Baseline Level and Reduction in PaCO2 are Associated with the Treatment Effect of Long-Term Home Noninvasive Positive Pressure Ventilation in Stable Hypercapnic Patients with COPD: A Systematic Review and Meta-Analysis of Randomized Controlled Trials

Zhipeng Wu1,*, Zujin Luo2,*, Zengtao Luo2, Jingyi Ge3, Jiawei Jin4, Zhixin Cao2, Yingmin Ma1

1Department of Respiratory and Critical Care Medicine, Beijing Youan Hospital, Capital Medical University, Beijing, People’s Republic of China; 2Department of Respiratory and Critical Care Medicine, Beijing Chaoyang Hospital, Capital Medical University, Beijing, People’s Republic of China; 3Department of Respiratory Medicine, Beijing Rehabilitation Hospital, Capital Medical University, Beijing, People’s Republic of China; 4The Clinical Research Center, Beijing Chaoyang Hospital, Capital Medical University, Beijing, People’s Republic of China

*These authors contributed equally to this work

Correspondence: Yingmin Ma, Department of Respiratory and Critical Care Medicine, Beijing Youan Hospital, Capital Medical University, Beijing, People’s Republic of China, Tel +86 10-13501185982, Email mayingmin@ccmu.edu.cn

Purpose: The evidence of long-term home noninvasive positive pressure ventilation (LTHNIPPV) in patients with stable hypercapnic chronic obstructive pulmonary disease (COPD) is controversial. In this meta-analysis study, we sought to establish whether a baseline level and reduction in partial pressure of arterial carbon dioxide (PaCO2) were associated with the treatment effect of LTHNIPPV in these patients.

Patients and Methods: Six electronic databases were comprehensively searched from January 1980 until June 2020. Randomized clinical trials (RCTs) comparing LTHNIPPV with control treatment were included. Two authors independently extracted data, assessed the study quality, and used the GRADE approach to evaluate evidence quality. The main outcome was mortality.

Results: Nineteen studies involving 1482 patients (LTHNIPPV, n = 730; control, n = 752) were included. LTHNIPPV significantly reduced mortality (relative risk [RR] = 0.76; 95% confidence interval [CI]: 0.61–0.95; p = 0.02; I² = 14%), the frequency of hospital admissions, PaCO2, and improved partial pressure of oxygen (PaO2) compared to control treatment. LTHNIPPV also relieved dyspnea and improved exercise capacity and health-related quality of life (HRQL) but showed no significant benefit for improving the forced expiratory volume in one second in predicted (FEV1% pred). Subgroup analysis revealed that the baseline level and reduction in PaCO2 were associated with decreased mortality (baseline PaCO2 ≥ 55 mmHg RR = 0.69, P = 0.02; vs baseline PaCO2 < 55 mmHg RR = 0.87, P = 0.32; and higher dPaCO2 RR = 0.42, P < 0.0001; vs lower dPaCO2 RR = 0.91, P = 0.38).

Conclusion: LTHNIPPV significantly reduced mortality. The baseline level and reduction in PaCO2 were associated with the treatment effect of LTHNIPPV in patients with stable hypercapnic COPD. Large-scale, multicenter RCTs are needed to confirm our results.

Keywords: COPD, long-term home noninvasive positive pressure ventilation, mortality, meta-analysis study, RCTs, stable hypercapnic

Introduction

Chronic obstructive pulmonary disease (COPD) is a common disease with nonreversible airflow limitation and persistent respiratory symptoms, such as cough. COPD is a primary cause of morbidity and mortality worldwide and represents a great social and economic burden for humans. 1–3
Noninvasive positive pressure ventilation (NIPPV) is recommended as the first ventilation method for patients with COPD with acute exacerbation, but for stable patients with COPD, the evidence is still controversial. The latest GOLD guideline (2019) reported that the evidence for regular use of long-term home NIPPV (LTHNIPPV) in stable COPD patients remains insufficient. However, the guideline also emphasized that LTHNIPPV can probably be used in patients with daytime hypercapnia. Studies have found that patients with stable hypercapnia have poor prognosis compared to those with normocapnia and reversible hypercapnic COPD.

Randomized clinical trials (RCTs) comparing LTHNIPPV with control treatment in patients with stable hypercapnic COPD have been conducted. Köhnlein et al. found that LTHNIPPV could reduce mortality, but Murphy et al. found that the mortality rate was not significantly different between the LTHNIPPV group and the control group. Previous meta-analysis studies combined these RCTs and showed that LTHNIPPV reduced PaCO2 but did not decrease all-cause mortality or improve PaO2 compared to control treatment for patients with stable hypercapnic COPD. The most recent meta-analysis study combined RCTs and observational studies and found that LTHNIPPV reduced all-cause mortality compared to control treatment for patients with stable hypercapnic COPD. However, it remains unclear whether baseline level and reduction in PaCO2 are associated with the treatment effect of NIPPV on reducing mortality.

