Association between lung function and exacerbation frequency in patients with COPD

Martine Hoogendoorn 1
Talitha L Feenstra 2,3
Rudolf T Hoogenveen 2
Maiwenn Al 1
Maureen Rutten-van Mölken 1

1 Institute for Medical Technology Assessment, Erasmus University, Rotterdam; 2 Department for Prevention and Health Services Research, National Institute for Public Health and the Environment, Bilthoven; 3 Department of Epidemiology, University Medical Centre Groningen, Groningen, The Netherlands

Purpose: To quantify the relationship between severity of chronic obstructive pulmonary disease (COPD) as expressed by Global Initiative for Chronic Obstructive Lung Disease (GOLD) stage and the annual exacerbation frequency in patients with COPD.

Methods: We performed a systematic literature review to identify randomized controlled trials and cohort studies reporting the exacerbation frequency in COPD patients receiving usual care or placebo. Annual frequencies were determined for total exacerbations defined by an increased use of health care (event-based), total exacerbations defined by an increase of symptoms, and severe exacerbations defined by a hospitalization. The association between the mean forced expiratory volume in one second (FEV1)% predicted of study populations and the exacerbation frequencies was estimated using weighted log linear regression with random effects. The regression equations were applied to the mean FEV1% predicted for each GOLD stage to estimate the frequency per stage.

Results: Thirty-seven relevant studies were found, with 43 reports of total exacerbation frequency (event-based, n = 19; symptom-based, n = 24) and 14 reports of frequency of severe exacerbations. Annual event-based exacerbation frequencies per GOLD stage were estimated at 0.82 (95% confidence interval 0.46–1.49) for mild, 1.17 (0.93–1.50) for moderate, 1.61 (1.51–1.74) for severe, and 2.10 (1.51–2.94) for very severe COPD. Annual symptom-based frequencies were 1.15 (95% confidence interval 0.67–2.07), 1.44 (1.14–1.87), 1.76 (1.70–1.88), and 2.09 (1.57–2.82), respectively. For severe exacerbations, annual frequencies were 0.11 (95% confidence interval 0.02–0.56), 0.16 (0.07–0.33), 0.22 (0.20–0.23), and 0.28 (0.14–0.63), respectively. Study duration or type of study (cohort versus trial) did not significantly affect the outcomes.

Conclusion: This study provides an estimate of the exacerbation frequency per GOLD stage, which can be used for health economic and modeling purposes.

Keywords: COPD, exacerbations, disease severity, GOLD, review, regression

Introduction

The progression of chronic obstructive pulmonary disease (COPD) is often accompanied by periods of increasing symptoms, such as dyspnea, cough, and sputum production, known as exacerbations. Exacerbations are important events because they are associated with an increase in mortality,1,2 significant impairment of health-related quality of life,3–5 and an increase in health care use and associated costs.6,7 especially in the event of a hospitalization.8 The exacerbation frequency is therefore an important outcome parameter in COPD.9,10 However, quantification of the average exacerbation frequency is difficult. Many studies report the exacerbation frequency, but results
cannot be compared directly because different definitions are used, exacerbations are measured in different seasons, or data come from different types of studies, eg, clinical trials or cohort studies, each using specific inclusion criteria. Use of different definitions in particular seems to have a large influence.

Definitions of exacerbation can be roughly divided into two groups, ie, symptom-based definitions and event-based definitions. Studies defining exacerbations as self-reported changes in symptoms (symptom-based definition) generally result in higher estimates than studies using event-based definitions, because they also include exacerbations which do not present to physicians. When symptoms are closely monitored using diaries, these “unreported” exacerbations are thought to account for about 50% of all exacerbations. Event-based definitions use more objective criteria, such as a doctor’s visit, use of antibiotics and/or systemic steroids, or hospitalization. However, event-based definitions are sensitive to differences in treatment patterns between settings.

