Long-term outcomes in chronic obstructive pulmonary disease patients: exploring the effects of inhalatory devices and their influence on the outcome

Valeria Fadda
Dario Maratea
Department of Pharmaceutical Sciences, University of Florence, Sesto Fiorentino, Fi, Italy

Purpose: Numerous systematic reviews have examined the outcomes in patients with chronic obstructive pulmonary disease managed with different therapeutic strategies. However, no such studies have specifically focused on the effect of inhalation devices.

Methods: A standard PubMed search was carried out in which we identified all randomized placebo-controlled trials conducted in patients with moderate-to-severe or severe chronic obstructive pulmonary disease. The clinical end points were exacerbations rate, incidence of pneumonia, and mortality. Meta-regression was employed to assess the effect of the device. For the incidence of exacerbations, an equivalence analysis was also carried out.

Results: A total of 37 studies were analyzed. Four different devices were used across these trials (Respimat®, HandiHaler®, Diskus, and Turbuhaler®). Our meta-regression analysis failed to show any significant difference between devices with regard to exacerbation rate. Equivalence was shown for some comparisons (HandiHaler® vs Respimat®), but not for others. In analyzing mortality, Respimat® was shown to worsen this end point in comparison with Turbuhaler® and HandiHaler®. Moreover, Turbuhaler® showed a protective effect over Diskus in the incidence of pneumonia.

Conclusion: The results of our analysis represent the first attempt to explore the effect of the type of device on long-term outcomes. One important limitation was that most drugs were associated with one particular device, and so the effects of drugs and devices could not be reliably differentiated from one another.

Keywords: COPD, inhalation device, moderate-severe, meta-analysis, equivalence

Introduction
Chronic obstructive pulmonary disease (COPD) is the fourth leading cause of death in most industrialized countries and affects about three million people worldwide. COPD is not one single disease, but represents different chronic lung diseases that cause lung airflow limitations. The most common symptoms are breathlessness, excessive sputum production, and a chronic cough.1

Several therapeutic strategies are available for treating COPD and the effectiveness of the different approaches has been investigated by a large number of trials and numerous systematic reviews.2–9 On the other hand, different types of inhalers have been developed for the delivery of these agents (eg, Diskus/Accuhaler, Turbuhaler®, HandiHaler®, and Respimat®), but their role has not yet been fully investigated.

Despite the wide literature on this topic, no analysis has been specifically focused on whether the different devices can influence the clinical outcomes observed in these
patients (eg, exacerbation rates or incidence of pneumonia). Some papers in this area have only investigated patients’ preferences or manageability, or the causal relationship between a single device (Respinimat®) and mortality. However, further studies are clearly needed.

The goal of the present analysis is to summarize the overall evidence available on the therapeutic strategies used in COPD and to investigate the role of the different devices on hard outcomes. In particular, the primary objective of the analysis was to test whether or not the inhalation devices currently in use to deliver topical agents differ from one another in terms of clinical outcomes. Three end points were evaluated (exacerbation rates, mortality, and incidence of pneumonia). The following topical agents were examined: 1) long-acting muscarinic agonists (LAMAs) (tiotropium); 2) long acting beta antagonist (LABA) inhaled corticosteroids (ICS). The secondary objective of our study was to carry out a formal statistical test of equivalence (proof of no difference) between the different devices with regard to the end point of exacerbations.

Methods
Criteria of study eligibility for inclusion in meta-analyses and meta-regressions
Eligible studies included all placebo-controlled randomized controlled trials (RCTs) indexed in PubMed that evaluated topical pharmacological treatments in patients with moderate-to-severe or severe COPD.

Literature search
Our initial search on PubMed was conducted by combining the first index term focused on the disease condition (COPD OR chronic obstructive pulmonary disease) with the second index term focused on the pharmacological treatments (LABA OR long-acting beta agonist) OR (LAMA OR long-acting antimuscarinic) OR (ICS OR inhaled corticosteroid) along with a PubMed filter restricting the extraction to “systematic reviews” and/or “randomized controlled trials”. These searches were repeated using variants of these terms and supplemented by the analysis of the Cochrane Library.

