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)	Identification of antithrombotic drugs related to total joint replacement using anonymised free text notes: a search strategy in the Clinical Practice Research Datalink.
AUTHORS	Nielen, Johannes; van den Bernt, Bart; Boonen, Annelies; Dagnelie, Pieter; Emans, Pieter; Veldhorst, Nicole; Lalmohamed, Arief; van Staa, Tjeerd-Pieter; de Vries, Frank

VERSION 1 - REVIEW

REVIEWER	Amitava Banerjee University of Birmingham. UK
REVIEW RETURNED	29-Jul-2015

GENERAL COMMENTS	Only one very minor change required. In the strengths/limitations section, it should read: "We were unable to determine the specificity and sensitivity of our method."

REVIEWER	Domenico Prisco Dept of Experimental and Clinical Medicine University of Florence Italy
REVIEW RETURNED	02-Aug-2015

GENERAL COMMENTS	This is a useful study to improve the identification of exposure to NOACs or LMWHs in TJR surgery. The merits of the study, as far as I can understand, are clear and well depicted by authors. Strenghts and limitations of the study are reported. However, because the paper is rather technical, an effort to make it more friendly for an average medical readership should be made. Moreover I'm not able the evaluate the correctness of methodology, so that a statistician should be involved in the reviewing.

REVIEWER	Kristina Harris University of Oxford
REVIEW RETURNED	28-Aug-2015

GENERAL COMMENTS	The authors describe a validation study where they design and test a method to extract additional information related to the use of anti-thrombotic drug use from the anonymised free text in the CPRD.

	General comments: The main objective of this paper is to develop and validate the algorithm, but the study mostly focuses on the relevance of clinical findings. The focus of the paper should be the development of the algorithm and the validity of the methods, whilst clinical findings should be separately discussed or even in another paper. I suggest the authors use the approach that was outlined in the paper of Shah et al they cite. Methods: There is not enough information in the methods that would allow the replication of the algorithm. If authors want to keep to the word limit, they should look into including this information as a supplement. As my comment above. If authors want to assess the validity of the test in general, they should calculate sensitivity and specificity and test it on the random sub sample. Please avoid "hypothesize" unless you are using formal hypothesis testing-which is not the case here. Abstract: The objectives section reads as introduction. The first sentence is not necessary. The second sentence talks about aims and the third about hypothesis which was never formally tested. This needs restructuring as well as the Article summary.

REVIEWER	Yohei Kawasaki University of Shizuoka, Japan
REVIEW RETURNED	10-Oct-2015

GENERAL COMMENTS	This paper may fail to reach the aims and scope of BMJ Open as no new methods provided. I think that the big contribution does not include it in this study.

REVIEWER	LM Ho School of Public Health The University of Hong Kong HKSAR
REVIEW RETURNED	13-Oct-2015

GENERAL COMMENTS	1) Although it was addressed in the Discussion that false/true negatives could not be calculated, it is very important to calculate sensitivity and specificity of drug identification algorithm. These parameters are the basis for assessing an algorithm, without which we are unable to assess its accuracy. Is it possible/practical to randomly select some cases (say a thousand) to estimate these important parameters? 2) The meaning of the significance is not clear in Table 2. The footnote said that significant difference was based on the comparison to TJP patients with unknown exposure (line 7, p19), but where are the % for the corresponding TJP patients in Table 2? 3) Regarding that “identified users of NOACs and LMWH users appear to be reasonably similar to patients without identified drug use” (lines 35, p13), please indicate the tables/figures showing this point.

	4) A table summarizing the results in Figures 2 and 3 with p-values seems more concise. 5) Suggest putting the keywords (not just “available upon request” line 52, p14) on the web to improve accessibility. 6) This aim of the study is to design an algorithm for drug identification. Suggest adding a section on this algorithm, ie the procedure of anonymised free text analysis. It seems that it was written under “Selection of comparison groups” (line 43, p.6), but not definitely sure this is the algorithm.

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1
Reviewer Name: Amitava Banerjee

Institution and Country: University of Birmingham. UK

Please state any competing interests or state ‘None declared’: None

Please leave your comments for the authors below:
1) Only one very minor change required. In the strengths/limitations section, it should read:

"We were unable to determine the specificity and sensitivity of our method."

We have changed this, Lines 315-316:
“As a result, were unable to determine the specificity and sensitivity of our method.’

