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

<table>
<thead>
<tr>
<th>TITLE (PROVISIONAL)</th>
<th>Identification of Reciprocal Causality between Non-alcoholic Fatty Liver Disease and Metabolic Syndrome by Simplified Bayesian network in Chinese Population</th>
</tr>
</thead>
<tbody>
<tr>
<td>AUTHORS</td>
<td>Zhang, Yongyuan; Zhang, Tao; Zhang, Chengqi; Tang, Fang; Zhong, Nvjuan; Li, Hongkai; Song, Xinhong; Lin, Haiyan; Liu, Yanxun; Xue, Fuzhong</td>
</tr>
</tbody>
</table>

VERSION 1 - REVIEW

<table>
<thead>
<tr>
<th>REVIEWER</th>
<th>Dr Victoria Allgar</th>
</tr>
</thead>
<tbody>
<tr>
<td>University</td>
<td>University of York, England</td>
</tr>
<tr>
<td>REVIEW RETURNED</td>
<td>11-May-2015</td>
</tr>
</tbody>
</table>

GENERAL COMMENTS

Although utilising more complex statistical methodology, the paper is well written and conveys the findings well. The general reader may struggle with the statistical terms.

<table>
<thead>
<tr>
<th>REVIEWER</th>
<th>Amedeo Lonardo, M.D.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Outpatient Liver</td>
<td>Outpatient Liver Clinic - NOCSAE - Azienda USL Modena - Italy</td>
</tr>
<tr>
<td>REVIEW RETURNED</td>
<td>11-May-2015</td>
</tr>
</tbody>
</table>

GENERAL COMMENTS

1) This manuscript is difficult to follow owing to the use of non-idiomatic English and excessive use of statistical (as opposed to medical) information.

2) Introduction - The rationale of the study and pre-test hypothesis are poorly detailed. This may be due to incomplete/outdated referencing. Discussion – This section is too long and fails to illustrate the chief findings and the putative biological basis underlying it.

SPECIFIC COMMENT

MAJOR

1) This manuscript needs revision by a mothertongue native. Here follow just a few examples of the many typos and sentences in non-idiomatic English of which the manuscript abounds "In recent years, due to the lifestyle and economic changes, the prevalence of NAFLD and MetS has been rapidly increasing in Chinese populations, which has become one of the major public-health challenges". Written as it is, the sentence includes a typo
(tos) and a wrong sentence (it seems as if the Chines populations rather than NAFLD & the MS are a health challenge.

The following sentence "In pathogenesis, insulin resistance (IR) plays a critical role in the development of both NAFLD and MetS, expecting the inherent link between them." Who is the subject of the word "expecting"? Is "to expect" the most appropriate verb to be used here?

"Although accumulating epidemiologic evidences support an association between NAFLD and MetS/its components, it remained unclear whether NAFLD was a cause or consequence of MetS". This sentence includes a) the plural of "evidence" which does not exist in idiomatic English and b) two different tenses (support; remained; was) rather than one (support; remains; is).

Lines 75-88 are quite hard to follow. This occurs owing to poor English and excess statistical details being given in the absence of robust physiopathological basis of why NAFLD should be located in the cascade of physiopathological eventually culminating in the development of NAFLD. The Authors should be encouraged to read those recent studies which fully substantiate how come that NAFLD may be a precursor of the MetS rather than merely its hepatic manifestation (Lonardo A, Ballestri S, Marchesini G, Angulo P, Loria P. Nonalcoholic fatty liver disease: a precursor of the metabolic syndrome. Dig Liver Dis. 2015;47:181-90; Machado MV, Cortez-Pinto H. Management of fatty liver disease with the metabolic syndrome. Expert Rev Gastroenterol Hepatol. 2014;8:487-500.; Yki-Järvinen H. Non-alcoholic fatty liver disease as a cause and a consequence of metabolic syndrome. Lancet Diabetes Endocrinol. 2014;2:901-10. Loria P, Lonardo A, Anania F. Liver and diabetes. A vicious circle. Hepatol Res. 2013;43:51-64.).