Therefore, we performed a comprehensive systematic meta-analysis study comparing NIPPV with control treatment for patients with stable hypercapnic COPD. The analysis had no language restrictions and included more RCTs than previous meta-analysis studies. The main outcome was all-cause mortality, and the secondary outcomes included frequency of hospital admissions, gas exchange, FEV1% pred, 6-min walk distance (6MWD), dyspnea, and health-related quality of life (HRQL). Subgroup analysis based on the baseline level of PaCO2 and the degree of dPaCO2 was performed to establish whether the baseline level of PaCO2 and reduction in PaCO2 were associated with treatment effect of LTHNIPPV in patients with stable hypercapnic COPD.

**Materials and Methods**

We conducted this meta-analysis study in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement and the Cochrane Handbook.

**Literature Search Strategy**

We conducted a comprehensive search of six databases: PubMed, EMBASE, Cochrane Central Register, China National Knowledge Infrastructure (CNKI), WANFANG DATA, and Wei Pu (China Science and Technology Journal Database). The publication time was limited from January 01, 1980, to June 30, 2020, and this study had no language restrictions. Three items were included in the search terms and combined with the Boolean operator “AND”: Patients (including “nasal ventilat*” or “COPD,” “Chronic Obstructive Pulmonary Disease” or “COAD”); intervention (including “mechanical ventilat*” or “noninvasive ventilat*” or “noninvasive ventilat*”); and study design (including RCTs). The detailed search strategies of all databases are presented in Appendix 1. The references of each included study were also searched for additional papers.

**Inclusion/Exclusion Criteria and Study Selection**

Types of studies: RCTs were included, and any other type of study was excluded.

Patients: Adults > 18 years old, diagnosed with moderate to severe COPD according to GOLD 2019, in a stable hypercapnic state.

Intervention: Patients in treatment groups underwent NIPPV treatment at least 4 h per day for > 3 months.

Predefined outcomes: All-cause mortality, exacerbation frequency, gas exchange, FEV1%predicted, 6MWD, dyspnea, and HRQL.

Two investigators (ZPW and ZJL) independently identified the title/abstract of each record that was eligible for further review, before retrieving the full text of the eligible studies. Subsequently, two reviewers independently read the full texts and decided which records to include. The two reviewers resolved any disagreements through consensus with a third investigator (YMM).
Data Extraction, Quality Assessment, and Publication Bias

We designed a data form recommended by Cochrane to extract and record the relevant information of each study. The form included the author, study time, study type, study location, interventions, sample size, and outcomes. If there were any data or information unavailable, we contacted the corresponding authors by e-mail to request the data.

The bias risks of each study were assessed according to the Cochrane risk of bias tool implemented in RevMan 5.3 software. Studies were assessed for seven domains as follows: selection bias, random sequence generation; selection bias, allocation concealment; performance bias, blinding of participants and personnel; detection bias, blinding of related outcomes assessment; attrition bias, incomplete outcome data; reporting bias, selective reporting; and other biases. Each study was graded as low, high, or unclear risk of bias.

Two reviewers (ZPW and ZJL) independently performed the data extraction and quality assessment; any disagreements were solved through discussion with a third investigator (YMM).

We used RevMan 5.3 software to generate Funnel plots to assess the publication bias. Funnel plot asymmetry was evaluated by the Egger regression test using Stata 15.1 (StataCorp LP, TX, USA).

Quality of the Evidence

We used GRADE to judge the quality of evidence for outcomes. Five factors (risk of bias, indirectness, inconsistency, imprecision, and publication bias) were used to upgrade or downgrade the quality of evidence. The quality of evidence was defined as high, moderate, low, or very low using GRADE profiler 3.6.1.

Statistical Analysis

The Cochrane systematic review software Review Manager software (RevMan; Version 5.3.5) was used for all statistical analyses. The risk ratio (RR) and 95% confidence intervals (CIs) were calculated for dichotomous data, mortality, and exacerbation rates. The weighted mean difference (WMD) with 95% CIs was adopted for continuous variables, PaCO2, PaO2, FEV1%predicted, exercise capacity, dyspnea, and HRQL. RRs with 95% CIs were calculated using the Mantel-Haenszel method using a random-effects model. WMD with 95% CIs was estimated by inverse variance method using a random effects model. We used a weighted Mantel-Haenszel $X^2$ and the $I^2$ to test statistical heterogeneities; when $P < 0.1$ and $I^2 > 50\%$, we regarded the heterogeneity of the study to be significant. If sufficient data were available, we conducted subgroup and sensitivity analyses to establish the source of heterogeneity. The interaction between subgroups was evaluated by the subgroup difference test described by Deeks et al. When $P_{interaction} < 0.05$, we considered that the difference was significant between subgroups and that there was a significant interaction between subgroups.

Subgroup Analysis

We performed subgroup analysis based on baseline PaCO2 level, degree of dPaCO2, whether PaCO2 was decreased or increased in the control treatment group, prior screening time, and duration of follow-up. The higher baseline PaCO2 subgroup included trials with patients with a mean PaCO2 ≥ 55 mmHg at the beginning, while the baseline PaCO2 lower subgroup included trials with patients with a mean 45 mmHg ≤ PaCO2 < 55 mmHg at the beginning.

The definition of dPaCO2 was as follows:

$dPaCO2 = (\text{baseline PaCO2 level of NIPPV group} – \text{endpoint PaCO2 level of NIPPV group}) – (\text{baseline PaCO2 level of control treatment group} – \text{endpoint PaCO2 level of control treatment group}).$

As shown in the definition, dPaCO2 could be a suitable index for two aspects: 1. the degree of change in PaCO2 in the control treatment group, and 2. the treatment ability of NIPPV to reduce PaCO2 compared to control treatment. For outcomes, we distributed the studies into two or three subgroups in high-to-low order.

