PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Cohort Profile: Aussiedler Mortality (AMOR) – Cohort Studies on Ethnic German Migrants from the Former Soviet Union
AUTHORS	Winkler, Volker; Kaucher, Simone; Deckert, Andreas; Leier, Valentina; Holleczek, Bernd; Meisinger, Christa; Razum, Oliver; Becher, Heiko

VERSION 1 – REVIEW

REVIEWER	Raphael Simon Peter Institute of Epidemiology and Medical Biometry, Ulm University, Germany
REVIEW RETURNED	13-Aug-2018

GENERAL COMMENTS	Winkler et al. describe a cohort of ethnic German migrants from the former Soviet Union. While the cohort data is retrospective in nature, it is based on high quality registries. Overall, this is a well written manuscript. I only have few minor comments: 1. "Acute myocardial infarction" is abbreviated as "AMI" within the abstract. Please spell in full. 2. Page 7, lines 41-44 "They were at least 15 years of age at immigration and were and were full samples from municipalities/counties randomly selected from those which could access their population registries". This sentence sounds odd. Consider restructuring. 3. Page 9, lines 40-46. The MIR provides data on incident events for the age range 25 to 74 years. And cohort members were censored when they moved out of the MIR catchment area. Have cohort members also been censored at age 75 years? - Please clarify. 4. Page 11, line 29. The abbreviation "SMR" is used before it is first defined in line 40. 5. Pages 16 & 17. The figure legends have been interchanged during upload of the figure files.

REVIEWER	Philip A Anglewicz Tulane University, United States
REVIEW RETURNED	08-Sep-2018

GENERAL COMMENTS	- There is editing needed throughout the manuscript. References usually appear after the punctuation, not before. Should define the acronym before the first time it's used in the text. Figures 1 and 2 should be labelled, and don't appear in proper sequence. The words "and were" are repeated on pg. 7. Spacing between paragraphs is not consistent. This paper would greatly benefit from a through review of the writing and formatting.

- Sampling- some aspects of sampling are not clear- how did the study identify and recruit participants (how were the specific 34,3939 of 281,356 participants in NRW selected)? How did the study determine the sample sizes? What were the response rates for the initial cohorts?

- It appears that there are only limited sociodemographic measures for participants in the cohorts. But for the limited characteristics, how do they compare for the main study and the nested study participants?

- What were the procedures for obtaining consent for participants of the nested study?

REVIEWER	Prof Laurence Gruer University of Edinburgh, UK
REVIEW RETURNED	27-Sep-2018

GENERAL COMMENTS	This is an interesting and generally well-written paper about cohorts representing a wave of immigration few outside Germany will know about. It would be improved by more clarity on how the cohorts were established, their representativeness, the linkage methods and the governance of the studies. The numbers below refer to the checklist. 2. Most of the main paper is taken up with the background to the cohorts (3 pages) and a description of the four cohorts (4 pages), with 1.5 pages on the findings to date. The abstract has 3 lines on the background, 3 on the participants and 10 on the findings to date. As this is a cohort profile and most of the findings to date are published elsewhere, the abstract should be rebalanced to reflect the paper as a whole. The abstract refers to "the cohort" and to four sub-cohorts. The main paper refers to "all parts of the AMOR cohort" or the AMOR cohorts, to the "NRW cohort", the" Saarland cohort" and the "Augsburg cohort". In P7 L29 Munster is referred to as a "fourth part of the cohort". It is clear to me that although they all involve resettlers, these are four separate cohorts recruited in different regions, at different times and with different purposes. I suggest they are thus collectively called the AMOR cohorts and each is referred to as a cohort in its own right. 4. How the cohorts were recruited is currently not clear enough. Together, the cohorts include 92k individuals out of an estimated 2.1million FSU settlers i.e about 4.3%. To be able to generalise from the cohorts to the whole population of FSU resettlers, it is important to have confidence that the cohorts are genuinely representative samples. At present, I don't think the paper provides that. In P6 L34-35 it states: "It is important to emphasise that the resettlers' German citizenship prevents their direct identification in German registry data such as civil registers or cause of death statistics." More information is thus needed on how the resettlers were identified so they could be included in the samples. The NRW cohort includes only 12.2% of the resettlers in NSW. These were "full samples from municipalities/counties randomly selected from those which could access their population registries". How were the resettlers identified? In addition, it would be helpful to give the total number of counties, how many could access their registries and how many of those were randomly selected". Why could some counties not access their registries?

