Endorsement of the TRIPOD statement and the reporting of studies developing contrast-induced nephropathy prediction models for the coronary angiography/percutaneous coronary intervention population: a cross-sectional study

Simeng Miao,1,2 Chen Pan,1 Dandan Li,1 Su Shen,1 Aiping Wen

ABSTRACT
Objective Clear and specific reporting of a research paper is essential for its validity and applicability. Some studies have revealed that the reporting of studies based on the clinical prediction models was generally insufficient based on the Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD) checklist. However, the reporting of studies on contrast-induced nephropathy (CIN) prediction models in the coronary angiography (CAG)/percutaneous coronary intervention (PCI) population has not been thoroughly assessed. Thus, the aim is to evaluate the reporting of the studies on CIN prediction models for the CAG/PCI population using the TRIPOD checklist.

Methods PubMed and Embase were systematically searched from inception to 30 September 2021. Only the studies on the development of CIN prediction models for the CAG/PCI population were included. The data were extracted into a standardised spreadsheet designed in accordance with the ‘TRIPOD Adherence Assessment Form’. The overall completeness of reporting of each model and each TRIPOD item were evaluated, and the reporting before and after the publication of the TRIPOD statement was compared. The linear relationship between model performance and TRIPOD adherence was also assessed.

Results We identified 36 studies that developed CIN prediction models for the CAG/PCI population. Median TRIPOD checklist adherence was 60% (34%–77%), and no significant improvement was found since the publication of the TRIPOD checklist (p=0.770). There was a significant difference in adherence to individual TRIPOD items, ranging from 0% to 100%. Moreover, most studies did not specify critical information within the Methods section. Only 5 studies (14%) explained how they arrived at the study size, and only 13 studies (36%) described how to handle missing data. In the Statistical analysis section, how the continuous predictors were modelled, the cut-points of categorical or categorised predictors, and the methods to choose the cut-points were only reported in 7 (19%), 6 (17%) and 1 (3%) of the studies, respectively. Nevertheless, no relationship was found between model performance and TRIPOD adherence in both the development and validation datasets (r=−0.260 and r=−0.069, respectively).

Conclusions The reporting of CIN prediction models for the CAG/PCI population still needs to be improved based on the TRIPOD checklist. In order to promote further external validation and clinical application of the prediction models, more information should be provided in future studies.

INTRODUCTION
Complete and transparent reporting is fundamental in health-related research. Clear, detailed reporting helps the reader understand how a study was designed and conducted, judge the reliability of the findings, and allows for the replication of the study methods and procedures within clinical practice.1 As Professor Douglas Altman said, ‘Readers should not have to infer what was probably done, they should be told explicitly’.2
Many reporting guidelines have been developed for various types of studies in order to improve the reporting of health research, and can be found on the website at www.equator-network.org. The ‘Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD)’ statement was reported in 2015 and involved a checklist of 22 items considered essential for informative reporting of diagnostic or prognostic prediction model studies. Some studies used the TRIPOD checklist to assess the reporting of published prediction models. The results were unsatisfactory; for example, a recent study of 170 published prediction models showed the median compliance of the TRIPOD checklist items was only 44%.

Contrast-induced nephropathy (CIN) is an acute decline in kidney function caused by the use of contrast agents, and is described as the third most common cause of new acute kidney injury in hospitalised patients. Patients with CIN are at a higher risk for death, long-term hospitalisation and other adverse outcomes, including early or late cardiovascular events. Since there is no definitive treatment, it is ideal for the early screening of high-risk patients as well as taking measures to prevent CIN. With the development of imaging medicine and the application of interventional diagnosis and treatment technologies, an increasing number of patients receiving coronary angiography (CAG)/percutaneous coronary intervention (PCI), and CIN is one of the major complications when undergoing CAG and PCI due to the use of iodine contrast agents. Some CIN prediction models have been developed in the CAG/PCI population. However, there are different opinions on the recommendations for CIN prediction models within the guidelines. The Cardiological Society of India practice guidelines in 2012 recommended using the Mehran risk score to screen for high-risk patients. In contrast, CIN prediction models were not recommended in either the European Society of Urogenital Radiology Contrast Medium Safety Committee guidelines in 2018 or the guideline on the use of iodinated contrast media in patients with kidney disease in 2020. The guidelines pointed out that the usefulness of the models had not been thoroughly investigated and still needed external validation. Clear and detailed reporting is essential to facilitate a thorough investigation and further external validation. However, to our knowledge, the reporting of the studies on CIN prediction models for the CAG/PCI population has not been assessed. Therefore, our primary purpose is to thoroughly evaluate the reporting of these studies on CIN prediction models for the CAG/PCI population using the TRIPOD checklist.