2) Introduction - The classical view maintains that NAFLD is the hepatic manifestation of the MetS. Stated otherwise, it postulates NAFLD as the consequence of pre-existing MetS. More recent data, however, highlight the bidirectional relationship existing between these 2 different conditions. It think the Authors need to shortly summarize the state of art concerning the relationship between NAFLD and the Metabolic; next they need to state what they expected to find and which specific statistical method they followed. Discussion – Start the discussion by shortly summarizing the chief findings. Next, discuss if such data are in agreement/conflict with previous studies. Do not forget to highlight novel findings and to succinctly illustrate th eputative biological basis underlying these findings. Conclude by listing strengths and limitations of the study. Among the latter, please include that NAFLD was not subjected to any semiquantitative indices (Ballestri S, Nascimbeni F, Romagnoli D, Lonardo A. Do ultrasonographic semiquantitative indices predict histological changes in NASH irrespective of steatosis extent? Liver Int. 2015; Ballestri S, Romagnoli D, Nascimbeni F, Francica G, Lonardo A. Role of ultrasound in the diagnosis and treatment of nonalcoholic fatty liver disease and its complications. Expert Rev Gastroenterol Hepatol. 2015 May;9(5):603-27; Ballestri S, Lonardo A, Romagnoli D, Carulli L,
Losi L, Day CP, Loria P. Ultrasonographic fatty liver indicator, a novel score which rules out NASH and is correlated with metabolic parameters in NAFLD. Liver Int. 2012 Sep;32(8):1242-52.

MINOR
How was alcohol consumption quantified and what was the threshold followed to discriminate alcoholic from nonalcoholic fatty liver disease?

REVIEWER
Prof. Aldo Paolicchi
University of Pisa - Italy
Dep. of Translational Research and New Technologies in Medicine and Surgery

REVIEW RETURNED
25-May-2015

GENERAL COMMENTS
The object of this work is to clarify the reciprocal causality between non-alcoholic fatty liver disease (NAFLD) and metabolic syndrome (MetS) to establish if NAFLD is a cause or a consequence of MetS. This manuscript proposes an original approach both in the set up of the design of the study and in the statistical methods. In fact, to my knowledge, this is the first epidemiological study based on two bi-directional longitudinal cohorts, where “new” (for the field) statistical methods have been applied.

The manuscript is extremely rich of data, maybe to many, which hide the most important results. Moreover I think the inclusion/exclusion criteria are not clear as well as the outcomes, which might confound the reader.

A more detailed description of my concerns follows.

***Meaning of the variable “MetS/its component”***
- Pag 4, line 76; Table 1: What does the variable “MetS/its components” exactly mean? Maybe metabolic syndrome AND/OR any of its components?

When this variable is 0, does it mean that metabolic syndrome as well as any of its components is absent? Or at least two components can be present?

***Cohort definition***
I think that cohort definitions are not clear, please clarify the inclusion/exclusion criteria:
- Pag 5, line 101: sub-cohort A is defined as free of “MetS/Its component”, does it mean that any of the four criteria for metabolic syndrome are present or, at the limit, two can be present? What about for cohort B?

From the Introduction I expected two cohorts, one including subjects all with NAFLD but not MetS at baseline and followed up to MetS onset and viceversa for the second one. This concept seems to be reiterated in Materials and Methods (pag 8, line 150 and 156) where statistical methods to study the reciprocal causality pathway from NAFLD to MetS and viceversa are described.

Instead, from the provided description of inclusion criteria (pag 5, line 98-104) it seems that the first cohort include subjects without MetS or any of its components who will develop NAFLD and then MetS, while the second cohort is composed by subjects with one to four MetS components who will develop MetS before than NAFLD.
Which interpretation is the correct one? The second one seems in accordance with Table 2.

