data collected from the control group? Will any information on refusals be recorded? How will the effects of non-response be taken into account? • The sample size seems to be quite low to ensure diversity of the sample regarding origins, age, gender etc., especially if the plan is to adjust for baseline characteristics as mentioned under statistical analysis. Can the authors specify how this will be addressed in the analysis? How will eligible persons be invited to both qualitative and quantitative interviews or are these groups different? What is the sample size for the control group? Is the data collection similar for the study group and the control group? The study flow could also be easier to follow in a figure format. • Material incentives are mentioned under retention, but these are also an ethical issue, which should be specified in the manuscript. What kind of incentives will be given? Do the participants get any feedback or will they be referred to get help if needed? • If the participants may withdraw from the study for any reason at any time, how all data will be included until the last participants’ questionnaire? Will answers given at earlier stages be kept in the data set even though the participant will want to withdraw his/her consent? • The interviews have been pretested, so some information could be available on the planned length of each interview. Based on the listed instruments and variables in Table 1, the interviews might take quite a long time: is this feasible? • I’m not a native English speaker, but in my opinion the manuscript could be improved by language checking.

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1

The very particular circumstances of the reforms that this study proposes to examine make this a potentially fascinating 'natural experiment'.

1. Some further details could be given on what will constitute an adequate 'ability to communicate' for inclusion in the study, since 'communication' could cover a very basic level of ability. (page 8, line 15).

We mean holding a basic conversation and clarified this in the text (page 7, line 208)

2. Is it appropriate to allocate those who are refused legal status to the control group? (page 8, line 28). Might those who wanted but were refused legal status be rather different from those who did not want to apply in the first place?

Thank you for this interesting comment, which refers to the criteria of inclusion and/or exclusion related to the case and control groups and highlights a mistake we made. In fact, the actual regularization policy means that only undocumented migrants who meet all the eligibility criteria can apply for a residency permit and that the permit will be systematically provided. Therefore, there won't be rejected claims. We clarified this on page 6, line 208. The control group is made up of migrants lacking one or more eligibility criteria or those who prefer not to engage into the administrative process of regularization.

3. Are there any implications of getting legal status for healthcare for pregnancy and birth? So, is it equally possible to have ante-natal and post-natal care without and with legal status? If not, then does there need to be special attention paid to women (and perhaps men too?) of child-bearing age? (page 9)

Universal access to prenatal and perinatal care is provided by the University Hospitals of Geneva irrespective of the legal status, and receiving pre-, peri- or postnatal care won't change depending on the legal status of women. Therefore, we do not foresee major differences between groups.

4. How is the likely drop out rate calculated and is it high enough? (page 9 ,line 42)

We expect drop-out mainly in the control group, i.e. among migrants living in the absence of stable legal status. Yet, we purposely selected "stable" undocumented migrants who had been living in Geneva for at least 3 years, therefore less at risk of leaving the country soon. We empirically calculated the rate at roughly 10%.

Minor matters

- Remove UM as an unnecessary and unattractive acronym that does not aid comprehension.
- page 3, line 38 restrict rather than restricts
- page 3, line 51, impact rather than impacts
- page 5, line 60, hypothesise rather than hypothesis
- page 9, line 43 - secure, rather than a secured server

Changes have been made

Reviewer: 2

1. The study is unique and focuses on important and timely topics for the health and wellbeing of undocumented migrants. The manuscript shows that the study has a well-defined purpose and background. However, some key facts on the protocol need more details and specification, while some parts of the manuscript could be abbreviated. The abstract is quite long but lacks some key

details on the protocols: target group, age and selection/eligibility criteria for study and control groups, as well as sample size.

We added more details as requested and, at the same time, reduced a bit the length of the abstract.

2. The dates of the study are not reported: when will the data collection begin (or has begun) and when it's planned to be finalized?

We added this information in the method section (page 6, line 188)

3. Some details on the target group are missing. Based on the inclusion criteria the study will include migrants aged 18 and over, who have stayed in the country for at least 3 years and who plan to stay at least 3 years, have no criminal record and are not registered as asylum seekers. How will the researchers be able to check the criminal records and asylum seeker status? It's not clear if persons with a previous asylum seeker status are included. The target group should be better defined in the abstract and all parts of the manuscript. How will the control group be selected?

Thank you for this important remark. In Geneva, the sociodemographic and administrative profile of asylum seekers and undocumented migrants differ widely. No asylum seeker will be included in the present study. We did plan and conduct this study with a network of NGOs acting as gatekeepers into the regularization process. Information about previous asylum claim was thus accessible and we review the full migration history during the first wave of quantitative data collection. We specified in the methods section that migrants who registered as asylum seekers were not eligible for inclusion and added this information in the abstract (page 1, line 34 and page 6, line 203). Regarding criminal records, this criterion is also assessed by NGOs. While it is a criteria of ineligibility for regularization, it does not constitute a criteria of ineligibility for the study in the control group. We have better clarified this point in the text.