Was there anything about the selected counties that could introduce bias into the sample? Was the achieved sample size determined in advance or simply the largest that could be obtained in the circumstances.

How were the resettlers identified in the other cohorts? The Saarland cohort was composed of all resettlers whose date of birth was available. As that was only 64.2% of all resettlers, a large proportion were therefore excluded. Was any attempt made to check if there were systematic differences between those with a date of birth and those without? The Augsburg cohort was restricted to resettlers allocated to local transition hostels. What proportion of resettlers went to these hostels and could they be systematically different from those who did not? The Munster cohort is said to be "currently under investigation". It is unclear what this means and how this differentiates this cohort from the other three. If it means that follow-up is still ongoing then it would be clearer to say that.

The data on vital status, cancer and AMIs were linked to the cohorts by record linkage. Different methods of linkage were used in Saarland and Munster, with the Munster method "intended to increase the success rate of the record linkage." This implies that the record linkage rates in the other cohorts was relatively low: it would be helpful to know what the rates were. A high linkage failure rate could make the results less reliable. The AMI data on the resettlers in Augsburg were linked just using sex, initials of the names and date of birth, a much less rigorous procedure than in Saarland where phonetic code of first and last names and city of residence were also used. On P9 L40 it states that case finding completeness was 95%. If the level of record linkage was low, this would potentially affect the estimated AMI incident rates for the resettlers.

The Augsburg nested studies resulted in a 16% response rate for the questionnaire, of whom 32%, or 5% of all the survivors, responded to the invitation for an examination. The low response rates are mentioned as a weakness in P12 L43 and indeed they seems far too low to draw any meaningful conclusions from the results. However, in P10 L57 publications are referred to which suggest these results have been used to infer higher risk factors than in the general population, in contrast to the apparently lower CVD mortality. Could errors in case-finding and risk factor estimation potentially explain this apparent mismatch, rather than "factors that have not yet been studied"?

5. No information is given about the governance of the cohort studies. Is there a single coordinating group which controls all the cohorts or some other arrangement? In the absence of informed consent, what procedures are in place to ensure the interests of the individuals in the cohorts are safeguarded and their privacy protected? In P8 L9 it states that only 50% of local registry offices approved the identification procedures. This suggests there were some concerns about the procedures used. What were these concerns and did they also affect recruitment to the other cohorts?

In P9 L15 it states that vital status was assessed "by record linkage or manually". These two processes are very different - with different levels of error and risk. What proportion of records were searched manually and why?

10. Table 4 shows SMRs "for all cohorts combined". Does this exclude the Munster cohort and if so this should be stated.

11. In P11 L25-30, reference is made to investigating the effect of deprivation on mortality. If possible, it would be useful to briefly summarise what the findings were.

12. Given the questions I raised in 4., I think more caution should be expressed about the representativeness of the cohorts. 13. Has a STROBE checklist been completed? 15. While the standard of English is high overall, there are a number of minor errors in the use of definite or indefinite articles, prepositions or word order which should be corrected. E.g. the first sentence of the Intro would read better as : "Studies on migrant populations contribute to knowledge on the aetiology of diseases and reveal differences in the health status of migrants compared with the autochthonous population." T

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1

Reviewer Name: Raphael Simon Peter

Institution and Country: Institute of Epidemiology and Medical Biometry, Ulm University, Germany

Please state any competing interests or state 'None declared': None declared

Overall, this is a well written manuscript. I only have few minor comments:

1.   "Acute myocardial infarction" is abbreviated as "AMI" within the abstract. Please spell in full.

☐ We changed this accordingly

2.   Page 7, lines 41-44 "They were at least 15 years of age at immigration and were and were full samples from municipalities/counties randomly selected from those which could access their population registries".