Results

A total of 1556 citations were identified by the electronic literature search: PubMed = 279, Embase = 585, Cochrane = 323, CNKI = 86, Wan Fang = 170, and Wei Pu = 113. Three articles were identified from reviewing the reference list. A total of 1559 records were included for initial screening; among them, 478 duplicates were removed, 758 were
excluded based on screening of titles, and 249 were excluded after abstract screening. Finally, 55 studies were discarded through full-text review for the following reasons: not an RCT: 32, insufficient duration of NIPPV: 13, critical digital data unavailable: 5, not stable hypercapnic: 3, treated in hospital: 1, and pilot study: 1. Finally, 19 articles were included for quality assessment and analysis. The study selection procedure details were shown in Figure 1.

Study Description
All included articles were RCTs, 13 studies provided data on mortality, 17 studies showed data on PaCO2; 8 studies discussed the MRC-dyspnea scale, and 10 studies discussed SGRQ scores; further details are provided in Table 1. The characteristics of the included studies are shown in Table 1, and the baseline characteristics of patients in the included studies are shown in Table 2.

Study Quality and Publication Bias
All studies were evaluated as having an unclear risk of bias. All studies with the exception of one were judged as low risk of bias for random sequence generation. Four studies used proper allocation and concealment methods, while the others

Figure 1 Flow diagram of the study selection.
Abbreviation: NIPPV, noninvasive positive pressure ventilation.
<table>
<thead>
<tr>
<th>Author (Year)</th>
<th>Study Design</th>
<th>Location</th>
<th>Duration of Screening Time</th>
<th>Study Design</th>
<th>Location</th>
<th>Duration of Screening Time</th>
<th>Main Characteristics of the Subjects</th>
<th>Population (n); Randomized (Treatment/Control)</th>
<th>Intervention (Treatment/Control)</th>
<th>Duration</th>
<th>Time of Use NIPPV Per Day</th>
<th>IPAP</th>
<th>EPAP</th>
<th>dPCO2 (mmHg)</th>
<th>Main Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>Xiang et al. (2007)</td>
<td>RCT</td>
<td>China</td>
<td>4 W</td>
<td>FEV1/FVC &lt; 70%; FEV1 &lt; 50% predicted; PaCO2 ≥ 55 mmHg</td>
<td>20/20</td>
<td>NPPV + LTOT/LTOT</td>
<td>2 Y</td>
<td>&gt; 8 h</td>
<td>16–20</td>
<td>2–4</td>
<td>12.4</td>
<td>1, 2, 3, 4, 5, 7, 8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zhou et al. (2012)</td>
<td>RCT</td>
<td>China</td>
<td>NR</td>
<td>FEV1/FVC &lt; 70%; 30% &lt; FEV1 &lt; 50% predicted; stable hypercapnic</td>
<td>26/32</td>
<td>NPPV + LTOT/LTOT</td>
<td>3 M</td>
<td>6–9 h</td>
<td>8–12</td>
<td>2–4</td>
<td>NR</td>
<td>2, 3, 7, 9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zhou et al. (2008)</td>
<td>RCT</td>
<td>China</td>
<td>NR</td>
<td>FEV1/FVC &lt; 70%; Chronic respiratory failure</td>
<td>15/21</td>
<td>NPPV + LTOT/LTOT</td>
<td>1 Y</td>
<td>9 h</td>
<td>12–16</td>
<td>2–4</td>
<td>7.24</td>
<td>1, 2, 3, 4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tang et al. (2010)</td>
<td>RCT</td>
<td>China</td>
<td>Frequent exacerbation</td>
<td>FEV1/FVC &lt; 70%; PaCO2 ≥ 55 mmHg</td>
<td>12/13</td>
<td>NPPV + LTOT/LTOT</td>
<td>6 M</td>
<td>4–8 h</td>
<td>14–20</td>
<td>2–4</td>
<td>3.8</td>
<td>1, 2, 3, 6, 7, 9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shang et al. (2009)</td>
<td>RCT</td>
<td>China</td>
<td>4 W</td>
<td>FEV1/FVC &lt; 70%; FEV1 &lt; 50% predicted; PaCO2 ≥ 55 mmHg</td>
<td>20/20</td>
<td>NPPV + LTOT/LTOT</td>
<td>2 Y</td>
<td>&gt; 8 h</td>
<td>12–18</td>
<td>4–8</td>
<td>6</td>
<td>1, 2, 3, 4, 5, 7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Li et al. (2015)</td>
<td>RCT</td>
<td>China</td>
<td>NR</td>
<td>FEV1/FVC &lt; 70%; FEV1 &lt; 50% predicted</td>
<td>30/30</td>
<td>NPPV + LTOT/LTOT</td>
<td>6 M</td>
<td>&gt; 5 h</td>
<td>10–20</td>
<td>4–10</td>
<td>18.9</td>
<td>2, 3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lin et al. (2015)</td>
<td>RCT</td>
<td>China</td>
<td>NR</td>
<td>FEV1/FVC &lt; 70%; Relative severe; stable hypercapnic</td>
<td>39/39</td>
<td>NPPV + LTOT/LTOT</td>
<td>1 Y</td>
<td>4–9 h</td>
<td>12–18</td>
<td>2–4</td>
<td>NR</td>
<td>5, 6, 9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Li et al. (2011)</td>
<td>RCT</td>
<td>China</td>
<td>NR</td>
<td>FEV1/FVC &lt; 70%; FEV1 &lt; 50% predicted; type II respiratory failure</td>
<td>24/24</td>
<td>NPPV + LTOT/LTOT</td>
<td>1 Y</td>
<td>&gt; 8 h</td>
<td>16–20</td>
<td>4–6</td>
<td>11.5</td>
<td>1, 2, 3, 4, 5, 7, 8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Guo et al. (2015)</td>
<td>RCT</td>
<td>China</td>
<td>2 W</td>
<td>FEV1/FVC &lt; 70%; 30% ≤ FEV1 &lt; 50% predicted; PaCO2 ≥ 45 mmHg</td>
<td>30/30</td>
<td>NPPV + LTOT/LTOT</td>
<td>1 Y</td>
<td>6–9 h</td>
<td>10–14</td>
<td>4–5</td>
<td>15</td>
<td>2, 3, 4, 6, 7, 8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wang et al. (2010)</td>
<td>RCT</td>
<td>China</td>
<td>1 M</td>
<td>FEV1/FVC &lt; 70%; 30% ≤ FEV1 &lt; 50% predicted; PaCO2 ≥ 45 mmHg</td>
<td>18/18</td>
<td>NPPV + LTOT/LTOT</td>
<td>1 Y</td>
<td>4–14 h</td>
<td>12–20</td>
<td>4–8</td>
<td>17</td>
<td>2, 3, 4, 6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zhou et al. (2017)</td>
<td>RCT</td>
<td>China</td>
<td>4 W</td>
<td>FEV1/FVC &lt; 70%; FEV1 &lt; 50% predicted; PaCO2 &gt; 50 mmHg</td>
<td>57/58</td>
<td>NPPV + LTOT/LTOT</td>
<td>3 M</td>
<td>5.6 ± 1.4 h</td>
<td>17.8</td>
<td>4.2</td>
<td>6.06</td>
<td>1, 2, 3, 8, 9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Köhnlein et al. (2014)</td>
<td>RCT</td>
<td>Germany, Austria</td>
<td>4 W</td>
<td>FEV1/FVC &lt; 70%; FEV1 &lt; 30% predicted; PaCO2 &gt; 51.9 mmHg</td>
<td>102/102</td>
<td>NPPV + LTOT/LTOT</td>
<td>1 Y</td>
<td>5.9 h</td>
<td>21.6</td>
<td>4.8</td>
<td>7.45</td>
<td>1, 2, 4, 9</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(Continued)
Table 1 (Continued).