Reviewer: 2
Reviewer Name: Domenico Prisco

Institution and Country:
Dept of Experimental and Clinical Medicine
University of Florence
Italy

Please state any competing interests or state ‘None declared’: None declared

Please leave your comments for the authors below
This is a useful study to improve the identification of exposure to NOACs or LMWHs in TJR surgery. The merits of the study, as far as I can understand, are clear and well depicted by authors. Strengths and limitations of the study are reported.

1) However, because the paper is rather technical, an effort to make it more friendly for an average medical readership should be made. Moreover I'm not able the evaluate the correctness of methodology, so that a statistician should be involved in the reviewing.

We agree with the reviewers comment regarding the “readability” of this paper. However the goal of this paper is to present a new research method. This inherently reflects to the writing style of the

manuscript. Therefore, we lost some of the “readability” in order to describe the methodology correctly.
Nonetheless, we have made an effort to include sentences/sections that reflect the clinical relevance of this study. Introduction: Lines 115.
“In order to assess safety and efficacy of the new oral anticoagulants (NOACs).”
Discussion: Lines 340-342
“Our method is a useful tool to identify exposure to NOACs or LMWHs related to total TKR or THR surgery in the CPRD, and increases statistical power to evaluate potential side effects of these drugs in pharmacoepidemiological studies. “

Reviewer: 3
Reviewer Name: Kristina Harris

Institution and Country: University of Oxford

Please state any competing interests or state ‘None declared’: None declared

Please leave your comments for the authors below
The authors describe a validation study where they design and test a method to extract additional information related to the use of anti-thrombotic drug use from the anonymised free text in the CPRD.

1) General comments: The main objective of this paper is to develop and validate the algorithm, but the study mostly focuses on the relevance of clinical findings. The focus of the paper should be the development of the algorithm and the validity of the methods, whilst clinical findings should be separately discussed or even in another paper. I suggest the authors use the approach that was outlined in the paper of Shah et al they cite.
We agree with the reviewer that this is in essence a methodological study. Therefore most of the writing is generally technical. However, in order to address other clinical researchers we felt to stress the clinical relevance throughout the manuscript. Furthermore, we used some baseline characteristics (clinical findings) as a surrogate measure, since we were unable to calculate sensitivity and specificity.

2) Methods:
There is not enough information in the methods that would allow the replication of the algorithm. If authors want to keep to the word limit, they should look into including this information as a supplement. As my comment above.
In order to provide more information for replication of this study we have now included the used keywords as a supplementary table in the appendix (Appendix Table 1). Furthermore, we have provided the product codes used to determine exposure to NOACs, LMWHs and aspirin and medical codes for total hip (THR) and total knee replacement (TKR) in the CPRD (Appendix Table 2). With the provided information in the method section, these keywords and code lists the analyses should be reproducible.

3) If authors want to assess the validity of the test in general, they should calculate sensitivity and specificity and test it on the random sub sample.
We understand the concern regarding the importance of sensitivity and specificity calculations in these types of studies. Unfortunately, this is practically impossible to do using this database. Therefore, we could only present positive predictive values. In order to evaluate whether documentation/registration of the exposure was differential between patients with a positive hit and patients without a hit we applied two methods. First, we assessed distribution of NOAC and LMWH

use to an external source, the National Joint Registration (NJR). Second, we assessed differences in baseline characteristics of the various exposure groups to the group without a hit. With these surrogate measurements we have generated reassuring information concerning the “potential differential detection of exposure”.

4) Please avoid "hypothesize" unless you are using formal hypothesis testing-which is not the case here.
Both statements (abstract and introduction) regarding a hypothesis have been deleted.

5) Abstract:
The objectives section reads as introduction. The first sentence is not necessary. The second sentence talks about aims and the third about hypothesis which was never formally tested. This needs restructuring as well as the Article summary.

The first and third sentence of the objective section of the abstract have been deleted. It now reads:

“We aimed to design and test a method to extract information on antithrombotic therapy from anonymised free text notes in the Clinical Practice Research Datalink (CPRD).”

Reviewer: 4
Reviewer Name: Yohei Kawasaki

Institution and Country: University of Shizuoka, Japan

Please state any competing interests or state ‘None declared’: None

Please leave your comments for the authors below
1) This paper may fail to reach the aims and scope of BMJ Open as no new methods provided. I think that the big contribution does not include it in this study.