Study selection and data extraction
Studies potentially suitable for our analysis were identified from the abstract and, when necessary, from full texts. Information on the end points was extracted in duplicate by VF and DM; differences were resolved by consensus.

Quality assessment
Two reviewers (VF and DM) assessed the risk of bias in included studies by determining the Jadad score. This method addresses three domains (randomization, blinding of participants and personnel, and incomplete outcome data) and generates a score on a scale from 0 to 5.

Data analysis
Our analyses were separately carried out on each of the clinical end points under examination. Each of these analyses was divided into two steps: 1) traditional pair-wise meta-analysis, in which treated patients were compared with controls given placebo according to the design of the original RCTs, and of the clinical end point concerned, and 2) meta-regression, in which we tested whether the type of device, handled as a covariate, had any influence on the clinical end points. Three clinical end points were analyzed: exacerbation rate, mortality, and incidence of pneumonia; all of these end points were dichotomous.

Step 1
Our pair-wise meta-analyses were stratified according to the type of device and were therefore performed in the form of subgroup meta-analyses. The model was a random-effect model implemented according to the Mantel–Haenszel method. All comparisons were expressed as relative risk (RR) along with the respective 95% confidence intervals (CIs). Heterogeneity was quantified using the F statistic test. The statistical computer program used for this purpose was the Open Meta-Analyst (OMA) software (version 4.16.12; Tufts University, Medford, MA, USA).

Step 2
For each of the three clinical end points, a meta-regression analysis was carried out to investigate whether or not the specific devices had any influence on the end point under examination. Standard techniques of meta-regression were used. Also in this case, the OMA software was employed for all statistical calculations. The devices used for drug delivery in the various RCTs were handled as a covariate for the meta-regression analysis. Each of these meta-regressions was preceded by a standard pair-wise meta-analysis in which the pooled indexes of outcome, along with their 95% CIs, were separately estimated for the patient subgroups corresponding to the devices examined. All meta-regressions were performed by handling the type of device as a categorical covariate; in particular, in keeping with the design of our analysis, a...
single device was identified as a common comparator, in each meta-regression, for the remaining devices. Finally, the data on exacerbation rates were incorporated into an analysis of equivalence that was aimed at testing whether the proof of no difference could be demonstrated for the pair-wise comparisons between specific devices. These equivalence tests were based on a well-known approach that, in a Forest plot, combines the traditional horizontal bars (indicating the 95% CI for individual RRs) with an equivalence interval that is between the lower and the upper margins of equivalence (see Ahn et al16 for further details). The margins employed for these analyses were directly obtained from the statistical power sections reported in the original RCTs. In this framework, the superiority of two-sided margins employed in the randomized trials were assumed to represent, at the same time, the margins of therapeutic equivalence.

Results

Literature search and study selection

Our initial search on PubMed retrieved 868 studies (including RCTs and systematic reviews). Among these publications, we identified 37 placebo-controlled RCTs that met all the criteria required for inclusion in our analysis. In the selection process that led us to identify the above 37 RCTs, some trials (N=37) were scored 4.

Four different devices were used in these RCTs, namely, Diskus/Accuhaler (eight trials), Turbuhaler® (ten trials), Respimat® (three trials), and HandiHaler® (16 trials). The devices Diskus/Accuhaler and Turbuhaler® were used to deliver fluticasone/salmeterol (eight trials) and budesonide/formoterol (ten trials), respectively; tiotropium (19 trials) was delivered using either Respimat® (three trials) or HandiHaler® (17 trials).

Meta-regression analysis

There were three meta-regression analyses that were investigated, namely, exacerbations, mortality, and pneumonia. In each analysis, a single device was identified as a common comparator for the remaining devices; whenever possible, Respimat® was the reference device; however, in a single analysis in which Respimat® was absent, Turbuhaler® was selected.

Exacerbations

A total of 24 studies were analyzed.17–40 Three therapeutic strategies vs placebo were compared: 1) fluticasone plus salmeterol; 2) budesonide plus formoterol; and 3) tiotropium. More importantly, four devices were used in these 24 RCTs.

