

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)	Identifying patterns of non-communicable diseases in developed eastern coastal China: A longitudinal study of electronic health records from 12 public hospitals
AUTHORS	Yu, Dehua; Shi, Jianwei; Zhang, Hanzhi; Wang, Zhaoxin; Lu, Yuan; Zhang, Bin; Pan, Ying; Wang, Bo; Sun, Pengfei

VERSION 1 - REVIEW

REVIEWER	Ruth Webster The George Institute for Global Health, AUstralia
REVIEW RETURNED	25-Jan-2017

GENERAL COMMENTS	<p>The concept of using longitudinal EHR records to document the proportions of NCD admissions to eastern coastal China hospitals is good and can provide a reasonably simple method of looking at trends over time without the prohibitive cost of individual data collection over many years. The paper however is quite confusing to read with a lot of results presented and could be improved by addressing the following points:</p> <ol style="list-style-type: none">1. NCDs are of importance and concern particularly related to the global increase of the 'top 4' of cardiovascular disease, cancer, respiratory disease and diabetes mellitus primarily due to the epidemiological transition. It would be useful to specifically discuss the results of this study in relation to these 'top 4' global concerns rather than present multiple results, some of which may not be that interesting or relevant.2. When discussing the results, every result is presented which is confusing. It would perhaps simplify the paper if the key results that are of major clinical and economic interest were presented and discussed. For example, it is reported in the 0 to 10 year old group that the most common NCD was chronic tonsillitis – is this a concern? Is this something that needs to be addressed in public health programs? IF not, then perhaps this type of detail can be left in the tables and figures without confusing the text.3. The discussion and conclusions contain several very general statements about what needs to be done such as 'Public health campaigns should be targeted at the regional level and should consider differences'. This is unsupported by the evidence presented in the paper and is too general. More specific recommendations could be made based on the outcomes presented in the paper. It may be that the data is too general to make specific recommendations, in which case simply discussion of the next research questions to be addressed would be ideal.
-------------------------	--

	<p>4. Were any differences seen between the different hospitals representative of different socio-economic groups?</p> <p>5. For the hospitals that declined to participate, what were their characteristics? Were they all the lowest/middle/highest socio-economic hospitals? Did their non participation affect the results in anyway?</p> <p>6. The methods section is not very detailed. How were NCD conditions defined and who defined them before data was extracted? How were your results affected by patients who came back multiple times which is likely for patients with chronic disease? What proportion of visits were repeat visits? How did you define 'patients who received their first diagnosis of an NCD between 2003 and 2014'? How did you know this was their first diagnosis?</p> <p>7. What is known about the quality of data entry into hospital EHRs? Is the data reliable? Who coded the diagnoses to match ICD-9 and GBD NCD classification?</p> <p>8. The number of admissions over time has increased dramatically - is this due to a large population? Is there the possibility that over time patients have changed their practices and now come to larger hospitals more than community health centres thus appearing that proportions have changed?</p> <p>This is potentially a very interesting study into trends into NCD admissions over time but could be improved with further refining.</p>
--	--

REVIEWER	Peter Lloyd-Sherlock UEA, UK
REVIEW RETURNED	02-Mar-2017

GENERAL COMMENTS	<p>This paper offers an interesting analysis of a vast amount of data on inpatients in 12 Chinese hospitals. The data set offers important opportunities for analysis, but I do not feel that the authors fully exploit them.</p> <p>The paper would be much more effective (less of a square peg in a round hole), if the authors play to the strengths of the data. Rather than framing the study as a general review of NCD epidemiological trends (and there are many of these for China already), why not approach it as study of hospitals? You could include some early discussion about pressures on inpatient provision and could then assess the extent to which these are being driven by specific health conditions. It might also be revealing to make some comparisons across the 12 hospitals.</p> <p>I realise that this will lead to a rather different kind of paper, with a more explicit policy focus, but I feel that this would make a far better and more original contribution to what is already known in China.</p> <p>It would also be helpful to focus on older people who appear to account for a large and growing % of inpatients. Is this seen equally across all 12 hospitals?</p> <p>If the authors strongly object to taking this different approach, I will</p>
-------------------------	---