It seems to me that the main point of this study is the effect of NAFLD on single MetS components and vice versa rather than the reciprocal causality between NAFLD and MetS (i.e. the co-presence of at least three criteria).

***Alcohol consumption and NAFLD definition***
-Pag 5, line 99: Table 1: Authors should quantify the statement “drinking regularly” as well as the values assigned in Table 1 (i.e 1: seldom, 2: often).

The definition of non-alcoholic fatty liver disease is strictly dependent on alcohol consumption that must be lower than 20 g/day in women and 30 g/day in men (Watanabe S. et al., “Evidence-based clinical practice guidelines for nonalcoholic fatty liver disease/nonalcoholic steatohepatitis” Hepatology Research 2015;45:363-77). Thus alcohol intake must be better specified. If this information is not available in detail it must be discussed in as a limitation of the study.

***Statistical analysis***
- All variables have been reported as mean ± standard deviation independently from the frequency distribution. Non-Gaussian variables (e.g.: triglycerides, fasting glucose, GGT, creatinine, BMI…) should be better presented as median (25th – 74th percentile).

***Results, Tables and Figures***
-pag 11, line 188 Cohort A: 103 people in non-NAFLD group developed MetS, but also NAFLD (before the MetS)? Vice versa for non-MetS group of cohort B?
- I don't understand how incidence density has been calculated, can you explain numbers at denominator of the fractions?

- Table 1: some abbreviations and assignments are missing, e.g. DBP, SBP, TP, A/G, TG, FPG, MCHC, RDW-CV, RDW-SD

- Figure 1 legend: more details are needed to understand the precise meaning of this figure.
- Table 2: sample size does not correspond to the text; DBP and SBP are exchanged; check the WBC unit.
- Table S1: check the age of Sub-cohort B; DBP and SBP are exchanged; check the WBC unit.

- Figure 2: the legend is not clear. If I understand well, the figure represents the RR of developing MetS or one of its components having NAFLD at baseline (white diamond) or the RR of developing NAFLD having MetS or one of its components at baseline (black diamond). Is it correct?
- Table S2, S3: why the total cholesterol is considered, and not triglycerides and HDL-cholesterol, which are two criteria of MetS? Why the variables considered in the two tables are different?

- Table 3: what is the meaning of P(M)% and AR(%); What does 1; 0; 1 in the column “NAFLD”, “condition”, “M”*, respectively mean? Present/absent?
- Table 4: as before
Figure 3/5, S1/2: what is the meaning of the numbers “1” and “2” associated with the variables? Do they refer to two different observations (e.g. baseline and end of the study)?

Figure S2: is SD the same as RDW-SD?

VERSION 1 – AUTHOR RESPONSE

#Response to Reviewer #1:
1. Although utilising more complex statistical methodology, the paper is well written and conveys the findings well. The general reader may struggle with the statistical terms.

#Response: Thanks for your suggestion. We have revise the manuscript by adding related pathophysiological description about NAFLD and MetS in introduction part, to make it more readable for the general reader. Additionally, a thorough language revision by the English mothertongue native, have been made to ensure consistent verb tense usage, smooth phrasing, and correct article use.

#Response to Reviewer #2:

GENERAL COMMENT
1. This manuscript is difficult to follow owing to the use of non-idiomatic English and excessive use of statistical (as opposed to medical) information.

#Response: Thanks for your suggestion. A thorough language revision by the English mothertongue native, have been made to ensure consistent verb tense usage, smooth phrasing, and correct article use.

2. Introduction - The rationale of the study and pre-test hypothesis are poorly detailed. This may be due to incomplete/outdated referencing.

Discussion – This section is too long and fails to illustrate the chief findings and the putative biological basis underlying it.

#Response: We have re-write the introduction and discussion according to your suggestion.