<table>
<thead>
<tr>
<th>Author et al. (Year)</th>
<th>Study Design</th>
<th>Location</th>
<th>Duration of Screening Time</th>
<th>Main Characteristics of the Subjects</th>
<th>Population (n); Randomized (Treatment/Control)</th>
<th>Intervention (Treatment/Control)</th>
<th>Duration</th>
<th>Time of Use NIPPV Per Day</th>
<th>IPAP</th>
<th>EPAP</th>
<th>dPCO2 (mmHg)</th>
<th>Main Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>Struik et al. (2014)</td>
<td>RCT</td>
<td>The Netherlands</td>
<td>48 h</td>
<td>FEV1/FVC &lt; 70%; FEV1 &lt; 50% predicted; PaCO2 &gt; 45 mmHg</td>
<td>101/100</td>
<td>NPPV + LTOT/LTOT</td>
<td>1 Y</td>
<td>&gt; 5 h</td>
<td>19.2</td>
<td>4.8</td>
<td>3</td>
<td>1, 2, 3, 5, 8, 9</td>
</tr>
<tr>
<td>McEvoy et al. (2009)</td>
<td>RCT</td>
<td>Australia</td>
<td>6 M</td>
<td>FEV1/FVC &lt; 60%; FEV1 &lt; 1.5 L or FEV1 &lt; 50% predicted; PaCO2 &gt; 46 mmHg</td>
<td>72/72</td>
<td>NPPV + LTOT/LTOT</td>
<td>2.21 Y</td>
<td>4.5 h</td>
<td>12.9</td>
<td>5.1</td>
<td>−1.1</td>
<td>1, 2, 3, 6</td>
</tr>
<tr>
<td>Clinin et al. (2002)</td>
<td>RCT</td>
<td>Italy, France</td>
<td>4 W</td>
<td>FEV1/FVC &lt; 60%; FEV1 &lt; 1.5 L or FEV1 &lt; 50% predicted; PaCO2 &gt; 50 mmHg, PaO2 &lt; 60 mmHg</td>
<td>43/47</td>
<td>NPPV + LTOT/LTOT</td>
<td>2 Y</td>
<td>9 ± 2 h</td>
<td>14</td>
<td>2</td>
<td>3.5</td>
<td>1, 2, 3, 4, 6, 7, 8, 9</td>
</tr>
<tr>
<td>Casanova et al. (2000)</td>
<td>RCT</td>
<td>Spain</td>
<td>3 M</td>
<td>FEV1/FVC &lt; 70%; FEV1 &lt; 45% predicted</td>
<td>26/26</td>
<td>NPPV + LTOT/LTOT</td>
<td>1 Y</td>
<td>&gt; 6 h</td>
<td>12</td>
<td>4</td>
<td>−1.3</td>
<td>1, 2, 3, 5, 6, 8</td>
</tr>
<tr>
<td>Gay et al. (1996)</td>
<td>RCT</td>
<td>America</td>
<td>Minimal 6 W</td>
<td>FEV1 &lt; 40% predicted; PaCO2 &gt; 45 mmHg</td>
<td>7/6</td>
<td>NPPV + LTOT/LTOT</td>
<td>3 M</td>
<td>5.9 h</td>
<td>10</td>
<td>2</td>
<td>NR</td>
<td>2, 3, 5, 6, 7</td>
</tr>
<tr>
<td>Duiverman et al. (2011)</td>
<td>RCT</td>
<td>The Netherlands</td>
<td>4 W</td>
<td>FEV1/FVC &lt; 70%, PaCO2 &lt; 50% predicted, PaCO2 &gt; 45 mmHg</td>
<td>31/35</td>
<td>NIPPV + PR/LTOT</td>
<td>2 Y</td>
<td>6.9 h</td>
<td>23</td>
<td>6</td>
<td>NR</td>
<td>1, 8, 9</td>
</tr>
<tr>
<td>Murphy et al. (2017)</td>
<td>RCT</td>
<td>UK</td>
<td>2–4 W</td>
<td>FEV1 &lt; 50% predicted, FEV1/FVC &lt; 60%, PaCO2 &gt; 53 mmHg</td>
<td>57/59</td>
<td>NIPPV + LTOT/sham–NIPPV +LTOT</td>
<td>1 Y</td>
<td>7.6 h</td>
<td>24</td>
<td>4</td>
<td>2</td>
<td>1, 2, 3, 9</td>
</tr>
</tbody>
</table>