	<p>review the current approach on its own merits. My main comments there are:</p> <p>the finding for diabetes is interesting to some extent, but you need to emphasise that other conditions account for a larger % of inpatients</p> <p>develop your explanations more carefully. do your findings reflect changes in overall population health or changes in health policy (eg whether people are admitted or treated as inpatients, etc.)?</p> <p>provide data on an age-specific and age-weighted basis</p>
--	---

REVIEWER	<p>Diem, Gunter Austrian Public Health Association Vienna, Austria</p>
REVIEW RETURNED	<p>02-Mar-2017</p>

GENERAL COMMENTS	<p>Only data from hospitalized patients have been used. Maybe it's worth mentioning if there is a possible confounding on diagnosis due to economic reasons like DRM system or similar (explanation: DRM systems tend to use more severe diagnoses to increase their remuneration)</p> <p>The study shows remarkable differences in NCD prevalences in comparison with western populations. The FHS is mentioned in the article, it could be interesting to compare with other international longitudinal studies.</p>
-------------------------	--

VERSION 1 – AUTHOR RESPONSE

Response to reviewers

To Reviewer 1

12 April, 2017

Dear Professor,

We appreciate your review of our work. Based on your suggestions, we have revised the manuscript, particularly the Results and Discussion sections, to make it more clear and focused. In addition, we have provided additional detailed information on the Chinese electronic health records and introduced other issues related to the characteristics of China to make this manuscript easier to understand. The changes in the revised manuscript have been highlighted in yellow. Additionally, our point-by-point responses are presented below. We hope that the revisions increase the persuasiveness of the manuscript.

Sincerely,

Zhaoxin Wang

Associate Professor, Ph.D.

School of Medicine

Tongji University

Address: 1239 Si Ping RD, Yangpu District, Shanghai, 200092, China

E-mail: supercell002@sina.com;M.P.: 86-13918537473

The concept of using longitudinal EHR records to document the proportions of NCD admissions to eastern coastal China hospitals is good and can provide a reasonably simple method of looking at trends over time without the prohibitive cost of individual data collection over many years. The paper however is quite confusing to read with a lot of results presented and could be improved by addressing the following points:

1. NCDs are of importance and concern particularly related to the global increase of the 'top 4' of cardiovascular disease, cancer, respiratory disease and diabetes mellitus primarily due to the epidemiological transition. It would be useful to specifically discuss the results of this study in relation to these 'top 4' global concerns rather than present multiple results, some of which may not be that interesting or relevant.

Answer: Thank you for your comment! We agree that the results should be more focused and concise. According to your suggestion, we have substantially revised the Results and Discussion sections, and the top NCDs including diabetes, cardiovascular and circulatory diseases, urogenital diseases, digestive diseases and cancer have been further emphasized and discussed in this study. Please see the Results and Discussion sections from line 216 to line 349.

2. When discussing the results, every result is presented which is confusing. It would perhaps simplify the paper if the key results that are of major clinical and economic interest were presented and discussed. For example, it is reported in the 0 to 10 year old group that the most common NCD was chronic tonsillitis – is this a concern? Is this something that needs to be addressed in public health programs? If not, then perhaps this type of detail can be left in the tables and figures without confusing the text.

Answer: According to your helpful suggestions, we simplified our paper and depicted the key results with a stronger focus on the NCDs with major clinical and economic impact. The descriptions of NCDs such as chronic tonsillitis and other non-primary diseases were deleted to make the discussion more consistent with the results and to shift the focus to the meaningful public health programs for certain age groups or vulnerable populations. Please see lines from 221 to 238 in the Results section.

3. The discussion and conclusions contain several very general statements about what needs to be done such as 'Public health campaigns should be targeted at the regional level and should consider differences'. This is unsupported by the evidence presented in the paper and is too general. More specific recommendations could be made based on the outcomes presented in the paper. It may be that the data is too general to make specific recommendations, in which case simply discussion of the next research questions to be addressed would be ideal.