Another source of variation between studies is the method used to classify the severity of an exacerbation. Most studies classify exacerbations based on the treatment required, ie, either an increase of short-acting bronchodilator or maintenance medication use, additional antibiotics and/or systemic corticosteroids, or hospitalization. However, event-based definitions are sensitive to differences in treatment patterns between settings.

Despite the difficulties in measuring exacerbations, the general pattern is that the frequency of exacerbations increases with decreasing lung function. However, as far as we know, no studies have quantified this relationship. The present study aimed to quantify the relationship between degree of airflow obstruction, expressed as the forced expiratory volume in one second (FEV1)% predicted, and the annual exacerbation frequency, using previously published data. The association was estimated separately for symptom-based and event-based exacerbations and for total and severe exacerbations. Furthermore, we explored the impact of study duration and type of study, ie, clinical trial or cohort study, on this relationship. This study arose out of the need to estimate the average exacerbation frequency for the different COPD severity stages as defined by the Global Initiative for Chronic Obstructive Lung Disease (GOLD) that were used as input parameters in a COPD disease progression model. Because this model aims to simulate the long-term cost-effectiveness of interventions which successfully prevent exacerbations compared with minimal care, the exacerbation frequency in patients receiving minimal care was essential.

Methods
A systematic literature review was performed to identify randomized controlled trials and cohort studies reporting the exacerbation frequency in patients receiving care as usual or placebo. MEDLINE, EMBASE, and the Cochrane database were searched using the key words “chronic obstructive pulmonary disease”, “COPD”, or “chronic bronchitis” in combination with “exacerbat*” and the specification “cohort or survey” or “observation*”, or the selection “clinical trial”. Studies were included if they were published after 1990, had a follow-up of at least three months, used an event- or symptom-based definition for an exacerbation, and included a group of patients that received either usual care or placebo (eg, the placebo arm of a long-acting bronchodilator trial or a combination treatment trial). Studies that included a subgroup of COPD patients selected based on criteria other than lung function were excluded (eg, studies including only patients admitted to hospital or patients with an acute exacerbation at baseline). Retrospective studies based on administrative or claims data were excluded because the algorithms to identify exacerbations in these databases are often quite different from the definitions used in prospective cohort studies or clinical trials. Finally, references of the studies that met the inclusion and exclusion criteria were checked.

Primary outcomes
The three main outcomes of the study were the annual frequency of total exacerbations using an event-based definition, the annual frequency of total exacerbations using a symptom-based definition, and the annual frequency of severe exacerbations as defined by a hospitalization. One study could provide more than one estimate of exacerbation frequency by presenting separate rates for total and severe exacerbations or rates based on both a symptom- and an event-based definition, or by presenting rates for different lung function classes.

Data extraction
Because the comparator arm in our model needed to reflect minimal care, we only extracted exacerbation data for the groups of patients that received either usual care or placebo. The following data were extracted: percentage males, mean age, mean lung function (FEV1% predicted of the study population), follow-up duration, definition of exacerbation used (symptom- or event-based), and the annual exacerbation frequency. If the mean FEV1 was only given in liters, the mean FEV1% predicted for the study population was calculated using the association between the absolute value and percentage...
predicted from other studies. If the exacerbation frequency was presented for different classes of the FEV$_1$% predicted and the mean within-class FEV$_1$% predicted was not specified, the mean FEV$_1$% predicted was estimated based on the mean and standard deviation of the FEV$_1$% predicted in the total population assuming a normal distribution, or it was assumed to be the middle FEV$_1$% predicted for that specific class.

Data on the exacerbation frequency were recalculated to annual exacerbation rates, if necessary. The annual exacerbation rate was calculated by dividing the total number of exacerbations by the total number of patient-years, on the assumption that dropouts count for half of the follow-up time.