Notes: Main outcomes: 1, all-cause mortality; 2, PaCO2; 3, PaO2; 4, frequency of hospitalization; 5, FEV1 (L); 6, FEV1%pred; 7, 6-MWD; 8, MRC-dyspnea scale; 9, SGRQ.
Abbreviations: RCT, randomized controlled trial; NIPPV, noninvasive positive pressure ventilation; LTOT, long-term oxygen therapy; NR, not reported; PR, pulmonary rehabilitation; IPAP, intermittent positive airway pressure; EPAP, expiratory positive airway pressure; PaCO2 = (NIPPV group baseline PaCO2 − NIPPV group endpoint PaCO2) − (control group baseline PaCO2 − control group endpoint PaCO2); PaCO2, partial pressure of arterial carbon dioxide; PaO2, partial pressure of arterial oxygen; FEV1 (L), forced expiratory volume in 1 s; FEV1% pred, forced expiratory volume in one second in predicted; 6-MWD, 6-min walk distance; MRC-dyspnea scale, Medical Research Council scale; SGRQ, St. George's Respiratory Questionnaire; W, week; M, month; Y, year; h, hour.
did not properly state their methods.12,13,15,41 Due to the lack of a suitable sham option for NIPPV, only one study that used sham NIPPV was evaluated as low risk of performance bias.40 Six studies used appropriate methods for outcome assessment, while the others did not.13–16,18,41 One study did not provide baseline data and was judged as having another bias. The risk of bias graph and the risk of bias summary are shown in Figures 2 and 3, respectively.

Publication bias for mortality was evaluated, and no significant publication bias was found by inspection of funnel plots (additional file: Figure s1). Asymmetry was tested by the Egger’s regression test, which showed no significant asymmetry (RR, −0.63; 95% CI: −1.56–0.31; P = 0.166) (additional file: Figure s2). Publication bias for other endpoints was also evaluated by funnel plots (additional file: Figures s3–s9).
<table>
<thead>
<tr>
<th>Study</th>
<th>Random sequence generation (selection bias)</th>
<th>Allocation concealment (selection bias)</th>
<th>Blinding of participants and personnel (performance bias)</th>
<th>Blinding of outcome assessment (detection bias)</th>
<th>Incomplete outcome data (attrition bias)</th>
<th>Selective reporting (reporting bias)</th>
<th>Other bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>Casanova et al (2000)(^{14})</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
</tr>
<tr>
<td>Clini et al (2002)(^{15})</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
</tr>
<tr>
<td>Duiverman et al (2011)(^{17})</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
</tr>
<tr>
<td>Gay et al (1996)(^{40})</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
</tr>
<tr>
<td>Guo et al (2015)(^{30})</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
</tr>
<tr>
<td>Köhnelein et al (2014)(^{13})</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
</tr>
<tr>
<td>Li et al (2011)(^{32})</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
</tr>
<tr>
<td>Li et al (2015)(^{33})</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
</tr>
<tr>
<td>Lin et al (2015)(^{31})</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
</tr>
<tr>
<td>McEvoy et al (2009)(^{12})</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
</tr>
<tr>
<td>Murphy et al (2017)(^{18})</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
</tr>
<tr>
<td>Shang et al (2009)(^{34})</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
</tr>
<tr>
<td>Struijk et al (2014)(^{16})</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
</tr>
<tr>
<td>Tang et al (2010)(^{35})</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
</tr>
<tr>
<td>Wang et al (2010)(^{36})</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
</tr>
<tr>
<td>Xiang et al (2007)(^{37})</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
</tr>
<tr>
<td>Zhou et al (2008)(^{39})</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
</tr>
<tr>
<td>Zhou et al (2012)(^{38})</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
</tr>
<tr>
<td>Zhou et al (2017)(^{41})</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
</tr>
</tbody>
</table>