SPECIFIC COMMENT
1. This manuscript needs revision by a mothertongue native. Here follow just a few examples of the many typos and sentences in non-idiomatic English of which the manuscript abounds "In recent years, due to the lifestyle and economic changes, the prevalence of NAFLD and MetS has been rapidly increasing in Chinese populations, which has become one of the major public-health challenges". Written as it is, the sentence includes a typo (tos) and a wrong sentence (it seems as if the Chines populations rather than NAFLD & the MS are a health challenge.

The following sentence "In pathogenesis, insulin resistance (IR) plays a critical role in the development of both NAFLD and MetS, expecting the inherent link between them." Who is the subject of the word "expecting"? Is "to expect" the most appropriate verb to be used here?

Although accumulating epidemiologic evidences support an association between NAFLD and MetS/its components, it remained unclear whether NAFLD was a cause or consequence of MetS". This sentence includes a) the plural of "evidence" which does not exist in idiomatic English and b) two different tenses (support; remained; was ) rather than one (support; remains; is).

Lines 75-88 are quite hard to follow. This occurs owing to poor English and excess statistical details being given in the absence of robust physiopathological basis of why NAFLD should be located in the cascade of physiopathological eventually culminating in the development of NAFLD. The Autors should be encouraged to read those recent studies which fully substantiate how come that NAFLD may be a precursor of the MetS rather than merely its hepatic manifestation (Lonardo A, Ballestri S,

#Response: Thanks very much for your suggestion. We have revised the manuscript according to your comments (see introduction part lines 74-94, and other related lines in the revised manuscript). Additionally, a thorough language revision by the English mother tongue native, have been made to ensure consistent verb tense usage, smooth phrasing, and correct article use.

2. Introduction - The classical view maintains that NAFLD is the hepatic manifestation of the MetS. Stated otherwise, it postulates NAFLD as the consequence of pre-existing MetS. More recent data, however, highlight the bidirectional relationship existing between these 2 different conditions. It think the Authors need to shortly summarize the state of art concerning the relationship between NAFLD and the Metabolic; next they need to state what they expected to find and which specific statistical method they followed.

Discussion – Start the discussion by shortly summarizing the chief findings. Next, discuss if such data are in agreement/conflict with previous studies. Do not forget to highlight novel findings and to succinctly illustrate th eputative biological basis underlying these findings. Conclude by listing strengths and limitations of the study. Among the latter, please include that NAFLD was not subjected to any semiquantitative indices (Ballestri S, Nascimbeni F, Romagnoli D, Lonardo A. Do ultrasonographic semiquantitative indices predict histological changes in NASH irrespective of steatosis extent? Liver Int. 2015; Ballestri S, Romagnoli D, Nascimbeni F, Francica G, Lonardo A. Role of ultrasound in the diagnosis and treatment of nonalcoholic fatty liver disease and its complications. Expert Rev Gastroenterol Hepatol. 2015 May;9(5):603-27; Ballestri S, Lonardo A, Romagnoli D, Carulli L, Losi L, Day CP, Loria P. Ultrasonographic fatty liver indicator, a novel score which rules out NASH and is correlated with metabolic parameters in NAFLD. Liver Int. 2012 Sep;32(8):1242-52.

#Response: Thanks for your suggestion. We have re-write the introduction (lines 74-94) and discussion (lines 266-275, and 312) according to your comments in the revised manuscript.

3. How was alcohol consumption quantified and what was the threshold followed to discriminate alcoholic from nonalcoholic fatty liver disease?

#Response: The regular alcohol intake was defined in revised manuscript (lines 138-143). “Questions about alcohol intake included the type of alcohol consumed, the frequency of alcohol consumption per week and the usual amount per day (≥20 g/day). Based on these questions, alcohol intake was coded as an ordered categorical variable as follows: 0, never; 1, seldom; 2, often, wine; 3, often, beer; 4, often, Chinese spirits; and 5, often, mixed/all types. Persons with a value greater than one were considered regular alcohol users”.