Answer: Based on your suggestions, we have revised the Results section and pointed out more specific results that are consistent with the evidence proposed in the Discussion section. For example, regarding your comment that statements in the Discussion are too general, such as "Public health campaigns should be targeted at the regional level and should consider differences in demographic factors across regions", we have revised this section to read as follows: " These factors imply that specific risk factors for frequently occurring NCDs in different gender groups should be monitored in addition to the diseases or groups that are currently monitored and screened for due to their significant burden. Other frequently occurring NCD diseases and a wider population should also be targeted, such as the prevention and screening of digestive diseases in females or increased preventative measures for diabetes and blood, endocrine, cardiovascular and circulatory diseases among the younger and population. In addition, all these call for the improvement and more invest in prevention and control of NCDs by community health institutions, which is lagged even in the economic-developed eastern coastal or averagely the whole China compared with the western countries[30]". Please see lines from 270 to 349.

4. Were any differences seen between the different hospitals representative of different socio-economic groups?

Answer: In this study, we did not include any region or socioeconomic factors of the hospitals in the

analysis because all of the hospitals were located in an urban region and the aim of this study was to reflect the whole disease spectrum in the relatively affluent region of eastern coastal China compared with other regions. Currently, to obtain a representative sample, we divided the area into different socioeconomic groups.

In addition, we tried to analyse the socioeconomic groups among the inpatients in eastern coastal China. However, disappointingly, we found that the electronic health records (EHRs) did not include income variables. Although the EHRs contain a profession variable, this data is difficult to analyse because the patients are not required to provide this information when they enter the hospital and there were too many missing values (the effective rate of this variable was 48.9%). This information has been added to the limitations section of the Discussion. Please see lines 357-359.

5. For the hospitals that declined to participate, what were their characteristics? Were they all the lowest/middle/highest socio-economic hospitals? Did their non participation affect the results in anyway?

Answer: In our study, the three hospitals that declined to participate were in Shandong, Jiangsu and Zhejiang Provinces. Specifically, the hospital in Shandong Province was in the low socioeconomic group, the hospital in Jiangsu Province was in the middle socioeconomic group, and the hospital in Zhejiang Province was in the high group. In total, the hospitals included in the analysis included 4 hospitals in each of the high, middle and low socioeconomic groups. Therefore, we think that the lack of participation of these hospitals would not influence the final results. Please see lines 152-157.

6. The methods section is not very detailed. How were NCD conditions defined and who defined them before data was extracted? How were your results affected by patients who came back multiple times which is likely for patients with chronic disease? What proportion of visits were repeat visits? How did you define 'patients who received their first diagnosis of an NCD between 2003 and 2014'? How did you know this was their first diagnosis?

Answer: Based on your helpful comment, we have revised the Methods section to provide additional details. In 2001, China began using the uniform electronic health record version in all large hospitals. When admitted to the hospital, patients are diagnosed by physicians according to the ICD-9 disease code classification. The disease codes were issued by the Ministry of Health in China. This has been explained in the Methods section on line 157-167.

To account for the patients with multiple hospital admissions, we excluded any duplicate patients by searching and analysing the Identification Card numbers, which are required to be entered into the system when a patient is admitted to the hospital. In lines of 184-186, we added that the data were analysed after excluding any duplicate patients.

Regarding the question of how to define 'patients who received their first diagnosis of an NCD between 2003 and 2014'? and 'How did you know this was their first diagnosis?' this information is essential in the EHRs and usually the patients' primary diagnoses were entered into the EHR by the physicians. Therefore, we are confident that the NCD diagnoses information was precise.

7. What is known about the quality of data entry into hospital EHRs? Is the data reliable? Who coded the diagnoses to match ICD-9 and GBD NCD classification?