The results of the subgroup meta-analysis focused on the different devices (Figure 2) showed the superiority of the agents delivered through the Respimat® and Handi-Haler® devices. In contrast, the treatments delivered through the Diskus or the Turbuhaler® devices failed to show any significant difference in comparison with placebo. In the traditional meta-analysis, the pooled rate of exacerbations calculated across the whole series of 24 trials generated a RR of 0.86 (95% CI: 0.80–0.92). The heterogeneity among the studies was at I²=65%. The results of our meta-regression found no significant differences in exacerbations across different devices. However, these results might be biased by the high degree of heterogeneity found among the included studies.

Mortality

A total of 33 pairs of patient arms (included in 29 studies)17–29,31–34,36–48 were studied for each comparison; the number of paired groups exceeded the number of trials included in our analysis because, as shown in Figure 3, some studies contributed four or more patient arms to this
Abbreviations:
- CI, confidence interval; Ev/Trt, treatment; Ev/Ctrl, control; RCTs, randomized controlled trials; RR, relative risk.

Notes:
- The Forest plot shows the difference in the event rates stratified according to the four devices. Symbols: $\text{I}=$ point estimate of the RR, horizontal bars $=95\%$ CIs of the RR, big diamond (in yellow) and vertical dotted line (in red) = pooled estimate calculated across all RCTs, and small diamond (in yellow) = pooled estimate for the specific device. All data refer to the comparison of active agents vs placebo. $p$ is a measure of heterogeneity.

Abbreviations: CI, confidence interval; Ev/Trt, treatment; Ev/Ctrl, control; RCTs, randomized controlled trials; RR, relative risk.

Figure 2 Subgroup meta-analysis: incidence of exacerbations in patients treated with budesonide/formoterol delivered through Turbuhaler$^{\text{®}}$ device, tiotropium delivered through Respimat or HandiHaler$^{\text{®}}$ device, and fluticasone/formoterol delivered through Diskus device, compared to placebo.

Notes: The Forest plot shows the difference in the event rates stratified according to the four devices. Symbols: $\text{I}=$ point estimate of the RR, horizontal bars $=95\%$ CIs of the RR, big diamond (in yellow) and vertical dotted line (in red) = pooled estimate calculated across all RCTs, and small diamond (in yellow) = pooled estimate for the specific device. All data refer to the comparison of active agents vs placebo. $p$ is a measure of heterogeneity.

Abbreviations: CI, confidence interval; Ev/Trt, treatment; Ev/Ctrl, control; RCTs, randomized controlled trials; RR, relative risk.

Figure 3 Subgroup meta-analysis: mortality in patients treated with budesonide/formoterol delivered through Turbuhaler$^{\text{®}}$ device, tiotropium delivered through Respimat$^{\text{®}}$ or HandiHaler$^{\text{®}}$ device, and fluticasone/formoterol delivered through Diskus device, compared to placebo.

Notes: The Forest plot shows the difference in the event rates stratified according to the four devices. Symbols: $\text{I}=$ point estimate of the RR, horizontal bars $=95\%$ CIs of the RR, big diamond (in yellow) and vertical dotted line (in red) = pooled estimate calculated across all RCTs, and small diamond (in yellow) = pooled estimate for the specific device. All data refer to the comparison of active agents vs placebo. $p$ is a measure of heterogeneity.

Abbreviations: CI, confidence interval; Ev/Trt, treatment; Ev/Ctrl, control; RCTs, randomized controlled trials; RR, relative risk.
meta-analysis. Active treatments were the same as in the previous analysis. According to the pair-wise meta-analysis, no device showed a significant protective effect on mortality in comparison with placebo. The Respimat® device showed a significant detrimental effect for this end point. For the mortality outcome, heterogeneity was absent (I²=0%). The results of our meta-regression showed a significant higher mortality rate of Respimat® over Turbuhaler® (RR = 1.68 [95% CI: 1.13–2.48]; P=0.009) and HandiHaler® (RR = 1.57 [95% CI: 1.08–2.28]; P=0.018), but not versus Diskus. No significant difference was found between Turbuhaler®, HandiHaler®, and Diskus.