METHODS

Study design
A cross-sectional study was conducted by evaluating the reporting of studies on developing CIN prediction models for the CAG/PCI population.

Search strategy and study selection
We systematically searched for studies on CIN prediction models in PubMed and Embase with predefined search terms from inception to 30 September 2021 (online supplemental file 1). Only the studies on developing CIN prediction models for the CAG/PCI population were included. The studies on external validity, or those that evaluated the incremental value of adding a predictor to an existing model were excluded. Related reviews and references of the original articles were also checked to identify any missed studies. Only English publications were included.

Data extraction
Data were extracted into a standardised spreadsheet, designed based on the ‘TRIPOD Adherence Assessment Form’, which can be found on the TRIPOD statement website (www.tripod-statement.org/). Only the items involving the predictive model development were included in the spreadsheet. For the items containing subitems, the information of the subdivisions was extracted. Moreover, the following data of the studies were also extracted, including the year of publication, country of the first author, study population, single or multi-centre study, the sample size of the development/validation dataset, model performance (discrimination and calibration) in development/validation datasets. Study selection and data extraction were conducted by two independent reviewers (SM, CP); any discrepancies were resolved by further discussion with a third reviewer (DL).

Data analyses
Based on the extracted data elements, completeness of reporting of each TRIPOD item, overall completeness of reporting of each model and overall completeness of reporting of each TRIPOD item were assessed. The requested information for all elements of the TRIPOD items were assessed in order to evaluate the completeness of the reporting of each TRIPOD item. If all required information was available, the reporting of the TRIPOD item was judged to be complete. For the elements of TRIPOD items 4b, 5a, 6a and 7a, a reference to information in another article was considered acceptable. The subportions that were regarded as ‘Not applicable’ were excluded in the evaluation.

The sum of the adhered TRIPOD items was divided by the total number of applicable TRIPOD items in order to evaluate the overall completeness of reporting of each model. The total number of applicable TRIPOD items or subitems for model development is 30. Although not included in the overall scoring, the supplementary information data (item 21) was extracted.

The number of studies that adhered to a specific TRIPOD item were divided by the total number of studies in which the specific TRIPOD item was applicable in order to assess the overall completeness of the reporting of each TRIPOD item.
Overall completeness of reporting was compared between the studies before 2015 and after 2015 (including 2015) using a Mann-Whitney U test in order to evaluate whether there was an improvement of the reporting since the publication of the TRIPOD statement (January 2015). For the models which reported the area under the curve (AUC) in the Results section, linear regression was applied to investigate the relationship between the model performance and the adherence to the TRIPOD checklist. IBM SPSS V.26.0 was used, and a p value <0.05 was considered statistically significant.

**Patient and public involvement statement**
Patients and/or the public were not involved.

**RESULTS**

One thousand four hundred and five citations were identified via an electronic search and 40 potentially eligible articles were retrieved for a full-text screen. Finally, 36 studies developing CIN prediction models for the CAG/PCI population were included (figure 1). Four of the included studies were multi-centre studies. Most of the studies were conducted in China (20 studies), followed by the USA (6 studies), Italy (3 studies), Greece (2 studies) and one study each in Japan, India, Kuwait, Turkey and Thailand. The sample size of the studies ranged from 208 to 947,012. Detailed information is shown in table 1.

**Overall completeness of reporting of each model**

The data of reporting of each TRIPOD item and the subportions are illustrated within online supplemental file 2. None of the models reported all of the TRIPOD items. Overall, the median TRIPOD adherence was 60% (34%–69%) and 61% (40%–77%), respectively. No statistically significant differences were found (p=0.770).

**Overall completeness of reporting of each TRIPOD item**

The completeness of reporting of each TRIPOD item varied. Eleven items were reported in more than 80% of the studies, six items were reported in 60%–80% of the studies. Thirteen items were reported in less than 60% of the studies, among which seven items were reported in less than 20% of the studies. Details are illustrated in figure 2 and online supplemental file 2. The most noteworthy findings for each section of the TRIPOD checklist (title and abstract, introduction, methods, results, discussion, and other information) are described below.