Answer: The EHRs of hospitals in China contain two parts. The first part contains the patients' personal information, including their gender, age, identification card number, profession, address, etc. This information is usually provided by the patients or their family. In China, a citizen's ID contains their precise birth information. To ensure that the personal data were reliable, we obtained the patients' information from their ID. The second part contains the inpatients' hospitalization information, including their diagnoses code, discharge status, pathologic diagnosis (if possible), operation code (if possible), etc. This information is provided by the patient's physician, which ensures its reliability. In terms of the diagnoses code, each inpatient is coded with an ICD-9 disease code by their physician. In addition, the GBD has well-defined categories for NCDs. Therefore, we extracted the inpatients' NCDs using their ICD-9 codes and classified them into different categories. Realizing that many

readers may not be familiar with the Chinese health information system, we have now explained the coding system in much more detail from line 157 to 167.

8. The number of admissions over time has increased dramatically - is this due to a large population? Is there the possibility that over time patients have changed their practices and now come to larger hospitals more than community health centres thus appearing that proportions have changed?

Answer: The increase in the number of NCD inpatients from 2003 to 2014 also caught our attention during the analysis. However, we did not explain this finding in the original manuscript. There are several possible explanations. First, as you mentioned, due to improvements in economics, health insurance and convenient transportation in China, more people are more likely to go to larger tertiary hospitals because they are equipped with better devices and doctors. In addition, there are no strict referral policies. Second, the increase in the number of severe NCD patients may result from increased stress a lack of exercise, and air pollution, among other factors. According to your comment, we have added these possible reasons to the Discussion section. Please see lines 273-284.

This is potentially a very interesting study into trends into NCD admissions over time but could be improved with further refining.

Answer: Thank you for your kind suggestion. We hope our revisions have made the article more clear and persuasive.

Response to Reviewer 2

12 April, 2017

Dear Professor,

We greatly appreciate your valuable comments and suggestions, which have helped us to improve our manuscript. Based on your comments and suggestions, we have specifically revised the Results and Discussion sections. The changes in the revised manuscript have been highlighted in yellow. Additionally, our point-by-point responses are presented below. We hope that the revisions increase the persuasiveness of the manuscript.

Sincerely,

Zhaoxin Wang

Associate Professor, Ph.D.

School of Medicine

Tongji University

Address: 1239 Si Ping RD, Yangpu District, Shanghai, 200092, China

E-mail: supercell002@sina.com; M.P.: 86-13918537473

This paper offers an interesting analysis of a vast amount of data on inpatients in 12 Chinese hospitals. The data set offers important opportunities for analysis, but I do not feel that the authors fully exploit them.

The paper would be much more effective (less of a square peg in a round hole), if the authors play to the strengths of the data. Rather than framing the study as a general review of NCD epidemiological trends (and there are many of these for China already), why not approach it as study of hospitals?

You could include some early discussion about pressures on inpatient provision and could then assess the extent to which these are being driven by specific health conditions. It might also be revealing to make some comparisons across the 12 hospitals. I realise that this will lead to a rather different kind of paper, with a more explicit policy focus, but I feel that this would make a far better and

more original contribution to what is already known in China.

Answer: Thank you for your comment! We appreciate your suggestions about how to make this paper more interesting and a more original contribution. In this study, all of the hospitals were chosen from the eastern coastal region of China, where the economy is relatively affluent and the geography is similar, resulting in a relatively homogenous group of hospitals and patients in these provinces. Therefore, we consider the 12 hospitals in eastern coastal China as a single group. However, we believe your comment brings up a valid point of comparison and feel that a comparison of hospitals between the eastern, central, and western regions of China would be better because these regions have more economic, political, and geopolitical differences that may directly lead to differences in the prevalence of NCDs. In this case, we keep the original idea and hope you can understand.

It would also be helpful to focus on older people who appear to account for a large and growing % of inpatients. Is this seen equally across all 12 hospitals?