4. Parts of the manuscript refer to children (eg. under the Introduction, as statistics and previous literature). In my opinion it would be logical to focus on adults throughout the manuscript.

We reduced the number of references to children in the text. We opted yet to keep the mention about the DACA policy in the US as it is the largest source of evidence on the effect of regularization on undocumented young migrants' health.

5. The statistics presented in the introduction are quite old: the situation in Europe changed a lot between 2014-2016 and this should be taken into account. Also several references are quite old and the authors should update the manuscript with newer studies and literature. As the situation and regularization policies are so different in USA than in Europe, it might be better to focus on the situation and policies in European countries.

We agree some of our bibliographic references date from some years now. This is a key point in regards to the documentation of irregular migration. Most recent data (for instance in the Eurostat database or the 2019 WHO report on migrants health) refer to the arrival of migrants with uncertain legal status (asylum seekers vs undocumented) yet, we could not find any recent and valid count of undocumented residents despite a thorough literature search. For example, the International Organization for Migration portal (https://migrationdataportal.org/?i=stock_abs_&t=2017) does not include any data on irregular migrants. We therefore took care to use older data that have been

generally accepted. As undocumented migration affects most Western countries and represent a comparable policy challenge, we opted to also include North America in our discussion while keeping in mind the political and administrative differences.

6. The introduction, theoretical perspectives and references to previous literature could be summarized (now over 3 pages, could be max 2 pages) to focus more on the protocols of the Parchemins study.

Considering the multidisciplinary nature of the study and the different fields under investigation, it seemed important to provide a comprehensive background. To respond to the query, we did our best to shorten the text.

7. The description of theoretical perspectives could be abbreviated and specified in a figure of the conceptual framework of the study to give better background for the primary and secondary outcome measures and additional variables. Now it's not clear why satisfaction with current life is the secondary outcome while several items on quality of life are listed under the additional variables. It's not clear if the Table 1 presents only these additional variables. It would also be more informative to include references to chosen standard instruments in Table 1 to link the variables to the instruments.

The theoretical perspectives have been shortened. The main outcomes have been removed from Table 1 and it was clarified in the text that Table 1 only presents the additional variables. References have been added in the table.

8. The validity of the instruments in different cultures of the potential participants should be addressed. It is not clear if the questionnaires in the quantitative data collection will be translated to the languages mentioned under exclusion criteria (languages spoken by investigators). The authors should foresee how will the limitations in used languages, and different translations, as well cultural issues, affect the data collection and the results. This will also have an impact on the study ethics if it will also be possible to give the informed consent in different languages and if the information material will be available in different languages. It is not clear, if all questionnaires will be administered in face-to-face interviews or which questionnaires/instruments will be self-administered.

The Questionnaire and study documentation are presented face-to-face and were translated into the four most frequently spoken languages among undocumented migrants in Geneva. The survey and informed consents are available in all 4 languages. We took care to use scales validated in the four languages to reduce cultural issues. We will also look at results by culture/language in order to see whether there are differences in scale use (using differential item functioning analysis – DIF).

9. Under "Intervention" it's mentioned that "in the unlikely occurrence of refusal, participants will be allocated to the control group". What does this mean? Why does the study group assume that refusals will be unlikely? Is it ethically acceptable to allocate refused persons to the control group? How is data collected from the control group? Will any information on refusals be recorded? How will the effects of non-response be taken into account?

Our phrasing was confusing and we deleted this sentence. The current pilot regularization policy stipulates that in presence of all criteria of eligibility, the Government cannot refuse to grant a residency permit. We thought we should specify that in the unlikely occurrence of the refusal of a valid application in one of our participants, we would retain this participant in the study but allocate it to the

control (= non regularized) group. We specified this in the manuscript (page 7, inclusion criteria section).

10. The sample size seems to be quite low to ensure diversity of the sample regarding origins, age, gender etc., especially if the plan is to adjust for baseline characteristics as mentioned under statistical analysis. Can the authors specify how this will be addressed in the analysis? How will eligible persons be invited to both qualitative and quantitative interviews or are these groups different? What is the sample size for the control group? Is the data collection similar for the study group and the control group? The study flow could also be easier to follow in a figure format.

Regarding the sample size, we calculated it according to the usual guidelines taking into consideration the expected effect size of the intervention on the main variable. We agree that we do not have the sample size to specifically look at the effect of the intervention in specific subgroups. Unfortunately, though it would be nice to have a larger sample size, feasibility is also an important limitation since this programme only exists in Geneva and the total number of eligible person for the programme is limited. In Geneva, the population of undocumented migrants predominantly includes women of child-bearing age from Latin America and the Philippines. We planned our recruitment strategy to target these groups in priority in order to enhance generalizability as well as possible, We added a study flow diagram to clarify that the quantitative data collection is similar in both groups and that the qualitative sub-sample participants belong to the same study population.