Data analysis

Because almost all the studies provided only point estimates of exacerbation rates, uncertainty around the exacerbation rates was estimated assuming the exacerbations to follow a Poisson distribution within each study. To quantify the relationship between the FEV$_1$% predicted and the annual exacerbation frequency, weighted log linear regression analysis with random effects was performed. Log linear regression was chosen in order to symmetrize the skewed distribution of the exacerbation rates and approximate a normal distribution of residuals in the linear regression analysis. A random effect model was chosen to account for study heterogeneity. The logarithm of the annual exacerbation frequency was used as the dependent variable and the mean FEV$_1$% predicted of the study as the independent variable. The regression analysis was performed using the S-Plus routine general linear model for mixed-effects models. Analyses were performed separately for total event-based, total symptom-based, and severe exacerbations. From the resulting regression equation, the predicted log exacerbation rate for a specific FEV$_1$% predicted could be calculated. Simply taking the exponential function of the logarithm of the exacerbation rate, in order to retransform the data into a normal exacerbation rate introduces bias and inconsistency. Therefore, we have used the nonparametric smearing factor, which was calculated following the method of Duan et al. According to this method, the smearing factor, $\phi$, can be calculated as the weighted mean of the exponential of the differences between the logarithm of the observed and predicted exacerbation rates in the selected studies using the number of exacerbations in a study as a weight. This smearing factor is then multiplied by the uncorrected predicted exacerbation rates to find corrected predicted exacerbation rates for a given FEV$_1$% predicted. As a result, the relationship between the annual exacerbation frequency and the FEV$_1$% predicted is:

Annual exacerbation frequency $= \phi \cdot \exp[a + \beta \cdot \text{FEV1\% predicted}]$

whereby $\phi =$ smearing factor, $a =$ intercept (estimated in the regression analysis), $\beta =$ coefficient for FEV$_1$% predicted (estimated in the regression analysis).

This equation was used to calculate the annual exacerbation frequency in the four COPD severity stages according to the GOLD classification using a mean FEV$_1$% predicted of 90 for mild, 65 for moderate, 42 for severe, and 23 for very severe COPD. To include the uncertainty around the smearing factor jointly with the uncertainty around the regression coefficients, the uncertainty around the exacerbation rates per GOLD stage was estimated by Monte Carlo simulation, ie, 1000 random draws were taken from the joint distribution of the intercept and the coefficient for FEV$_1$% predicted. For each combination of intercept and coefficient, the accompanying smearing factor was calculated using the formula described above. The mean FEV$_1$% predicted per GOLD stage was then applied to each of the 1000 combinations of intercept, coefficient for FEV$_1$% predicted, and smearing factor, resulting in 1000 estimates of the exacerbation rate per GOLD stage. The 2.5% and 97.5% percentiles of these 1000 estimates formed the 95% uncertainty interval.

Additional regression analyses were performed adding follow-up duration (in months) and type of study (cohort versus trial) to FEV$_1$% predicted as dependent variables. The analyses were performed with Splus 8.1 (TIBCO Spotfire S+ Version 8.1.1 HF-001 for Microsoft Windows, 2008).

Results

The literature review identified 86 references for trials and cohort studies published after 1990 that seemed eligible based on their titles. Of these 86 references that were obtained in full, another 44 studies were excluded because they did not present exacerbation frequencies or numbers ($n = 13$), were based on a selective subgroup of COPD patients ($n = 11$), were based on a cross-sectional study or on administrative or claims data ($n = 8$), had a follow-up less than three months ($n = 9$), or used a deviant definition for an exacerbation ($n = 3$). The final 42 references referred to 37 unique studies, comprising 28 trials and nine cohort studies. This resulted in 43 estimates for the total exacerbation frequency and 14 estimates of the frequency of severe exacerbations. Of the 43 estimates of total exacerbation frequency, 19 used the event-based definition and 24 used the symptom-based definition. Characteristics of all the included studies with their annual exacerbation rates are presented in Table 1.
### Table 1 Characteristics of included studies