The following is cited from the BMJ Open aims and scope section:

“BMJ Open is a medical journal. We consider papers addressing research questions in … epidemiology. … research methods, …..
Our focus is on research that is relevant to patients and clinicians. All research study types are considered, ... “

To the best of our knowledge this study addresses a new research method to identify drug use in large population based medical databases, such as the Clinical Practice Research Datalink (CPRD) , often used in epidemiological research. Admittedly, we have based our study design on a previous study by Shah and colleagues. However, substantial adaptations have been made in order to ensure applicability for the intended use in de CPRD. In short, this is a new method based on previously proven effective design.

This study is limited to the identification of anticoagulants in free text notes of the CPRD. However this method may be applied to related issues in other databases that include free text notes. Furthermore, the methodology presented in this study may not directly affect patients and clinicians, but is likely to be a basis for further studies assessing safety and efficacy of various hospital prescribed drugs. These future studies could potentially affect regulators, clinicians, and consequently patients.

We therefore feel that this studies, describing a new method, does reach the aims and scope of BMJ Open.

Reviewer: 5
Reviewer Name: LM Ho

Institution and Country:
School of Public Health
The University of Hong Kong
HKSAR

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

Please leave your comments for the authors below
The present study aims to design an algorithm to identify those patients who took antithrombotic drug (drug identification) by analyzing free text notes in the Clinical Practice Research Datalink.

Comments:

1) Although it was addressed in the Discussion that false/true negatives could not be calculated, it is very important to calculate sensitivity and specificity of drug identification algorithm. These parameters are the basis for assessing an algorithm, without which we are unable to assess its accuracy. Is it possible/practical to randomly select some cases (say a thousand) to estimate these important parameters?
We understand the concern regarding the importance of sensitivity and specificity calculations in these types of studies. Unfortunately, this is practically impossible to do using this database. Therefore, we could only present positive predictive values. In order to evaluate whether documentation/registration of the exposure was differential between patients with a positive hit and patients without a hit we applied two methods. First, we assessed distribution of NOAC and LMWH use to an external source, the National Joint Registration (NJR). Second, we assessed differences in baseline characteristics of the various exposure groups to the group without a hit. With these surrogate measurements we have generated reassuring information concerning the "potential differential detection of exposure".

2) The meaning of the significance is not clear in Table 3. The footnote said that significant difference was based on the comparison to TJP patients with unknown exposure (line 7, p19), but where are the % for the corresponding TJP patients in Table 3?
This was not clearly describe in the footnote of Table 3. We have altered this as follows:

Statistically significant different as compared to patients with unknown exposure with regards to chemical thromboprophylaxis ($p<0.05$). THR patients using NOACs, LMWHs, or aspirin were compared to THR patients with unknown exposure. TKR patients using NOACs, LMWHs, or aspirin were compared to TKR patients with unknown exposure.

The difference between NOAC and unknown users were minor. Therefore, there appears to be no reason to believe we were dealing with a deviating group of patients in the NOAC group. Although, more differences were found when comparing LMWH users to unknown exposure, we still believe these differences were minor and that there is no reason to believe we were dealing with a deviating group.

3) Regarding that "identified users of NOACs and LMWH users appear to be reasonably similar to patients without identified drug use" (lines 35, p13), please indicate the tables/figures showing this

point.
Reference to Table 3 has been added to Lines 294, 299, and 335 in the Discussion section.

4) A table summarizing the results in Figures 2 and 3 with p-values seems more concise.

The information presented in Figure 2 and Figure 3 are too different to combine in one table. Figure 2 presents percentages of drug identification according to the used methods, whereas Figures 3a and 3b compare the ratio of LMWH and NOAC use in our method to the ratio of LMWH and NOAC use in the National Joint Registry. This is substantially different and we believe combining this would not improve the quality of the paper.

Furthermore, we feel the information in Figure 2 is visually more appealing when presented in a figure. Nonetheless, we have included a supplementary table (Appendix Table 3) presenting the same information in a table. We leave the decision to the editor to present either Figure 2, Appendix Table 3, or both. Calculating p-values for this table/figure would not be relevant for the research question of this paper. This figure/table is merely included to show that information regarding the prescription of hospital prescribed drugs, such as NOACS, is mainly (>80%) found in free text notes, whereas information regarding GP prescribed drugs, such as, aspirin is mainly recorded by product codes (~ 90%). Highlighting the importance of free text notes when investigating hospital prescribed drugs.