#Response to Reviewer #3:

1. ***Meaning of the variable “MetS/its component”***
   - Pag 4, line 76; Table 1: What does the variable “MetS/its components” exactly mean? Maybe metabolic syndrome AND/OR any of its components? When this variable is 0, does it mean that metabolic syndrome as well as any of its components is absent? Or at least two components can be present?

#Response: Sorry for not defining it clearly. “MetS/its component” means both MetS and its
components (obesity, dyslipidemia, hyperglycemia, and hypertension). These five variables all have dichotomous values, with 1=present and 0=absent. We have revised Table1 and the sentences including “MetS/its component”.

2. ***Cohort definition***
I think that cohort definitions are not clear, please clarify the inclusion/exclusion criteria: - Pag 5, line 101: sub-cohort A is defined as free of “MetS/Its component”, does it mean that any of the four criteria for metabolic syndrome are present or, at the limit, two can be present? What about for cohort B? From the Introduction I expected two cohorts, one including subjects all with NAFLD but not MetS at baseline and followed up to MetS onset and vice versa for the second one. This concept seems to be reiterated in Materials and Methods (pag 8, line 150 and 156) where statistical methods to study the reciprocal causality pathway from NAFLD to MetS and vice versa are described. Instead, from the provided description of inclusion criteria (pag 5, line 98-104) it seems that the first cohort include subjects without MetS or any of its components who will develop NAFLD and then MetS, while the second cohort is composed by subjects with one to four MetS components who will develop MetS before than NAFLD. Which interpretation is the correct one? The second one seems in accordance with Table 2. It seems to me that the main point of this study is the effect of NAFLD on single MetS components and vice versa rather than the reciprocal causality between NAFLD and MetS (i.e. the co-presence of at least three criteria).

#Response: We have revised the cohort definition to make it more clearly (shown in lines 112-122). The second interpretation of cohort definition is right. As shown in Figure 1, the sub-cohort A (the temporal sequence: from NAFLD to MetS) includes the participants with or without NAFLD at baseline to follow-up the incident of MetS, while the sub-cohort B (from NAFLD to MetS) includes the participants with or without NAFLD and its components at baseline to follow-up the incident of NAFLD. In order to perform casual inferences, we excluded the individuals violating the above temporal sequence in sub-cohort A and B, separately.

3. ***Alcohol consumption and NAFLD definition***
-Pag 5, line 99; Table 1: Authors should quantify the statement “drinking regularly” as well as the values assigned in Table 1 (i.e 1: seldom, 2: often).
The definition of non-alcoholic fatty liver disease is strictly dependent on alcohol consumption that must be lower than 20 g/day in women and 30 g/day in men (Watanabe S. et al., “Evidence-based clinical practice guidelines for nonalcoholic fatty liver disease/nonalcholic steatohepatitis”Hepatology Research 2015;45:363-77). Thus alcohol intake must be better specified. If this information is not available in detail it must be discussed in as a limitation of the study.

#Response: The regular alcohol intake was defined in revised manuscript (lines 138-143). “Questions about alcohol intake included the type of alcohol consumed, the frequency of alcohol consumption per week and the usual amount per day (≥20 g/day). Based on these questions, alcohol intake was coded as an ordered categorical variable as follows: 0, never; 1, seldom; 2, often, wine; 3, often, beer; 4, often, Chinese spirits; and 5, often, mixed/all types. Persons with a value greater than one were considered regular alcohol users”.

4. ***Statistical analysis***
- All variables have been reported as mean ± standard deviation independently from the frequency distribution. Non-Gaussian variables (e.g.: triglycerides, fasting glucose, GGT, creatinine, BMI…) should be better presented as median (25th – 74th percentile).

#Response: Thanks for your suggestion. We have revised the section of “statistical analysis” and Tables. The non-Gaussian variables, including triglycerides, GGT and creatinine were presented as...
median (25th, 75th percentile), and the \( p \) values were calculated using non-parametric test (shown in lines 167-171 and Table 2 in the revised manuscript).