**Figure 3** Risk of bias summary.
Quality of Evidence
The quality of evidence synthesis was evaluated using the GRADE approach according to GRADE Quality assessment, and the evidence syntheses are shown in additional file Table s1. 29

Primary Outcomes
Thirteen studies (n = 1168) that included mortality data were pooled and combined. Mortality was significantly lower in the NIPPV group compared to the control group (128/580 [22.1%] vs 172/588 [29.3%]; RR = 0.76; 95% CI: 0.61–0.95; P = 0.02; I² = 14%). There was no significant heterogeneity between the studies (chi² = 13.97, degrees of freedom [df] = 12, P = 0.30, I² = 14%) (Figure 4A).

Secondary Outcomes
Eight studies (n = 558) that included data on the frequency of Hospital admissions were included in the meta-analysis and indicated a significant reduction of hospitalization in the NIPPV group compared to the control group (WMD, −1.74; 95% CI: −2.30 to −1.17; P < 0.00001) and had a high heterogeneity (chi² = 85.85, df = 7; P < 0.00001; I² = 92%) (additional file: Figure s10).

Seventeen studies (n = 1042) that reported data on PaCO2 were combined and analyzed. PaCO2 decreased significantly in the NIPPV group compared to the control group (WMD, −7.07; 95% CI: −9.78 to −4.37; P < 0.00001) and had a high heterogeneity (chi² = 488.95; df = 16; P < 0.00001; I² = 97%) (additional file: Figure s11).

Sixteen studies (n = 831) included PaO2 data. These studies showed a significant improvement in PaO2 in the NIPPV group compared to the control group (WMD, 3.97; 95% CI: 1.86–6.08; P < 0.00001; I² = 85%), with a high heterogeneity (chi² = 100.03; df = 15; P < 0.00001; I² = 85%) (additional file: Figure s12).

Our study found that LTHNIPPV was also beneficial in improving exercise capacity and HRQL and in relieving dyspnea but showed no significant benefit for FEV1% pred (additional file 1: Figures s13–s16).

Association Between Baseline PaCO2 Level and the Treatment Effect of NIPPV versus Control Treatment
We performed subgroup analysis based on the baseline PaCO2 level and found that mortality was significantly lower in the NIPPV group compared to the control group in the higher baseline PaCO2-level subgroup but not in the lower baseline PaCO2-level subgroup (baseline PaCO2 ≥ 55 mmHg RR = 0.69; 95% CI: 0.51–0.94; P = 0.02; I² = 21% vs baseline PaCO2 < 55 mmHg RR = 0.87; 95% CI: 0.67–1.14; P = 0.32; I² = 0%); the interactions were not significant (P interaction = 0.26; Figure 4B).

The frequency of hospital admissions was significantly lower in the NIPPV group compared to the control group in both higher and lower baseline PaCO2-level subgroups (baseline PaCO2 ≥ 55 mmHg; WMD, −1.91; 95% CI: −2.47 to −1.35; P < 0.00001; I² = 91% vs baseline PaCO2 < 55 mmHg WMD −0.50; 95% CI: −1.26 to 0.26; P = 0.20); the interactions were significant (P interaction = 0.003; additional file: Figure s17).

End point PaCO2 was significantly lower in the NIPPV group compared to the control group in both higher and lower baseline PaCO2-level subgroups (baseline PaCO2 ≥ 55 mmHg; WMD, −8.76; 95% CI: −11.67 to −5.84; P < 0.00001; I² = 95% vs baseline PaCO2 < 55 mmHg; WMD −2.73; 95% CI: −4.52 to −0.94; P = 0.003; I² = 42%); the interactions were significant (P interaction = 0.0005; additional file: Figure s18).

Association Between Degree of dPaCO2 and Treatment Effect of NIPPV versus Control Treatment
We performed subgroup analysis based on the degree of dPaCO2 and found that mortality was significantly lower in the NIPPV group compared to the control group in the higher dPaCO2 subgroup but not in the lower dPaCO2 subgroup (higher dPaCO2 RR = 0.42; 95% CI: 0.28–0.63; P < 0.0001; I² = 0% vs lower dPaCO2 RR = 0.91; 95% CI: 0.75–1.12; P = 0.38; I² = 0%); the interactions were significant (P interaction = 0.0008; Figure 4C). We then distributed the studies into three subgroups based on the degree of dPaCO2 and found that mortality was significantly lower in the NIPPV group compared to the control group in the higher dPaCO2 subgroup but not in the moderate and lower dPaCO2 subgroups (higher dPaCO2 RR = 0.38; 95% CI: 0.24–0.60; P < 0.0001; I² = 0% vs moderate dPaCO2 RR = 0.94; 95% CI: 0.66–1.33; P = 0.72; I² = 0%;
Figure 4 Forest plot comparing mortality between NIPPV and control groups. (A) Overall result. (B) Subgroup analysis according to baseline PaCO2 level. (C) Subgroup analysis according to the degree of dPaCO2.