<table>
<thead>
<tr>
<th>Type of study</th>
<th>First author</th>
<th>n</th>
<th>% males</th>
<th>Mean age (years)</th>
<th>Mean FEV1% pred</th>
<th>Follow-up (months)</th>
<th>Definition used for an exacerbation</th>
<th>Annual total exacerbation rate</th>
<th>Annual severe exacerbation rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trial</td>
<td>Monninkhof et al&lt;sup&gt;11&lt;/sup&gt;</td>
<td>121</td>
<td>84</td>
<td>65</td>
<td>58</td>
<td>12</td>
<td>Event-based</td>
<td>1.51</td>
<td>0.14</td>
</tr>
<tr>
<td>Trial</td>
<td>Coultas et al&lt;sup&gt;22&lt;/sup&gt;</td>
<td>51</td>
<td>54</td>
<td>69</td>
<td>46</td>
<td>6</td>
<td>Event-based</td>
<td>0.20</td>
<td></td>
</tr>
<tr>
<td>Trial</td>
<td>Rea et al&lt;sup&gt;33&lt;/sup&gt;</td>
<td>52</td>
<td>41</td>
<td>68</td>
<td>50</td>
<td>12</td>
<td>Event-based</td>
<td>0.67</td>
<td></td>
</tr>
<tr>
<td>Trial</td>
<td>Littlejohns et al&lt;sup&gt;34&lt;/sup&gt;</td>
<td>65</td>
<td>63</td>
<td>63</td>
<td>50</td>
<td>12</td>
<td>Event-based</td>
<td>0.31</td>
<td></td>
</tr>
<tr>
<td>Trial</td>
<td>Gallefoss and Bakke&lt;sup&gt;25&lt;/sup&gt;</td>
<td>31</td>
<td>52</td>
<td>58</td>
<td>56</td>
<td>12</td>
<td>Event-based</td>
<td>0.14</td>
<td></td>
</tr>
<tr>
<td>Trial</td>
<td>Brussasco et al&lt;sup&gt;20&lt;/sup&gt;</td>
<td>400</td>
<td>76</td>
<td>65</td>
<td>39</td>
<td>6</td>
<td>Symptom-based</td>
<td>1.49</td>
<td>0.15</td>
</tr>
<tr>
<td>Trial</td>
<td>Casaburi et al&lt;sup&gt;37&lt;/sup&gt;</td>
<td>371</td>
<td>63</td>
<td>65</td>
<td>38</td>
<td>12</td>
<td>Symptom-based</td>
<td>0.95</td>
<td>0.16</td>
</tr>
<tr>
<td>Trial</td>
<td>Niewoehner et al&lt;sup&gt;38&lt;/sup&gt;</td>
<td>915</td>
<td>99</td>
<td>68</td>
<td>36</td>
<td>6</td>
<td>Symptom-based</td>
<td>1.05</td>
<td>0.25</td>
</tr>
<tr>
<td>Trial</td>
<td>Vincken et al&lt;sup&gt;39&lt;/sup&gt;</td>
<td>179</td>
<td>86</td>
<td>65</td>
<td>39</td>
<td>12</td>
<td>Symptom-based</td>
<td>0.96</td>
<td>0.16</td>
</tr>
<tr>
<td>Trial</td>
<td>Dusser et al&lt;sup&gt;40&lt;/sup&gt;</td>
<td>510</td>
<td>87</td>
<td>65</td>
<td>48</td>
<td>12</td>
<td>Event-based</td>
<td>0.15</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>280</td>
<td>–</td>
<td>–</td>
<td>67</td>
<td>12</td>
<td>Event-based</td>
<td>1.97</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>230</td>
<td>–</td>
<td>–</td>
<td>31</td>
<td>12</td>
<td>Event-based</td>
<td>2.70</td>
<td></td>
</tr>
<tr>
<td>Trial</td>
<td>Calverley et al&lt;sup&gt;31&lt;/sup&gt;</td>
<td>361</td>
<td>75</td>
<td>63</td>
<td>44</td>
<td>12</td>
<td>Event-based</td>
<td>1.30</td>
<td></td>
</tr>
<tr>
<td>Trial</td>
<td>Calverley et al&lt;sup&gt;32&lt;/sup&gt;</td>