5. ***Results, Tables and Figures***
- Pag 11, line 188 Cohort A: 103 people in non-NAFLD group developed MetS, but also NAFLD (before the MetS)? Vice versa for non-MetS group of cohort B?
- I don't understand how incidence density has been calculated, can you explain numbers at denominator of the fractions?

#Response: The incidence density is calculated as incident of new cases at follow-up period divided by the sum of length of time at risk for each individual in the population. In this study, the denominators were expressed in "person years" units. For example, for the 1243 participants with NAFLD at baseline in sub-cohort A, they were followed-up for 3767 person-years during the follow-up from 2005 to 2011.

6. - Table 1: some abbreviations and assignments are missing, e.g. DBP, SBP, TP. A/G, TG, FPG, MCHC, RDW-CV, RDW-SD

#Response: Thanks for your suggestion. Their abbreviations and assignments have been added in the revised manuscript.

7. - Figure 1 legend: more details are needed to understand the precise meaning of this figure.

#Response: Thanks for your suggestion. We revised the legend of Figure 1 to make it self-explanatory.

8. - Table 2: sample size does not correspond to the text; DBP and SBP are exchanged; check the WBC unit.
- Table S1: check the age of Sub-cohort B; DBP and SBP are exchanged; check the WBC unit.

#Response: Sorry for the mistakes. We have revised the sample size in Table 2, and the age in Table S1. The WBC unit should be 10^9/L, we have corrected it throughout the manuscript.

9. - Figure 2: the legend is not clear. If I understand well, the figure represents the RR of developing MetS or one of its components having NAFLD at baseline (white diamond) or the RR of developing NAFLD having MetS or one of its components at baseline (black diamond). Is it correct?

#Response: Yes, it is right. We have re-write the title and legend of Figure 2. Thanks for your suggestion.

10. - Table S2, S3: why the total cholesterol is considered, and not triglycerides and HDL-cholesterol, which are two criteria of MetS? Why the variables considered in the two tables are different?

#Response: Sorry for not describing it clearly. Simple GEE analyses were first performed to select the potential risk factors for MetS in sub-cohort A and NAFLD in sub-cohort B, separately. The variables with \( p \) value less than the significance level 0.05 were then included in the multiple GEE models (results shown in Tables S2 and S3), so the variables included in the two multiple GEE models might be different.

11. - Table 3: what is the meaning of \( P(M)\% \) and \( AR(\%)\); What does 1; 0; 1 in the column "NAFLD", "condition", "M\%", respectively mean? Present/absent?
- Table 4: as before
#Response: We have revised Table 3 and Table 4 as a new Table 3.

M is the abbreviation for MetS and its components. P(M)% is the conditional probability P(M|NAFLD=1)% of MetS and its components (M) given present of NAFLD. AR(%) is the attributable risks, which were calculated as P(M|NAFLD=1)-P(M|NAFLD=0). The values 1 or 0 of “NAFLD”, “M” and “condition” denote present or absent of these variables.

12. - Figure 3/5, S1/2: what is the meaning of the numbers “1” and “2” associated with the variables? Do they refer to two different observations (e.g. baseline and end of the study)?

#Response: Yes, the numbers “1” and “2” associated with the variables denote the status at baseline and the end of follow-up, respectively. We have added it on the legend of Figures.

13. - Figure S2: is SD the same as RDW-SD?

#Response: Sorry for the mistake. We have revised the Figure S2.

**VERSION 2 – REVIEW**

<table>
<thead>
<tr>
<th>REVIEWER</th>
<th>Amedeo Lonardo, M.D.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Italy</td>
<td></td>
</tr>
</tbody>
</table>

**REVIEW RETURNED**

26-Jun-2015

**GENERAL COMMENTS**

This is an innovative proof-of-concept study that NAFLD is both a cause and an effect of the Metabolic syndrome. The study is performed on a very large cohort of individuals and statistics follows novel avenues. As such, this submission has the potential for a large numbers of quotations. However, the manuscript is not ready for publication and extensive English editing needs to be done with the assistance of a mohtertongue expert.