This sentence sounds odd. Consider restructuring.

☐ We rephrased the sentence and extended the paragraph as follows to "The NRW cohort is a sample of 34,393 individuals from all 281,356 resettlers who were assigned to the federal state of NRW between 1990 and 2001. Information on all resettlers were available from the central reception centre of NRW. Sampled resettlers were at least 15 years of age at immigration and their first residence had been assigned to municipalities/counties with electronic population registries. The 34,393 resettlers represent 91% of all resettlers of the respective municipalities/counties. Since assignment to the different municipalities within a state has been done on a random basis at time of immigration, the selected cohort constitutes a random sample of all resettlers.".

3.   Page 9, lines 40-46. The MIR provides data on incident events for the age range 25 to 74 years. And cohort members were censored when they moved out of the MIR catchment area. Have cohort members also been censored at age 75 years? - Please clarify.

☐ We clarified and changed the sentence to "For the incidence follow-up cohort members were censored when they moved out of the MIR catchment area and at age 75."

4.   Page 11, line 29. The abbreviation "SMR" is used before it is first defined in line 40.

☐ We corrected that.

5.   Pages 16 & 17. The figure legends have been interchanged during upload of the figure files.

☐ We corrected that.

Reviewer: 2

Reviewer Name: Philip A Anglewicz

Institution and Country: Tulane University, United States

Please state any competing interests or state 'None declared': None declared

- There is editing needed throughout the manuscript. References usually appear after the punctuation, not before. Should define the acronym before the first time it's used in the text. Figures 1 and 2 should be labelled, and don't appear in proper sequence. The words "and were" are repeated on pg. 7. Spacing between paragraphs is not consistent. This paper would greatly benefit from a through review of the writing and formatting.

☐ We addressed all mentioned issues.

- Sampling- some aspects of sampling are not clear- how did the study identify and recruit participants (how were the specific 34,3939 of 281,356 participants in NRW selected)? How did the study determine the sample sizes? What were the response rates for the initial cohorts?

☐ We changed and extended the paragraph on the recruitment of the NRW cohort and added that information on all resettlers were available from the central reception centre of NRW. The cohort was not restricted by sample size calculations but represents 91% of all resettlers living in areas which had electronic population registries at that time. See also our answer to point 2. of reviewer 1.

Further, we want to point out the AMOR cohorts consist of secondary data and cohort members were identified through population registries. Therefore, response is not an appropriate measure. However, for further clarification we added the information that 91% of all resettlers could be identified in municipalities/counties with electronic population registries.

- It appears that there are only limited sociodemographic measures for participants in the cohorts. But for the limited characteristics, how do they compare for the main study and the nested study participants?

☐ Yes, since the cohorts are based on secondary data from the central reception centre and from the population registries, information on individuals are limited to sex, date of birth, age at immigration and date of immigration and country of origin, which is for early years of immigration unspecific by only stating former Soviet Union. However, all cohorts represent the expected distributions of sex, year and age of immigration when compared to national statistics of immigrated resettlers. We also this information to the text on page 8:" All cohorts reflect expected distributions of sex, year and age of immigration when compared to national statistics of immigrated resettlers which indicates that the cohorts are representative for this migrant group.".

With respect to the nested studies presented in table 3, this cannot be expected given the relatively low response which ranges from 16% to 38%. We added response information for the nested case-control study and the nested cross-sectional study on page 10.

- What were the procedures for obtaining consent for participants of the nested study?

☐ We extended the information on obtaining consent on page 10 by adding the following sentence:"

By signing a consent form or by a telephone call, the participants declared willingness to participate in the nested studies presented in table 3 and assured that they were informed sufficiently on the study and the further use of the data collected. Participation could be withdrawn at any time."

Reviewer: 3

Reviewer Name: Prof Laurence Gruer

Institution and Country: University of Edinburgh, UK

Please state any competing interests or state 'None declared': None

This is an interesting and generally well-written paper about cohorts representing a wave of immigration few outside Germany will know about. It would be improved by more clarity on how the cohorts were established, their representativeness, the linkage methods and the governance of the studies. The numbers below refer to the checklist.