Pneumonia

The information needed for our analysis was available from seven studies in which two devices were investigated. In the pair-wise meta-analysis, Turbuhaler® did not affect pneumonia (Figure 4). This end point was instead significantly worsened in the patient groups treated with Diskus. In the latter case, heterogeneity was at I²=43%.

Finally, our meta-regression found a significant difference favoring Turbuhaler® as compared with Diskus (RR = 0.55; 95% CI: 0.31–0.99; P=0.049).

Equivalence analysis

The secondary objective of our analysis was to test equivalence between devices according to the end point of exacerbations. Since all comparisons were based on RR, the lower and upper margins for equivalence were set, respectively, at RR = 0.76 and RR = 1.32 in keeping with the assumptions previously employed in the trial by Calverley et al (in which budesonide/formoterol, either alone or in combination, was compared to placebo). According to the design of our meta-regression, all values of RR that were firstly introduced in our equivalence testing relied on Respimat® as common comparator for all the devices. Hence, three values of RR (HandiHaler® vs Respimat®, Turbuhaler® vs Respimat®, and Diskus vs Respimat®) were directly derived from the results of meta-regression. Then, the values of RR for the three remaining comparisons (HandiHaler® vs Turbuhaler®, HandiHaler® vs Diskus, and Turbuhaler® vs Diskus) were determined by changing the reference device; finally, these values were incorporated into the equivalence analysis (Figure S1).

The results of our equivalence testing are presented in the Forest plot shown in Figure 5. According to the prespecified margins, equivalence was demonstrated for the comparison of HandiHaler® vs Respimat®, but not for the other five comparisons.

Discussion

To our knowledge, the previous literature on the effect of inhalation devices in COPD includes only a single study in which different inhalation devices were compared using FEV¹, as the outcome measure. Brocklebank et al evaluated terbutaline, salbutamol, and ipratropium bromide delivered through hand-held vs nebulizers; however, their results failed to show any significant difference among the devices.

The first RCT, specifically designed in terms of sample size to test the noninferiority between two devices delivering tiotropium, was carried out by Wise et al in 2013 (with 17,135 patients randomized to either Respimat® or HandiHaler®). The results of this trial, based on the risk

<table>
<thead>
<tr>
<th>Studies</th>
<th>Estimate (95% CI)</th>
<th>Ev/Trt</th>
<th>Ev/Ctrl</th>
</tr>
</thead>
<tbody>
<tr>
<td>Calverley et al³</td>
<td>4.031 (0.865, 18.800)</td>
<td>8254</td>
<td>2/256</td>
</tr>
<tr>
<td>Calverley et al³</td>
<td>0.596 (0.292, 3.399)</td>
<td>6257</td>
<td>5/256</td>
</tr>
<tr>
<td>Vaidya⁴</td>
<td>0.657 (0.370, 1.202)</td>
<td>16145</td>
<td>24/145</td>
</tr>
<tr>
<td>Subgroup Turbuhaler® (P=0.008)</td>
<td>1.313 (0.946, 1.822)</td>
<td>29656</td>
<td>31/657</td>
</tr>
<tr>
<td>Barnes et al⁵</td>
<td>3.265 (0.135, 7.878)</td>
<td>147</td>
<td>0/73</td>
</tr>
<tr>
<td>TORCH et al⁶</td>
<td>3.447 (1.263, 377)</td>
<td>2841.522 1901.564</td>
<td></td>
</tr>
<tr>
<td>TRISTAN et al⁷</td>
<td>2.896 (0.750, 10.610)</td>
<td>3/361</td>
<td></td>
</tr>
<tr>
<td>Matter et al⁸</td>
<td>5.492 (0.265, 113.349)</td>
<td>2/165</td>
<td>0/181</td>
</tr>
<tr>
<td>TRISTAN et al⁷</td>
<td>2.353 (0.613, 9.027)</td>
<td>7/158</td>
<td>3/361</td>
</tr>
<tr>
<td>Zheng et al⁹</td>
<td>3.600 (0.182, 87.311)</td>
<td>3297</td>
<td>0/148</td>
</tr>
<tr>
<td>Subgroup Diskus (P=0.91)</td>
<td>1.537 (1.301, 1.815)</td>
<td>3062.813 1906.2678</td>
<td></td>
</tr>
<tr>
<td>Overall (P=0.16)</td>
<td>1.463 (0.971, 2.269)</td>
<td>3353.469 2273.335</td>
<td></td>
</tr>
</tbody>
</table>