There are many published studies trying to highlight the connections linking Cholesterol with NASH (and the metabolic syndrome). The Authors may be willing to allude to this important line of research in their discussion.

**SPECIFIC COMMENTS**

There mistakes in the agreement of subject + verb eg (Abstract, last sentence) "The finding suggest".

There are statements which are rarely attributable to idiomatic English, e.g. "In pathogenesis", "rather than just hepatic manifestation".

There are unacceptable ideograms in the text, e.g. "NAFLD→MetS or MetS→NAFLD"; "NAFLD→MetS" "dysplipidemia→ hypertension → MetS"

There are very obscure sentences, e.g. "In pathogenesis, as the routine health check-up markers, GGT and persistent NAFLD play important roles in developing MetS, while dyslipidemia is the most important factor in developing NAFLD."; Among 36 causal pathways existed from NAFLD to MetS, the most important causal pathway in
The pathogenesis was that NAFLD led to elevated GGT: "the dominant 301 causal pathway in pathogenesis was beginning with dysliplidemia, and finally resulting in NAFLD." "low density lipoprotein synthesis obstructed with causing triglyceride accumulation in the liver and caused NAFLD". "was assessed by abdominal ultrasonography by experienced radiologists instead of pathologic finding"

Avoid continuous shifts from the past to the present tense throughout the manuscript.


| REVIEWER        | Prof. Aldo Paolicchi  
|-----------------| University of Pisa, Italy.  
| Dep. of Translational Research and New Technologies in Medicine and Surgery |
| REVIEW RETURNED | 14-Jul-2015 |
| GENERAL COMMENTS | The reviewer completed the checklist but made no further comments. |
#Response to Reviewer #2:

GENERAL COMMENT
1. This is an innovative proof-of-concept study that NAFD is both a cause and an effect of the Metabolic syndrome. The study is performed on a very large cohort of individuals and statistics follows novel avenues. As such, this submission has the potential for a large numbers of quotations. However, the manuscript is not ready for publication and extensive English editing needs to be done with the assistance of a mother-tongue expert.

#Response: Thanks for your suggestion. The manuscript has been revised by the English mother-tongue native, so as to ensure consistent verb tense usage, smooth phrasing, and correct article use.

2. There are many published studies trying to highlight the connections linking Cholesterol with NASH (and the metabolic syndrome). The Authors may be willing to allude to this important line of research in their discussion.

#Response: Thanks for your suggestion, we have added this in the discussion part of revised manuscript, lines 302-304 in page 16.

SPECIFIC COMMENTS
3. There mistakes in the agreement of subject + verb eg (Abstract, last sentence) "The finding suggest".

#Response: Sorry for the mistake, we have corrected it in the revised manuscript.

4. There are statements which are rarely attributable to idiomatic English, e.g. "In pathogenesis","rather than just hepatic manifestation".

#Response: Thanks for your comment, we have revised these in the revised manuscript.

5. There are inacceptable ideograms in the text, e.g. "NAFLD→MetS or MetS→NAFLD"; "NAFLD→MetS" "dysliplidemia→ hypertension → MetS"

#Response: Thanks for your suggestion, all these arrows have been replaced into acceptable texts in the revised manuscript.

6. There are very obscure sentences, e.g. "In pathogenesis, as the routine health check-up markers, GGT and persistent NAFLD play important roles in developing MetS, while dyslipidemia is the most important factor in developing NAFLD."; Among 36 causal pathways existed from NAFLD to MetS, the most important causal pathway in pathogenesis was that NAFLD led to elevated GGT. "the dominant 301 causal pathway in pathogenesis was beginning with dysliplidemia, and finally resulting in NAFLD." "low density lipoprotein synthesis obstacled with causing triglyceride accumulation in the liver and caused NAFLD". "was assessed by abdominal ultrasonography by experienced radiologists instead of pathologic finding"

#Response: Thanks for your comments, these sentences have been revised in the revised manuscript.