<td>256</td>
<td>75</td>
<td>65</td>
<td>36</td>
<td>12</td>
<td>Event-based</td>
<td>1.80</td>
<td></td>
</tr>
<tr>
<td>Trial</td>
<td>Szafranski et al&lt;sup&gt;33&lt;/sup&gt;</td>
<td>205</td>
<td>83</td>
<td>65</td>
<td>36</td>
<td>12</td>
<td>Event-based</td>
<td>1.87</td>
<td></td>
</tr>
<tr>
<td>Trial</td>
<td>Calverley et al&lt;sup&gt;34&lt;/sup&gt;</td>
<td>1524</td>
<td>76</td>
<td>65</td>
<td>44</td>
<td>36</td>
<td>Event-based</td>
<td>1.13</td>
<td>0.19</td>
</tr>
<tr>
<td>Trial</td>
<td>Dal Negro et al&lt;sup&gt;35&lt;/sup&gt;</td>
<td>6</td>
<td>83</td>
<td>40-76</td>
<td>50</td>
<td>12</td>
<td>Event-based</td>
<td>4.17</td>
<td></td>
</tr>
<tr>
<td>Trial</td>
<td>Wonsurakiat et al&lt;sup&gt;36&lt;/sup&gt;</td>
<td>125</td>
<td>95</td>
<td>68</td>
<td>60</td>
<td>12</td>
<td>Symptom-based</td>
<td>1.35</td>
<td>0.06</td>
</tr>
<tr>
<td>Trial</td>
<td>Allegra et al&lt;sup&gt;37&lt;/sup&gt;</td>
<td>218</td>
<td>71</td>
<td>59</td>
<td>70</td>
<td>6</td>
<td>Symptom-based</td>
<td>1.32</td>
<td></td>
</tr>
<tr>
<td>Trial</td>
<td>Bontognali&lt;sup&gt;38&lt;/sup&gt;</td>
<td>30</td>
<td>57</td>
<td>59</td>
<td>75</td>
<td>3</td>
<td>Event-based</td>
<td>1.27</td>
<td></td>
</tr>
<tr>
<td>Trial</td>
<td>Decramer et al&lt;sup&gt;39&lt;/sup&gt;</td>
<td>258</td>
<td>79</td>
<td>62</td>
<td>57</td>
<td>36</td>
<td>Event-based</td>
<td>1.31</td>
<td></td>
</tr>
<tr>
<td>Trial</td>
<td>Grassi et al&lt;sup&gt;40&lt;/sup&gt;</td>
<td>41</td>
<td>79</td>
<td>62</td>
<td>57</td>
<td>3</td>
<td>Event-based</td>
<td>5.37</td>
<td></td>
</tr>
<tr>
<td>Trial</td>
<td>Hansen et al&lt;sup&gt;41&lt;/sup&gt;</td>
<td>70</td>
<td>46</td>
<td>52</td>
<td>85</td>
<td>5</td>
<td>Event-based</td>
<td>1.95</td>
<td></td>
</tr>
<tr>
<td>Trial</td>
<td>Malerba et al&lt;sup&gt;42&lt;/sup&gt;</td>
<td>119</td>
<td>76</td>
<td>61</td>
<td>70</td>
<td>12</td>
<td>Symptom-based</td>
<td>0.87</td>
<td></td>
</tr>
<tr>
<td>Trial</td>
<td>Meister et al&lt;sup&gt;43&lt;/sup&gt;</td>
<td>124</td>
<td>41</td>
<td>58</td>
<td>79</td>
<td>6</td>
<td>Event-based</td>
<td>1.20</td>
<td></td>
</tr>
<tr>
<td>Trial</td>
<td>Moretti et al&lt;sup&gt;44&lt;/sup&gt;</td>
<td>61</td>
<td>75</td>
<td>68</td>
<td>59</td>
<td>8</td>
<td>Event-based</td>
<td>2.07</td>
<td></td>
</tr>
<tr>
<td>Trial</td>
<td>Pela et al&lt;sup&gt;45&lt;/sup&gt;</td>
<td>84</td>
<td>71</td>
<td>66</td>
<td>59</td>
<td>6</td>
<td>Event-based</td>
<td>3.50</td>
<td></td>
</tr>
<tr>
<td>Trial</td>
<td>Burge et al&lt;sup&gt;46&lt;/sup&gt;</td>
<td>370</td>
<td>74</td>
<td>64</td>
<td>50</td>
<td>36</td>
<td>Event-based</td>
<td>1.90</td>
<td></td>
</tr>
<tr>
<td>Trial</td>
<td>van Grunsven et al&lt;sup&gt;47&lt;/sup&gt;</td>