2. Most of the main paper is taken up with the background to the cohorts (3 pages) and a description of the four cohorts (4 pages), with 1.5 pages on the findings to date. The abstract has 3 lines on the background, 3 on the participants and 10 on the findings to date. As this is a cohort profile and most of the findings to date are published elsewhere, the abstract should be rebalanced to reflect the paper as a whole.

The abstract refers to "the cohort" and to four sub-cohorts. The main paper refers to "all parts of the AMOR cohort" or the AMOR cohorts, to the "NRW cohort", the" Saarland cohort" and the "Augsburg cohort". In P7 L29 Munster is referred to as a "fourth part of the cohort". It is clear to me that although they all involve resettlers, these are four separate cohorts recruited in different regions, at different times and with different purposes. I suggest they are thus collectively called the AMOR cohorts and each is referred to as a cohort in its own right.

☐ We rebalanced the abstract by deleting some of the presented results and extending the purpose and participants parts. We also changed the terminology of the cohorts according to the reviewer's suggestion throughout the text including the abstract and the title.

4. How the cohorts were recruited is currently not clear enough. Together, the cohorts include 92k individuals out of an estimated 2.1million FSU settlers i.e about 4.3%. To be able to generalize from the cohorts to the whole population of FSU resettlers, it is important to have confidence that the cohorts are genuinely representative samples. At present, I don't think the paper provides that.

☐ We added information on these aspects on page 8, see our answers to point 2. of reviewer 1 and to point 2 and 3 of reviewer 2, respectively.

In P6 L34-35 it states: "It is important to emphasize that the resettlers' German citizenship prevents their direct identification in German registry data such as civil registers or cause of death statistics." More information is thus needed on how the resettlers were identified so they could be included in the samples. The NRW cohort includes only 12.2% of the resettlers in NSW. These were "full samples from municipalities/counties randomly selected from those which could access their population registries". How were the resettlers identified? In addition, it would be helpful to give the total number of counties, how many could access their registries and how many of those were randomly selected". Why could some counties not access their registries? Was there anything about the selected counties that could introduce bias into the sample? Was the achieved sample size determined in advance or simply the largest that could be obtained in the circumstances.

☐ We added the information that data on all resettlers were available from the central reception centre of NRW on page 7 and explained about representativeness and sample size. See also our answers to point 2. of reviewer 1 and to point 2 of reviewer 2. Further, we want to highlight that municipalities/counties with electronic population registries were necessary to perform record linkages for the identification of resettlers within the population registries and for assessing their vital status. In 2002, there were 99 out of 396 municipalities/counties with electronic population registries in NRW. However, we did not add that information to the manuscript, because we think it might be confusing and is not be useful to understand that sampling procedure.

As a further information for the reviewer we would like to mention that at the beginning of the study we contacted the authorities of the largest federal states in Germany whether they could provide us with a list of all resettlers. This request was quite unique, and some states responded hesitantly, albeit not fully negative. At the end, we selected North Rhine-Westphalia who provided a positive answer first. Since the assignment to the different federal states was done randomly, we were satisfied with this procedure.

How were the resettlers identified in the other cohorts? The Saarland cohort was composed of all resettlers whose date of birth was available. As that was only 64.2% of all resettlers, a large proportion were therefore excluded. Was any attempt made to check if there were systematic differences between those with a date of birth and those without? The Augsburg cohort was restricted to resettlers allocated to local transition hostels. What proportion of resettlers went to these hostels and could they be systematically different from those who did not? The Munster cohort is said to be "currently under investigation". It is unclear what this means and how this differentiates this cohort from the other three. If it means that follow-up is still ongoing then it would be clearer to say that.

☐ With respect to the Saarland cohort, we added the following sentence to page 7:"Resettlers were identified by collecting application forms on obtaining German passports in all local refugee offices of the Saarland." The application is mandatory for resettlers and was usually done 3 to 6 months after arrival.