**Title (item 1) and abstract (item 2)**

Only seven studies (19%) completely reported the elements required by the title. All of the studies presented the outcomes to be predicted. The words of the prediction/risk prediction/prediction model/risk scores and the target population were included in more than 90% of the studies. However, only seven studies (19%) contained the words developing or development.

None of the studies completely reported the required elements within the abstract. The summary of objectives, participants, sample size, predictors, outcome, number of events, statistical analysis, results for model discrimination and conclusions were reported in 75%–97% of the studies. Only a minimal number of studies reported the study design (33%), setting (36%) and results of the model calibration (19%).

**Introduction (item 3)**

The medical context and rationale for developing the models (item 3a) and specific objectives (item 3b) were explained in 78% and 97% of the studies, respectively.

**Methods (items 4–11)**

There were 15 items or subitems in the Methods section. Four of the items were reported completely in more than 90% of the studies, including key study dates (item 4b, 92%), eligibility criteria for participants (item 5b, 94%), clear definition of the outcomes (item 6a, 97%) and detailed predictor definitions (item 7a, 92%). Five items were reported completely in 60%–80% of the studies, including study design or source of data (item 4a, 75%), essential elements of the study setting (item 5a, 78%), the details of the treatments received (item 5c, 61%), all measures used to assess model performance (item 10d, 69%) and details on how risk groups were created (item 11, 67%). The other six items were only thoroughly reported in less than 40% of the studies, including actions to blind the assessment of the outcomes (item 6b, 11%), predictors (item 7b, 11%), explaining how the study size was arrived at (item 8, 14%), describing how missing data were handled (item 9, 36%), explaining how predictors were handled (item 10a, 3%) and specifying the type of statement. Their median TRIPOD adherence was 60% (34%–69%) and 61% (40%–77%), respectively. No statistically significant differences were found (p=0.770).
Table 1  The characteristics of the included studies