Answer: In each hospital, older inpatients accounted for a higher percentage and larger trend in NCDs, which is consistent with the epidemiological features of NCDs. In this study, we analysed the NCD spectrum for all age groups, including young-aged individuals, because the threat of NCDs has increased in the young population in economically developed areas of China. However, based on your suggestions, we added a discussion of the most commonly observed conditions in the older population, such as cerebral infarction, coronary heart disease and hypertension, and indicated that these conditions should be monitored. Please see lines 316-320.

If the authors strongly object to taking this different approach, I will review the current approach on its own merits. My main comments there are:

the finding for diabetes is interesting to some extent, but you need to emphasise that other conditions account for a larger % of inpatients

Answer: Thank you for your helpful suggestion. In this study, cardiovascular and circulatory diseases, chronic respiratory diseases, digestive diseases, urogenital diseases and cancer were the top diseases and exhibited different tendencies from 2003 to 2014 in this region. Based on your suggestion, we have revised the Results and Discussion sections to focus on the most prevalent diseases. Please see lines 221-268 in the Results Section and lines 285-349 in the Discussion Section.

develop your explanations more carefully. do your findings reflect changes in overall population health or changes in health policy (eg whether people are admitted or treated as inpatients, etc.)?

Answer: We agree that the previous discussion was not clear. This paper aims to reflect changes in individuals who were diagnosed with severe NCDs and how the spectrum of NCDs was affected by the environment, economy, and lifestyle in this transforming region in China. In addition, we also provided suggestions for how to help the vulnerable populations.

The explanations in the Discussion section have been revised to be clearer and focused. Please see from line 269 to 349.

provide data on an age-specific and age-weighted basis

Answer: Thank you for this reminder. In this study, the analyses were performed using the proportion of inpatients with each NCD in order to relatively reflect the prevalence of severe NCDs in this region, it's the relative index instead of absolute index. Thereby, the age data were not weighted.

To Reviewer 3
12 April, 2017

Dear Professor,

We appreciate your review of our work. Based on your suggestions, we have now revised the

manuscript. The changes in the revised manuscript have been highlighted in yellow. Additionally, our point-by-point responses are presented below. We hope that the revisions increase the persuasiveness of the manuscript.

Sincerely,
Zhaoxin Wang
Associate Professor, Ph.D.
School of Medicine
Tongji University
Address: 1239 Si Ping RD, Yangpu District, Shanghai, 200092, China
E-mail: supercell002@sina.com;M.P.: 86-13918537473

Please leave your comments for the authors below
Only data from hospitalized patients have been used. Maybe it's worth mentioning if there is a possible confounding on diagnosis due to economic reasons like DRM system or similar (explanation: DRM systems tend to use more severe diagnoses to increase their remuneration).
Answer: Thank you for your comment! In China, the physicians at the hospital determine the patients' diagnoses. Given that the tertiary hospitals in China currently use the global budget payment system as their main mode of payment, they are less likely to use more severe diagnoses to increase their remuneration. We have added an explanation of the Chinese EHR and diagnosis procedure on lines 157-164.

The study shows remarkable differences in NCD prevalences in comparison with western populations. The FHS is mentioned in the article, it could be interesting to compare with other international longitudinal studies.

Answer: Thank you for your helpful reminder. According to your suggestions, we have added the international comparison in the Discussion section.

For instance, we found that a relationship between socioeconomic status and NCDs has been demonstrated in high-income countries. For instance, in Allen et al.(2017)'s systematic review, they showed that NCD behavioural risk factors is well established in high-income countries. Also Goryakin et al.(2017) showed that when shifting from rural to urban areas, the average body mass index (BMI), total cholesterol level and systolic blood pressure, increased[17], demonstrating that high urbanisation increases the occurrence of cardiovascular and circulatory diseases, which is in consistent with this study. Please see lines 277-284 and lines 289-294.