Yes, cohorts resemble the expected distributions based on national statistics (see also our answer to point 3 of reviewer 2).

With respect to the Augsburg cohort: All the resettlers firstly arrived at the transition hostels after being allocated to Augsburg by the government, the hostels were their local registration institution where they had to go through. From there, they were allowed to move and settle in the city, respectively in the region of Augsburg or elsewhere. We added a sentence in the text

We changed the wording for the Münster cohort as suggested.

The data on vital status, cancer and AMIs were linked to the cohorts by record linkage. Different methods of linkage were used in Saarland and Munster, with the Munster method "intended to increase the success rate of the record linkage." This implies that the record linkage rates in the other cohorts was relatively low: it would be helpful to know what the rates were. A high linkage failure rate could make the results less reliable. The AMI data on the resettlers in Augsburg were linked just using sex, initials of the names and date of birth, a much less rigorous procedure than in Saarland where phonetic code of first and last names and city of residence were also used. On P9 L40 it states that case finding completeness was 95%. If the level of record linkage was low, this would potentially affect the estimated AMI incident rates for the resettlers.

☐ The described implication of our wording that one of the record linkage procedures yielded low accuracy was not intended. In fact, record linkage of the Saarland cancer registry worked very well since the staff of this cancer registry is able to access internally names and other characteristics of

each individual diagnosed with cancer. This enables a manual check of all files that matched partly during the automated linkage procedure e.g. date of birth matched but the name not completely.

The record linkage done by the Münster cancer registry is more difficult since there is no possibility to get individual characteristics like names within the data of this cancer registry. Due to very strict federal laws of data protection this registry uses a stochastic method. However, it is not possible to provide a failure rate to assess the linkage quality since there is no data for comparison. But success rates for the two cancer registries are likely to be comparable since both result in similar incidence rates.

The linkage procedure in Augsburg was indeed based on the variables sex, initials and date of birth. Given the rather small number of people living in the area of Augsburg (registry area) with about 270,000 individuals chances to identify two people with the same date of birth and the same initials are low. However, there still might be wrong hits. For instance, one wrong match was identified where the AMI happened before the person moved in. Hence, the AMI incidence might have been slightly overestimated if wrong matches were not detected.

The wording that case finding completeness was 95% might be misleading. The completeness refers to the registry as a whole and not the resettler cohort. Therefore, we deleted this information.

The Augsburg nested studies resulted in a 16% response rate for the questionnaire, of whom 32%, or 5% of all the survivors, responded to the invitation for an examination. The low response rates are mentioned as a weakness in P12 L43 and indeed they seems far too low to draw any meaningful conclusions from the results. However, in P10 L57 publications are referred to which suggest these results have been used to infer higher risk factors than in the general population, in contrast to the apparently lower CVD mortality. Could errors in case-finding and risk factor estimation potentially explain this apparent mismatch, rather than "factors that have not yet been studied"?

☐ The mentioned conclusion is not only based on our studies, with response ranging up to 38% for the nested case-control study focusing on CVD risk factors (response has been added on page 10), but also on an older study by Aparicio et al (ref 36). Furthermore, all our publications discuss the underlying response in detail. But to avoid any misunderstandings that may arise, we rephrased this sentence as follows: "Given the limited knowledge which also arises from the generally low response of this group to actively take part in studies, the observed differences cannot be explained and may even be related to factors which have not yet been studied.".

5. No information is given about the governance of the cohort studies. Is there a single coordinating group which controls all the cohorts or some other arrangement?

☐ The AMOR cohorts are coordinated by the two PIs Prof. Dr. Heiko Becher and PD Dr. Volker Winkler. We added this information to page 14.

In the absence of informed consent, what procedures are in place to ensure the interests of the individuals in the cohorts are safeguarded and their privacy protected?

☐ The database to identify resettlers in population registries was at that time handed over with clearance from the federal ministry to set up register-based cohorts in NRW. The ministry as well as the ethical board of the University of Heidelberg approved the study approach which includes a procedure ensuring that resettler became anonymized as soon as they were censored e.g. in case of dying. The cause of death is then merged to an unique ID without names, city of residence, etc.