<table>
<thead>
<tr>
<th>No.</th>
<th>Study</th>
<th>Country</th>
<th>Population</th>
<th>Single/multi-centre</th>
<th>Sample size</th>
<th>Model performance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Development dataset</td>
<td>Validation dataset</td>
<td>Development dataset</td>
</tr>
<tr>
<td>1</td>
<td>Mehran et al^6</td>
<td>USA</td>
<td>PCI</td>
<td>Single</td>
<td>5571</td>
<td>2786</td>
</tr>
<tr>
<td>2</td>
<td>Bartholemew et al^7</td>
<td>USA</td>
<td>PCI</td>
<td>Single</td>
<td>10 481</td>
<td>9998</td>
</tr>
<tr>
<td>3</td>
<td>Marenzi et al^18</td>
<td>Italy</td>
<td>PCI</td>
<td>Single</td>
<td>208</td>
<td>—</td>
</tr>
<tr>
<td>4</td>
<td>Brown et al^19</td>
<td>USA</td>
<td>PCI</td>
<td>Single</td>
<td>11 141</td>
<td>—</td>
</tr>
<tr>
<td>5</td>
<td>Ghani and Tohamy^20</td>
<td>Kuwait</td>
<td>PCI</td>
<td>Single</td>
<td>247</td>
<td>100</td>
</tr>
<tr>
<td>6</td>
<td>Maioli et al^35</td>
<td>Italy</td>
<td>CAG/PCI</td>
<td>Single</td>
<td>1218</td>
<td>502</td>
</tr>
<tr>
<td>7</td>
<td>Chong et al^36</td>
<td>China</td>
<td>PCI</td>
<td>Single</td>
<td>770</td>
<td>—</td>
</tr>
<tr>
<td>8</td>
<td>Fu et al^37</td>
<td>China</td>
<td>PCI</td>
<td>Single</td>
<td>668</td>
<td>277</td>
</tr>
<tr>
<td>9</td>
<td>Tziakas et al^38</td>
<td>Greece</td>
<td>PCI</td>
<td>Single</td>
<td>488</td>
<td>200</td>
</tr>
<tr>
<td>10</td>
<td>Gurm et al^39</td>
<td>USA</td>
<td>PCI</td>
<td>Multi</td>
<td>48 001</td>
<td>20 572</td>
</tr>
<tr>
<td>11</td>
<td>Gao et al^40</td>
<td>China</td>
<td>CAG/PCI</td>
<td>Single</td>
<td>2764</td>
<td>1181</td>
</tr>
<tr>
<td>12</td>
<td>Tsai et al^41</td>
<td>USA</td>
<td>PCI</td>
<td>Multi</td>
<td>662 504</td>
<td>284 508</td>
</tr>
<tr>
<td>13</td>
<td>Chen et al^42</td>
<td>China</td>
<td>PCI</td>
<td>Single</td>
<td>1500</td>
<td>1000</td>
</tr>
<tr>
<td>14</td>
<td>Victor et al^43</td>
<td>India</td>
<td>PCI</td>
<td>Single</td>
<td>900</td>
<td>300</td>
</tr>
<tr>
<td>15</td>
<td>Liu et al^44</td>
<td>China</td>
<td>PCI</td>
<td>Single</td>
<td>495</td>
<td>233</td>
</tr>
<tr>
<td>16</td>
<td>Brown et al^45</td>
<td>USA</td>
<td>CAG/PCI</td>
<td>Multi</td>
<td>115 633</td>
<td>—</td>
</tr>
<tr>
<td>17</td>
<td>Liu et al^46</td>
<td>China</td>
<td>PCI</td>
<td>Single</td>
<td>356</td>
<td>273</td>
</tr>
<tr>
<td>18</td>
<td>Ji et al^47</td>
<td>China</td>
<td>PCI</td>
<td>Single</td>
<td>565</td>
<td>240</td>
</tr>
<tr>
<td>19</td>
<td>Inohara et al^48</td>
<td>Japan</td>
<td>PCI</td>
<td>Multi</td>
<td>3957</td>
<td>1979</td>
</tr>
<tr>
<td>20</td>
<td>Lazaros et al^49</td>
<td>Greece</td>
<td>PCI</td>
<td>Single</td>
<td>348</td>
<td>—</td>
</tr>
<tr>
<td>21</td>
<td>Lin et al^50</td>
<td>China</td>
<td>CAG/PCI</td>
<td>Single</td>
<td>461</td>
<td>231</td>
</tr>
<tr>
<td>22</td>
<td>Yin et al^51</td>
<td>China</td>
<td>CAG/PCI</td>
<td>Single</td>
<td>8800</td>
<td>—</td>
</tr>
<tr>
<td>23</td>
<td>Lian et al^52</td>
<td>China</td>
<td>CAG</td>
<td>Single</td>
<td>759</td>
<td>527</td>
</tr>
<tr>
<td>24</td>
<td>Duan et al^53</td>
<td>China</td>
<td>CAG/PCI</td>
<td>Single</td>
<td>1076</td>
<td>701</td>
</tr>
<tr>
<td>25</td>
<td>Hu et al^54</td>
<td>China</td>
<td>PCI</td>
<td>Single</td>
<td>192</td>
<td>—</td>
</tr>
<tr>
<td>26</td>
<td>Guo et al^55</td>
<td>China</td>
<td>CAG</td>
<td>Single</td>
<td>245</td>
<td>—</td>
</tr>
<tr>
<td>27</td>
<td>Fan et al^56</td>
<td>China</td>
<td>PCI</td>
<td>Single</td>
<td>57 630</td>
<td>24 656</td>
</tr>
<tr>
<td>28</td>
<td>Zeng et al^57</td>
<td>China</td>
<td>CAG/PCI</td>
<td>Single</td>
<td>771</td>
<td>386</td>
</tr>
<tr>
<td>29</td>
<td>Koowattanatianchai et al^38</td>
<td>Thailand</td>
<td>PCI</td>
<td>Single</td>
<td>217</td>
<td>1000</td>
</tr>
</tbody>
</table>

Continued
model, all model-building procedures, and the method used for internal validation (item 10b, 0%).

More specifically, in the Statistical analysis section, how the continuous predictors were modelled, the cut-points of categorical or categorised predictors, and the methods used to choose the cut-points were only reported in 7 (19%), 6 (17%) and 1 (3%) of the studies, respectively. The approach used for predictor selection before modelling was described in only two studies (6%). None of the studies clearly described the testing of the interaction terms. Other performance measures, including predictive values, sensitivity, specificity, AUC difference, net reclassification improvement and integrated discrimination improvement were described in only seven studies (19%).