7. Avoid continuous shifts from the past to the present tense throughout the manuscript.

#Response: Thanks for your suggestion. We have corrected the tense throughout the manuscript.

8. The following statement " if the elevated total cholesterol appeared in the pathway, the probability of MetS would become very low" is worth further comment based on the several published studies. The sentence "This abnormal phenomenon needed further investigation" accordingly should be deleted.

#Response: Thanks for your thoughtful suggestion. We have re-write the discussion according to your comments in the revised manuscript (lines 302-311 in page 16).
“The association between cholesterol and NAFLD (or the metabolic syndrome) has been fairly well established through long-term studies of high levels of serum cholesterol and the incidence of NAFLD, MetS and coronary heart diseases. Surprisingly, in this study, we found that the elevated total cholesterol appeared in the above pathways would result in a lower probability of MetS and NAFLD. This may be not in accord with the conventional viewpoints. Nevertheless, a meta-analysis reported that serum total cholesterol levels were significantly lower in nonalcoholic steatohepatitis (NASH) than in simple steatosis. This study concluded that lower cholesterol levels were independently associated with NASH, in addition to the well-known association with the MS and IR. However, the mechanistic explanations linking a lower cholesterol level with NAFLD and MetS still need further investigation.”

**VERSION 3 – REVIEW**

<table>
<thead>
<tr>
<th>REVIEWER</th>
<th>Amedeo Lonardo, M.D.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Azienda USL - NOCSAE - Baggiovara, Modena, Italy</td>
<td></td>
</tr>
<tr>
<td>REVIEW RETURNED</td>
<td>18-Aug-2015</td>
</tr>
</tbody>
</table>

**GENERAL COMMENTS**
The Authors have accepted most (though not all) of this Reviewer's suggestions. References 67, 68 and 69 are redundant: delete ref. 69.

**VERSION 3 – AUTHOR RESPONSE**
The Authors have accepted most (though not all) of this Reviewer's suggestions. References 67, 68 and 69 are redundant: delete ref. 69..

#Response: Thanks for your comments. We have deleted ref. 69 in the revised manuscript.
Identification of reciprocal causality between non-alcoholic fatty liver disease and metabolic syndrome by a simplified Bayesian network in a Chinese population

Yongyuan Zhang, Tao Zhang, Chengqi Zhang, Fang Tang, Nvjuan Zhong, Hongkai Li, Xinhong Song, Haiyan Lin, Yanxun Liu and Fuzhong Xue

*BMJ Open* 2015 5:
doi: 10.1136/bmjopen-2015-008204

Updated information and services can be found at:
http://bmjopen.bmj.com/content/5/9/e008204

These include:

**Supplementary Material**
Supplementary material can be found at:
http://bmjopen.bmj.com/content/suppl/2015/09/22/bmjopen-2015-008204.DC1

**References**
This article cites 59 articles, 6 of which you can access for free at:
http://bmjopen.bmj.com/content/5/9/e008204#BIBL

**Open Access**
This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See:
http://creativecommons.org/licenses/by-nc/4.0/

**Email alerting service**
Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

**Topic Collections**
Articles on similar topics can be found in the following collections

- Cardiovascular medicine (821)
- Epidemiology (2182)
- General practice / Family practice (678)
- Health informatics (217)
- Public health (2295)

**Notes**

To request permissions go to:
http://group.bmj.com/group/rights-licensing/permissions

To order reprints go to:
http://journals.bmj.com/cgi/reprintform

To subscribe to BMJ go to:
http://group.bmj.com/subscribe/