In P8 L9 it states that only 50% of local registry offices approved the identification procedures. This suggests there were some concerns about the procedures used. What were these concerns and did they also affect recruitment to the other cohorts?

☐ We would like to emphasize that it is legal in Germany to request this information from local registry offices. Our follow up procedure is in principle based on this law. This was also approved, see above. It is also important to mention the federal organisation of Germany in this context. With respect to the law on population registration this means that a general framework is set be a national law. Federal laws set concrete guidelines within that framework. In 2015, a new national law came into force which made some local registry offices in NRW (Münster is a part of NRW) and only in NRW stop the requests for our study.

In P9 L15 it states that vital status was assessed "by record linkage or manually". These two processes are very different - with different levels of error and risk. What proportion of records were searched manually and why?

☐ After the initial phase of the NRW cohort (the first follow up), we increased our efforts to additionally perform manual assessment of the vital status in municipalities without electronic population registries. This became more important since some resettlers moved to cities without electronic registries and would otherwise have been lost to follow up. However, a manual follow up is very time consuming and costly and was therefore avoided whenever possible.

With time more and more cities switched to electronic registries which allowed for some cities to perform a record linkage even though a manual follow up was done earlier. A comparison of the two methods showed similar results. Due to the changing conditions it is difficult to give an exact number of resettler that were followed up manually. Overall, we sent about 200,000 individual requests to registry offices and approximately 10,000 were answered manually. Altogether the 20 years follow up of the NRW cohort resulted in only 5.3% lost to follow up, while it needs to be highlighted that 41.9% of those who became lost to follow up moved abroad. Therefore, we are very confident about the high quality of the follow up.

10. Table 4 shows SMRs "for all cohorts combined". Does this exclude the Munster cohort and if so this should be stated.

☐ We stated now clearly that the Münster cohort is not included.

11. In P11 L25-30, reference is made to investigating the effect of deprivation on mortality. If possible, it would be useful to briefly summarise what the findings were.

☐ We extended this paragraph to give more explanation as follows: "Since the allocation of the resettlers was done quasi-randomly within Germany, the cohort allowed to investigate the effect of regional deprivation on individual mortality by making use of this natural experiment. The NRW cohort was used to assess the effect of regional deprivation on individual mortality by aggregating the 54 counties of NRW in six deprivation clusters. Standardized mortality ratios (SMRs) comparing resettler mortality with the mortality of the autochthonous population resulted in highest SMRs for regions with the highest level of regional deprivation."

12. Given the questions I raised in 4., I think more caution should be expressed about the representativeness of the cohorts.

☐ Please, see our answer to 4.

13. Has a STROBE checklist been completed?

☐ We attached a completed STROBE checklist for cohort studies.

15. While the standard of English is high overall, there are a number of minor errors in the use of definite or indefinite articles, prepositions or word order which should be corrected. E.g. the first sentence of the Intro would read better as :

"Studies on migrant populations contribute to knowledge on the aetiology of diseases and reveal differences in the health status of migrants compared with the autochthonous population." T

☐ We changed the sentence accordingly and did some more language editing.

VERSION 2 – REVIEW

REVIEWER	Raphael Simon Peter Institute of Epidemiology and Medical Biometry, Ulm University, Germany
REVIEW RETURNED	16-Dec-2018

GENERAL COMMENTS	The authors have adequately addressed all my points.

REVIEWER	Philip Anglewicz Tulane University
REVIEW RETURNED	12-Dec-2018

GENERAL COMMENTS	The authors have addressed my previous comments.

REVIEWER	Laurence Gruer University of Edinburgh
REVIEW RETURNED	12-Dec-2018

GENERAL COMMENTS	Thank you for satisfactorily addressing the points I made. I think this is now an excellent paper which merits publication. A few minor errors I noted on re-reading. p5 L 11 replace "constitute of" with "consist of" p8 L11 better as "where all resettlers were first assigned" p12 L15 "with to" delete "to" p12 L18 delete "as SMR". The table immediately below makes it clear that SMRs are used.