Results (items 13–16)
There were seven items or subitems in the Results section. Three of the items were entirely reported in more than 80% of the studies, including specifying the number of participants and outcome events in each analysis (item 14a, 100%), the unadjusted associations between each predictor and outcome (item 14b, 81%), as well as explaining how to use the prediction model (item 15b, 92%). The other four items were reported completely in less than 50% of the studies, including describing the flow of participants through the study (item 13a, 33%), the characteristics of the participants (item 13b, 36%), presenting the full prediction model to allow predictions for individuals (item 15a, 25%) and reporting performance measures for the prediction model (item 16, 47%). The results of discrimination were reported in more studies (94%) than that of calibration (64%).

Discussion (items 18–20) and other information (items 21, 22)
Reporting of the discussion section was generally good. Limitations of the study were discussed in 35 studies (item 18, 97%) and overall interpretation of the results was illustrated in all of the studies (item 19b, 100%). Potential clinical use of the model and the implications for future research were discussed in 32 studies (item 20, 89%).
Supplementary resources were provided in nine studies (item 21, 25%). The source of funding and the role of the funders for the study was reported in 14 studies (39%).

Relationship between TRIPOD adherence and performance of the model
There were 23 and 28 studies reporting the discrimination within the development and validation datasets, respectively, which was expressed using the AUC. Median AUC values in the development and validation datasets were 0.82 (ranging from 0.70 to 0.96) and 0.82 (ranging from 0.61 to 0.95), respectively.

The linear correlation of AUC versus TRIPOD adherence was not statistically significant both in the development and the validation datasets, \( r = -0.260 \) (\( p = 0.231 \)), \( r = -0.069 \) (\( p = 0.728 \)), respectively (figure 3).

DISCUSSION
This study assessed the reporting of current CIN prediction models for the CAG/PCI population according to the TRIPOD checklist. The results were somewhat disappointing. None of the studies reported all of the items. Median adherence to the TRIPOD checklist was only 60%, and no significant improvement was demonstrated after the publication of the TRIPOD statement. Moreover, there was a significant difference in adherence to individual TRIPOD items, ranging from 0% to 100%. Nevertheless, no relationship was found between model performance and TRIPOD adherence.

Comparison with other studies
Other studies evaluated the reporting of prediction models using the TRIPOD checklist,6–8 23 and their results were similar to our study. Heus et al6 selected 10 journals with the highest impact factor within 37 clinical domains. He finally included a total of 146 publications, from which 170 were evaluated regarding the prediction model studies that included model development, external validation, incremental value, combined development or external validation of the same model, and found that the median adherence to TRIPOD was 44% (16%–81%). Park et al7 assessed the reporting of 77 radiomics studies on both model development and external validation, and observed a mean adherence of 57.8% (33%–78%). Jiang et al8 evaluated the reporting of 27 melanoma prediction model studies, and found that the median adherence was 61% (10%–93%). Yang et al23 evaluated the reporting of 22 external validation studies for hepatocellular carcinoma prediction models, and found the compliance with TRIPOD ranged from 59% to 90%, with a median of 74%. In general, there is still much room for the improvement of the reporting of prediction models.

In addition, we identified some poorly reported items, including the blind assessment of outcome and predictors (items 6b and 7b, 11%), handling of missing data (item 9, 36%), measures for model performance (item 10d, 69%) in the Methods section and regression coefficients for each predictor in the model and the intercept (item 15a, 25%), and performance measures (item 16, 47%) in the Results section. These were similar to other studies.6–8 23 For example, Jiang et al8 also found that a limited number of studies described the handling of missing data (item 9, 26%), measures for model performance (item 10d, 37%), model specifications (item 15a, 46%) and model performance (item 16, 26%).

Implications for practice and future research
Many clinical prediction models have been developed in recent years,24–26 and only a very limited number of prediction models,24 such as Framingham, QRISK2 and CHADS risk scores, have been widely used in clinical practice.27–29 One of the reasons may be that there are reporting deficiencies in publications regarding the development and validation of prediction models.23 Incomplete reporting may limit the application of a prediction model.14

The Methods section is one of the most important sections for research articles, allowing other researchers to judge the rationality of the research design and facilitate other researchers in verifying the results.30 However, some important information was not specified in most
of our included studies, including blind assessment of outcome and predictors (items 6b and 7b, 11%), handling of missing data (item 9, 36%) and measures for model performance (item 10d, 69%). More attention should be paid to the reporting of these information in future research.