VERSION 2 – REVIEW

REVIEWER	Ruth Webster The George Institute for Global Health, Australia
REVIEW RETURNED	05-May-2017

GENERAL COMMENTS	I think the authors misunderstood the comment about socioeconomic comparisons. The methods section (line 144) stated that 3 cities were selected that represented high, middle and low socioeconomic status but results are not compared between these 3 regions. Were there any differences between NCD admissions across these 3 regions? The new paragraph at line 248 to 253 does not make sense. Why are non-significant results presented and commented as being 'higher and found in more stable population'?
-------------------------	---

REVIEWER	Peter Lloyd-Sherlock University of East Anglia, UK
-----------------	---

REVIEW RETURNED	21-Apr-2017
------------------------	-------------

GENERAL COMMENTS	My previous suggestions have been incorporated into this resubmission. I only have one minor suggestion. The authors need to consider the extent to which their findings are affected to changing protocols and practice of NCD hospital admissions over time. Policy change, as well as the availability of new treatments or therapies, may mean that a person who presents with a given condition in the past may be more or less likely to become an inpatient in the present day. This could introduce some bias.
-------------------------	--

VERSION 2 – AUTHOR RESPONSE

To Reviewer 1

Reviewer Name: Ruth Webster

Institution and Country: The George Institute for Global Health, Australia

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

Please leave your comments for the authors below

I think the authors misunderstood the comment about socioeconomic comparisons. The methods section (line 144) stated that 3 cities were selected that represented high, middle and low socioeconomic status but results are not compared between these 3 regions. Were there any differences between NCD admissions across these 3 regions?

Answer: Thank you for your helpful suggestion. We apologize for the misunderstanding. The comparison of the disease spectrums between the regions with high, middle and low socioeconomic status was not significant. The cities in eastern coastal China are of high socioeconomic status, and there is not much difference among the GDPs in these cities. When conceiving this study, we performed a regional comparison within eastern coastal China, but the results were not significant. Because of this lack of significance, we instead focused on comparing personal characteristics. Following your suggestion, we have now added the following sentence to the revised manuscript: Because there was no significant difference in disease spectrum between regions with high, middle and low socioeconomic status within eastern coastal China, we did not include region as a factor (lines 195-198).

The new paragraph at line 248 to 253 does not make sense. Why are non-significant results presented and commented as being higher and found in more stable population?

Answer: We apologize for the lack of clarity in this section. We intended to show the difference between men and women and intended to state the significant results. We have revised the text as follows: "We found a significant increase in mental and behavioural disorders ($Z=5.130$, $P<0.001$) and musculoskeletal disorders ($Z=6.896$, $P<0.001$) among women, but no significant changes among men. Also, it is worth noting that the percentage of men who were diagnosed with digestive diseases ($Z=-4.284$, $P<0.001$) and sensory organ diseases ($Z=-3.342$, $P<0.001$) reduced significantly from 2003 to 2014, but there was no significant change for women." In addition, we performed a more thorough examination to increase the validity of our work. Please see lines 253-258 in the Results section.

To Reviewer 2

Reviewer Name: Peter Lloyd-Sherlock

Institution and Country: University of East Anglia, UK

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

Please leave your comments for the authors below

My previous suggestions have been incorporated into this resubmission. I only have one minor suggestion. The authors need to consider the extent to which their findings are affected to changing protocols and practice of NCD hospital admissions over time. Policy change, as well as the availability of new treatments or therapies, may mean that a person who presents with a given condition in the past may be more or less likely to become an inpatient in the present day. This could introduce some bias.

Answer: Thank you for your helpful suggestions. Indeed, policy changes as well as improvements in medical technology and treatments may have influenced disease status. Based on your suggestion, we have added the following sentence to the revised manuscript in the limitations section: Third, because the period of data collection spanned a long time, there may be some bias related to disease diagnosis caused by changes in policy as well as changes in the availability of new treatments or therapies. Such changes could have resulted in patients with a given condition in the past becoming more or less likely to become inpatients in the present day (lines 364-368).

VERSION 3 – REVIEW

REVIEWER	Peter Lloyd-Sherlock University of East Anglia, UK
REVIEW RETURNED	12-Jun-2017

GENERAL COMMENTS	All my previous comments have been adequately addressed.
-------------------------	--