Figure 3a and 3b are replaced by Table 2. Ratio of NOAC/LMWH use is now presented, and differences in sources were calculated by means of the chi square statistic for independent samples ($p<0.05$).

Changes have been made to:

Method section Lines 202-203: by means of the chi square statistic for independent samples ($p<0.05$).

Results section Lines 228-232: Use of NOACs was higher with our method in CPRD as compared to the NJR reports in both THR and TKR patients when ratio of NOAC and LMWH use was compared. NOAC/LMWH ratio was only statistically significantly different in TKR patients in 2009 and 2010. All other groups were not statistically significantly different (Table 2).

Discussion section Lines 286-287: The ratio of NOAC and LMWH use in our CPRD analysis appeared to be different compared to the NJR reports in TKR patients in 2009 and 2010 only.

5) Suggest putting the keywords (not just “available upon request” line 52, p14) on the web to improve accessibility.
Keywords were added as a supplementary table (Appendix Table 1)

6) This aim of the study is to design an algorithm for drug identification. Suggest adding a section on this algorithm, ie the procedure of anonymised free text analysis. It seems that it was written under “Selection of comparison groups” (line 43, p.6), but not definitely sure this is the algorithm.

Indeed, details regarding the algorithm and the procedure were described in the “selection of comparison group” and the “data analysis” section of the meth

VERSION 2 – REVIEW

REVIEWER	LM Ho School of Public Health The University of Hong Kong
REVIEW RETURNED	09-Nov-2015

GENERAL COMMENTS	It is a re-submission. The present study aims to design an algorithm to identify those patients who took antithrombotic drug (drug identification) by analyzing free text notes in the Clinical Practice Research Datalink. The authors have satisfactorily addressed the statistical problems with the previous tables. The details of the searching strategies have been given in the Supplementary file. However one very important comment about sensitivity and specificity has not been handled properly, and its limitation has only been described. It would be useful and clearer if the authors could give the explanation of why these indices could not be computed in the manuscript, just like what had been written in the “Author’s Response”. Without calculating sensitivity and specificity, the use of “validation study” in the title can be misleading because there was no formal validation of the results indeed. According to what have been described in “Selection of comparison groups”, the method described in this section is simply a searching strategy based on a list of keywords and medical/product codes. More precisely, it is “keyword searching", rather than an “algorithm”. The authors may consider the use of a more appropriate term, other than "algorithm".

VERSION 2 – AUTHOR RESPONSE

Reviewer Name: LM Ho

Institution and Country:
School of Public Health
The University of Hong Kong

Please state any competing interests or state ‘None declared’: None

Please leave your comments for the authors below
It is a re-submission. The present study aims to design an algorithm to identify those patients who took antithrombotic drug (drug identification) by analyzing free text notes in the Clinical Practice Research Datalink. The authors have satisfactorily addressed the statistical problems with the previous tables. The details of the searching strategies have been given in the Supplementary file.

1) However one very important comment about sensitivity and specificity has not been handled properly, and its limitation has only been described. It would be useful and clearer if the authors could give the explanation of why these indices could not be computed in the manuscript, just like what had been written in the “Author’s Response”.

We added the following to the discussion. The wording is similar to what has been described in our first “Authors Response”.
Discussion lines 314-325:
“Our study also had limitations. We were practically unable to calculate false or true negatives. As a result, were unable to determine the specificity and sensitivity of our method. In order to evaluate whether documentation of the exposure was differential between patients with a positive hit and patients without a hit we applied two methods. First, we assessed distribution of NOAC and LMWH

use to an external source, the National Joint Registration (NJR). Second, we assessed differences in baseline characteristics of the various exposure groups to the group without a hit. With these surrogate measurements we have generated reassuring information concerning the potential differential detection of exposure”

2) Without calculating sensitivity and specificity, the use of “validation study” in the title can be misleading because there was no formal validation of the results indeed.

The title has been changed and no longer contains “validation study”. It now reads: “Identification of antithrombotic drugs related to total joint replacement using anonymised free text notes: a search strategy in the Clinical Practice Research Datalink.”

3) According to what have been described in “Selection of comparison groups”, the method described in this section is simply a searching strategy based on a list of keywords and medical/product codes. More precisely, it is “keyword searching", rather than an “algorithm”. The authors may consider the use of a more appropriate term, other than "algorithm".

When referring to our study, the word “algorithm” has been replaced by “search strategy” throughout the manuscript . When referring to other studies, the term “algorithm” was retained.