In prediction model studies, blind assessment of outcome and predictors is recommended to reduce potential bias. However, the outcome or predictor requires no subjective judgments or assessments (eg, all-cause mortality as an outcome; objective predictors such as age, sex and quantitative laboratory values), blind assessment may not be an issue. This might be the reason for our study's low adherence to the blind assessments of outcomes and predictors (four studies, 11%). The outcome 'CIN' and all predictors in the included studies were objective indicators. Nevertheless, we think it is necessary to mention whether there are blind assessments in the study; if no blind assessments are conducted, the reasons should be specified.

Almost all prediction model studies have some missing data of the outcomes and predictors, especially for the studies that derive models based on retrospective cohorts. A common approach to handling the missing data is to exclude individuals with missing values on any variables and perform a complete case analysis. However, it may significantly reduce the sample size, and lead to biased results when the remaining individuals without missing data are not representative of the entire original study sample. Multiple imputations may be a better choice to handle the missing data, which could minimise bias that may often result from excluding such patients. Additionally, it remains valid even if the proportion of missing data is large. However, the reports often do not clarify how the missing data are handled. Only 13 studies (36%) described missing data in our study.

Model performance is essential for applying a prediction model, which is usually assessed by two measures: discrimination and calibration. Discrimination is usually quantified by calculating the area under the receiver operating characteristic curve. Calibration can be assessed via a calibration plot, calibration slope or intercept, calibration table, Hosmer-Lemeshow test and an O/E ratio. Nevertheless, only 25 studies (69%) described both discrimination and calibration measures in our study.

Furthermore, the Results section is one of the most important parts within research articles for the judgement of whether the model should be considered for clinical use. However, some important information was not provided in most of the reports in our study. For instance, only nine studies (25%) described both regression coefficients for each predictor in the model and the intercept. Furthermore, only 17 studies (47%) described both the results of discrimination (with CIs) and calibration measures. The reporting of this information needs to be improved in future research.

It should be noted that no significant relationship was observed between model performance and TRIPOD adherence, which was consistent with other research's results. The results are not surprising, because the TRIPOD statement is just a reporting guideline, which guides authors to report their research more clearly. Reporting quality cannot reflect model performance; however, our findings suggest that incomplete reporting of the models with excellent performance may lead to the absence of external validation and limit their clinical application. Only 18 of the included 36 models in our study were validated externally, and only the Mehran score has been validated by multiple studies. The reporting may be one of the reasons for the limited external validation of the models.

**Strengths and limitations**

To our knowledge, this is the first study to evaluate the reporting of studies on the development of CIN prediction models within the CAG/PCI population. We systematically searched the literature; therefore, a relatively comprehensive picture could be presented. Furthermore, we carried out a detailed analysis to gain further information on the reporting of the TRIPOD items and subitems. However, we restricted the language of the publications to English. This might exclude studies published in other languages, reporting of which may be potentially sufficient.

**CONCLUSION**

Based on the TRIPOD checklist, the reporting of CIN prediction models for the CAG/PCI population still has room for improvement. In order to promote further external validation and clinical application of the prediction models, more specific information should be provided in future studies.

**Acknowledgements** The authors thank AAM Academic Services (www.aimeditor.com) for the English language editing and review services.

**Contributors** AW was responsible for the overall content as the guarantor. All authors contributed to the design of the study. SM, CP and DL conducted study selection and data extraction. SM analysed the data and wrote the manuscript. AW and SS revised the manuscript. All authors read and approved the final manuscript.

**Funding** This study was supported by the Digestive Medical Coordinated Development Center of Beijing Municipal Administration of Hospitals (No. XXZD06).

**Competing interests** None declared.

**Patient consent for publication** Not applicable.

**Ethics approval** This study does not involve human participants.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data availability statement** All data relevant to the study are included in the article or uploaded as supplementary information.

**Supplemental material** This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.
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, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc-4.0/.

ORCID iD
Aiiping Wen http://orcid.org/0000-0002-9848-9953

REFERENCES