11. Material incentives are mentioned under retention, but these are also an ethical issue, which should be specified in the manuscript. What kind of incentives will be given? Do the participants get any feedback or will they be referred to get help if needed?

We clarified these points. Participants are compensated with a 15 euros worth material present for the time spent with investigators. In addition, information about the Papyrus process and about healthcare resources are provided.

12. If the participants may withdraw from the study for any reason at any time, how all data will be included until the last participants' questionnaire? Will answers given at earlier stages be kept in the data set even though the participant will want to withdraw his/her consent?

According to Swiss law, data can be kept except if the participant explicitly request that all data be deleted. This latter request is very rare. Thus, all data collected until withdrawal of study or study end will be kept.

13. The interviews have been pretested, so some information could be available on the planned length of each interview. Based on the listed instruments and variables in Table 1, the interviews might take quite a long time: is this feasible?

Interviews last between 45 and 90 minutes which is both feasible and acceptable by participants.

14. I'm not a native English speaker, but in my opinion the manuscript could be improved by language checking.

We did our best to improve the English

VERSION 2 – REVIEW

REVIEWER	Päivikki Koponen National Institute for Health and Welfare, Finland
REVIEW RETURNED	01-Mar-2019

GENERAL COMMENTS	The authors have well addressed the comments from both reviewers. I have only one concern: The references added to Table 1 include the Swiss surveys, but are there any other instruments / items in these surveys which originate from other international validated Instruments than those marked with references 69-72? In their response to item 8 of my previous review the authors state that thay have used scales validated in the four languages: have all these four languages been used in the Swiss surveys and have all translations for all questions been validated in this contex? Should the limited language skills of the participants be mentioned as one potential limitation for this study? In the follow-up period the language skills of the participants may inprove, could this also have an impact on the results? There is an unnecessary "and" in the last sentence on page1 (line 60). In the Acknowledgements there is a spelling mistake in developping (developing).

VERSION 2 – AUTHOR RESPONSE

Reviewer 2

1. The references added to Table 1 include the Swiss surveys, but are there any other instruments / items in these surveys which originate from other international validated Instruments than those marked with references 69-72? In their response to item 8 of my previous review the authors state that they have used scales validated in the four languages: have all these four languages been used in the Swiss surveys and have all translations for all questions been validated in this context?

References 69 to 72 are indeed the main international validated instruments we use in the section dedicated to health in our questionnaire.

The Swiss Health Survey questionnaire exists in German, French and Italian, the Swiss Household Panel Study questionnaire exists in German, French, Italian and English. The Parchemins study questionnaire exists in another set of languages, chosen to match the main groups of undocumented migrants in Geneva (French, English, Portuguese, Spanish). In those different contexts and over the years for the Swiss national surveys, a range of strategies have been used to validate the translations of questions into different languages: already existing translations are reproduced; for new questions, cognitive tests are used to assess the understanding of translations; test-and- revision strategies are also used in some cases. On the whole and considering the diversity of questions, a range of techniques are applied to provide the best possible translations. When available, we used the Swiss version of the validation, but this was not always the case. We now specify in the methods that not all questionnaires were validated for the specific Swiss context.

2. Should the limited language skills of the participants be mentioned as one potential limitation for this study? In the follow-up period the language skills of the participants may improve, could this also have an impact on the results?

The reviewer is correct in mentioning the risks related to the limited language skills of a subsample of participants who do not speak one of the four languages used in the study (French, English, Portuguese, Spanish). Indeed, we addressed this risk by defining one of the exclusion criteria as the lack of language skills allowing for adequate communication with investigators and explicitly mentioned it. The field experience along the first wave of data collection showed that this risk was very minor and no interview had to be interrupted for language problems. We nevertheless added a line about this this risk in the section Quantitative data collection (page 18). Concerning the limited language skills, we prefer to frame the issue in terms of overall education level. One solution we used to mitigate this difficulty is the presence of trained interviewer to accompany the respondents during his answers, in case some questions were hard to interpret. Their training was done in order to favor consistency of help to avoid an interviewer effect.

Concerning the impact of overall education level and language skill improvement, it is true that we cannot exclude that an evolution in a score may be due to improvement in language skill and not in the concept being measured. We will be careful to point this limitation in future publications, to tailor the risk of this limitation depending on the concept measured.

3. There is an unnecessary "and" in the last sentence on page1 (line 60).

We corrected this spelling error

4. In the Acknowledgements there is a spelling mistake in developping (developing).

We corrected the spelling error