Abbreviations: CI, confidence interval; M-H, Mantel–Haenszel.
vs lower dPaCO2 RR = 0.87; 95% CI: 0.69–1.11; P = 0.27; I² = 0%); the interactions were significant (P interaction = 0.004; additional file: Figure s19).

The frequency of hospital admissions was significantly lower in the NIPPV group compared to the control group in both higher and lower dPaCO2 subgroups (higher dPaCO2 WMD −2.31; 95% CI: −2.87 to −1.74; P < 0.00001; I² = 89% vs lower dPaCO2 WMD −1.13; 95% CI: −2.30 to −1.17; P < 0.00001; I² = 19%); the interactions were significant (P interaction = 0.004; additional file: Figure s20). We then distributed the studies into three subgroups based on the degree of dPaCO2 and found that the frequency of hospital admissions was significantly lower in the NIPPV group compared to the control group in all three subgroups (higher dPaCO2 WMD −2.36; 95% CI: −3.06 to −1.67; P < 0.00001; I² = 92% vs moderate dPaCO2 WMD −1.57; 95% CI: −2.32 to −0.83; P < 0.0001; I² = 55% vs lower dPaCO2 WMD −0.98; 95% CI: −1.75 to −0.21; P = 0.01; I² = 72%); the interactions were significant (P interaction = 0.03; additional file: Figure s21).

Others
As for mortality, we performed subgroup analyses based on change in PaCO2 in the control treatment group, period of prior stable time, and duration of follow-up time (additional file: Figures s22–s24). The subgroup analysis results are shown in additional file: Table s2.

Discussion
The main finding of our study was that LTHNIPPV significantly reduced mortality, the frequency of hospital admissions, PaCO2, and dyspnea and improved PaO2, exercise capacity, and HRQL but showed no significant benefit for FEV1% pred compared to control treatment. Moreover, patients with higher baseline PaCO2 level benefited from decreased mortality and frequency of hospital admissions when treated with LTHNIPPV, while patients with lower baseline PaCO2 level did not. Reduction in PaCO2 may be associated with decreased mortality and reduced frequency of hospital admissions.

Our results had some differences compared to those of previous meta-analyses of LTHNIPPV in patients with stable hypercapnic COPD. First, we included 1482 patients in 19 RCTs, while other studies included between six and nine RCTs.19,21,22 Second, we found that LTHNIPPV could reduce mortality, while previous meta-analysis found that NIPPV reduced PaCO2 but did not significantly reduce mortality in patients with stable hypercapnic COPD. The most recent meta-analysis study, which combined RCTs and observational studies, found that LTHNIPPV reduced all-cause mortality compared to control treatment for patients with stable hypercapnic COPD, but when evidence was limited to data from RCTs, the results were not statistically significant.23 We hypothesize that patient selection and intensity of ventilation setting were the two main contributors to the different results of our study and previous studies.

With regard to patient selection, the stable hypercapnia criteria varied among the trials included in previous meta-analysis studies, and some of the included patients may have had reversible hypercapnia.20,22 Reversible hypercapnic patients are those who have hypercapnia during acute exacerbation but return to normocapnia during follow-up.7 Studies had found that patients with stable hypercapnia have a poor prognosis compared to those with reversible hypercapnia and normocapnia, but patients with reversible hypercapnia had a similar outcome as those with normocapnia.6–11 It is possible that LTHNIPPV reduces mortality in patients with stable hypercapnic COPD but not in patients with reversible hypercapnic COPD. Trials with high numbers of patients with reversible hypercapnia may have diluted the treatment effect of LTHNIPPV on reducing mortality compared to control treatment. Murphy et al found that during a screening period of 2 weeks, as many as 50% of the patients had already become normocapnic. Struijk et al found that after a 1-year follow-up, PaCO2 decreased by 11.25 mmHg in the NIPPV group and by 8.25 mmHg in the control treatment group. The criteria of stable hypercapnia should be stricter and unified in future studies.

With regard to the intensity of ventilation setting, it is likely that many RCTs included in previous meta-analyses had insufficient ventilation settings to reduce PaCO2, leading to negative results. Two meta-analysis studies found that LTHNIPPV treatment, with the goal of reducing PaCO2, significantly decreased mortality in patients with stable hypercapnia when compared to control treatment.22,41 The definition of high-intensity NIPPV is NIPPV treatment with the goal of normalizing or maximally reducing PaCO2.42 Studies and reviews have proven that higher IPAP or high-intensity ventilation improves blood gas exchange, FEV1% predicted, and/or mortality.43–46
Baseline PaCO2 level may be a good index for reversible hypercapnia and is associated with the degree of PaCO2 reduction. After a period of screening, patients with a higher baseline PaCO2 level (PaCO2 ≥ 55 mmHg) seem less likely to have reversible hypercapnia than patients with a relatively lower baseline PaCO2 (45 mmHg ≤ PaCO2 < 55 mmHg). Our meta-analysis study found that patients with higher baseline PaCO2 levels may have benefited from decreased mortality when treated with LTHNIPPV, while patients with lower baseline PaCO2 levels did not. When treated with LTHNIPPV, patients with higher baseline PaCO2 levels achieved a higher degree of PaCO2 reduction than patients with lower baseline PaCO2 levels (P_{interaction} = 0.0005; additional file: Figure s18).