Figure 4 Subgroup meta-analysis: incidence of pneumonia in patients treated with budesonide/formoterol delivered through Turbuhaler® device, tiotropium delivered through Respimat® or HandiHaler® device, and fluticasone/formoterol delivered through Diskus device, compared to placebo.

Abbreviations: CI, confidence interval; Ev/Trt, treatment; Ev/Ctrl, control; RCTs, randomized controlled trials; RR, relative risk.
individual drugs were associated with one particular device; therefore, an unavoidable consequence was that the effects of drugs and devices could not be reliably differentiated from one another. In other words, the effects of the therapeutic interventions in our analysis were attributed to the devices, but these should actually be attributed to the device/drug combinations. While this may be merely a question of wording, in our view, the results of our analysis keep a large part of their interest because they provided a comprehensive and updated picture of the current therapeutic evidence.

As a clue in future perspective, to correctly quantify the device effect, specific studies would be needed, for example, in which tiotropium is delivered by HandiHaler®, Turbuhaler®, or Diskus, as well as studies where a LABA is also delivered by these three devices. At present, no such trials are available, and so the effects of devices cannot be separated from those of the inhaled drugs.

**Conclusion**

Conclusive results on this issue will require that RCTs are specifically designed to evaluate the causal effect of individual devices on hard outcomes. Meanwhile, our study can be considered a first step forward in this controversial area.

**Disclosure**

The authors declare that they have no conflicts of interest in this work.

**References**


20. Bogdan MA, Aizawa H, Fukuchi Y, Mishima M, Nishimura M, Ichinose M. Efficacy and safety of inhaled formoterol 4.5 and 9 


Supplementary material

A Summary
Meta-regression
Metric: relative risk
Model results

<table>
<thead>
<tr>
<th>Covariate</th>
<th>Level</th>
<th>Studies</th>
<th>Coefficients</th>
<th>Lower bound</th>
<th>Upper bound</th>
<th>SE</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td></td>
<td></td>
<td>-0.199</td>
<td>-0.262</td>
<td>-0.135</td>
<td>0.032</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Device</td>
<td>Respimat®</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>HandiHaler®</td>
<td>17</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Diskus</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Turbuhaler®</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

B Summary
Meta-regression
Metric: relative risk
Model results

<table>
<thead>
<tr>
<th>Covariate</th>
<th>Level</th>
<th>Studies</th>
<th>Coefficients</th>
<th>Lower bound</th>
<th>Upper bound</th>
<th>SE</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td></td>
<td></td>
<td>0.376</td>
<td>0.023</td>
<td>0.730</td>
<td>0.180</td>
<td>0.037</td>
</tr>
<tr>
<td>Device</td>
<td>Respimat®</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>HandiHaler®</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Diskus</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Turbuhaler®</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Respimat®</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

C Summary
Meta-regression
Metric: relative risk
Model results

<table>
<thead>
<tr>
<th>Covariate</th>
<th>Level</th>
<th>Studies</th>
<th>Coefficients</th>
<th>Lower bound</th>
<th>Upper bound</th>
<th>SE</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td></td>
<td></td>
<td>-0.176</td>
<td>-0.726</td>
<td>0.375</td>
<td>0.281</td>
<td>0.532</td>
</tr>
<tr>
<td>Device</td>
<td>Respimat®</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>HandiHaler®</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure S1 Results of the meta-regression analysis investigating the influence of the device on the incidence of three outcomes: exacerbation rate (A), mortality (B), and incidence of pneumonia (C).

Abbreviation: SE, standard error.