Compared to control treatment, the dPaCO2 could be a suitable index for the treatment potential of NIPPV to reduce PaCO2. In trials with fewer patients with reversible hypercapnia, PaCO2 would not significantly decrease in the control treatment group, but if the intensity of NIPPV was sufficient, PaCO2 would significantly reduce in the NIPPV group, together with a higher degree of dPaCO2. Our study found that dPaCO2 was associated with decreased mortality and reduced the frequency of hospital admissions (P_{interaction} = 0.0008 and 0.0004 respectively; Figure 4C and additional file: Figure s19). Taken together, our results suggest that baseline level and reduction in PaCO2 are associated with treatment effect of LTHNIPPV in patients with stable hypercapnic COPD in terms of decreasing mortality.

Several mechanisms might account for the association between decreased mortality and baseline level and reduction in PaCO2. First, previous studies have shown that the strength and endurance of the diaphragm decline as a result of hypercapnia and that a reduction in PaCO2 might relieve respiratory muscle dysfunction. Second, CO2-associated vasodilatation in patients with hypercapnic COPD results in edema, which may be relieved by reducing the elevated PaCO2. Finally, a reduction in PaCO2 would be helpful for the central breathing center in terms of re-setting CO2 sensitivity and improving blood gases during spontaneous breathing periods.

**Limitations**
This meta-analysis study has some limitations that warrant discussion. First, the prior screening time for the definition of stable hypercapnic varied, which may have affected our results. To address this problem, subgroup analysis was performed, and the results remained consistent. Moreover, there was no significant heterogeneity for the main outcome, mortality; so, our main findings were credible. Second, the measurement tools and data types were different in some outcomes; so, not all results were combined and analyzed for some secondary outcomes, and the quality of these results was evaluated as low or very low. Third, the quality of the included studies was inconsistent; many trials had uncertain bias for selection, performance, and detection, and these risks of bias were considered when assessing the quality of evidence. Finally, we concluded from our results based on clinical trials that the mechanisms of the association between the reduction in PaCO2 and decreased mortality were still unclear; therefore, further studies are necessary to confirm our results.

**Conclusion**
LTHNIPPV significantly reduced mortality. The baseline level and reduction in PaCO2 may be associated with the treatment effect of LTHNIPPV in patients with stable hypercapnic COPD. Further large-scale, multicenter RCTs are needed to confirm our results.

**Abbreviations**
LTHNIPPV, long-term home noninvasive positive pressure ventilation; COPD, chronic obstructive pulmonary disease; PaCO2, partial pressure of arterial carbon dioxide; RCT, randomized clinical trial; GRADE, Grading of Recommendations Assessment, Development, and Evaluation; RR, relative risk; CI, confidence interval; WMD, weighted mean difference; PaO2, partial pressure of oxygen; HRQL, health-related quality of life; FEV1% pred, forced expiratory volume in one second in predicted; 6-MWD, 6-min walk distance; dPaCO2 = (NIPPV group baseline PaCO2 - NIPPV group endpoint PaCO2) - (control group baseline PaCO2 - control group endpoint PaCO2); GOLD, global initiative for chronic obstructive lung disease; Df, degrees of freedom; LTOT, long-term oxygen therapy; NR, not reported; PR, pulmonary rehabilitation; IPAP, intermittent positive airway pressure; EPAP, expiratory positive airway pressure; FEV1 (L), forced expiratory volume in first second; MRC-dyspnea scale, Medical Research Council scale;
SGRQ, St. George’s Respiratory Questionnaire; W, weeks; M, month, Y, year; h, hour; SD, standard deviation; WMD, weighted mean difference; SE, standard error; ARF, acute respiratory failure.

Acknowledgments
The study was supported by the National Key Research and Development Program of China (2019YFC0121700), Beijing Hospitals Authority Clinical Medicine Development of Special Funding Support (XMLX201709), Beijing Municipal Commission of Science and Technology Cultivation Research Program of Biomedicine and Life Science Innovation (Z171100000417057), and Beijing Hospitals Authority Youth Program (QML20180303).

Author Contributions
All authors have contributed toward this paper. YMM, ZXC, ZPW, and ZJL proposed the concepts and designed the study. ZPW and ZJL performed the acquisition, analysis, and interpretation of data and statistical analysis. All authors contributed to data analysis, drafting or revising the article, have agreed on the journal to which the article was submitted, gave final approval of the version to be published, and agree to be accountable for all aspects of the work.

Disclosure
The authors declare no conflicts of interest in this work.

References
10. Cooper CB, Waterhouse J, Howard P. Twelve year clinical study of patients with hypoxic cor pulmonale given long term domiciliary oxygen therapy. Thorax. 1987;42(2):105–110. doi:10.1136/thx.42.2.105