<td>88</td>
<td>90</td>
<td>61</td>
<td>44</td>
<td>24</td>
<td>Event-based</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>Trial</td>
<td>Vestbo et al&lt;sup&gt;48&lt;/sup&gt;</td>
<td>145</td>
<td>62</td>
<td>59</td>
<td>87</td>
<td>36</td>
<td>Event-based</td>
<td>0.45</td>
<td></td>
</tr>
<tr>
<td>Cohort</td>
<td>Lior et al&lt;sup&gt;49&lt;/sup&gt;</td>
<td>136</td>
<td>96</td>
<td>70</td>
<td>49</td>
<td>24</td>
<td>Event-based</td>
<td>0.93</td>
<td></td>
</tr>
<tr>
<td>Cohort</td>
<td>Mittmann et al&lt;sup&gt;50&lt;/sup&gt;</td>
<td>609</td>
<td>58</td>
<td>69</td>
<td>44</td>
<td>12</td>
<td>Event-based</td>
<td>1.39</td>
<td>0.27</td>
</tr>
<tr>
<td>Cohort</td>
<td>Langsetmo et al&lt;sup&gt;51&lt;/sup&gt;</td>
<td>421</td>
<td>57</td>
<td>67</td>
<td>46</td>
<td>6</td>
<td>Event-based</td>
<td>2.70</td>
<td></td>
</tr>
<tr>
<td>Cohort</td>
<td>Hutchinson et al&lt;sup&gt;52&lt;/sup&gt;</td>
<td>92</td>
<td>63</td>
<td>72</td>
<td>40</td>
<td>Median 10.8</td>
<td>Event-based</td>
<td>1.79</td>
<td></td>
</tr>
<tr>
<td>Cohort</td>
<td>O’Reilly et al&lt;sup&gt;53&lt;/sup&gt;</td>
<td>127</td>
<td>62</td>
<td>69</td>
<td>50</td>
<td>12</td>
<td>Event-based</td>
<td>3.20</td>
<td></td>
</tr>
</tbody>
</table>
The left three graphs in Figure 1 show the logarithm of the annual total and severe exacerbation frequency plotted against the mean FEV₁% predicted for each study, as well as the estimated relationship between the two obtained from the regression analyses.

The estimated coefficients for the relationship between the mean FEV₁% predicted and the exacerbation frequency are shown in Table 2. Lung function was a predictor of borderline significance ($P = 0.053$) for event-based exacerbations only (symptom-based, $P = 0.19$; severe exacerbations, $P = 0.50$). The final association between the FEV₁% predicted and the exacerbation frequency after retransforming the predicted log exacerbation rate into a normal exacerbation rate are shown in the right three graphs in Figure 1. Results for the mean exacerbation frequencies for the different GOLD stages based on the regression equations are presented in Table 3. Using an event-based definition, the total exacerbation frequency was significantly higher in patients with an FEV₁% predicted below 50% compared with patients having an FEV₁% predicted above 50%.

Regression analyses with additional covariates showed no significant effect of duration of follow-up of the study or type of study (cohort versus trial). The duration of follow-up was of borderline significance only for total exacerbations using the symptom-based definition, with longer follow-up resulting in lower rates (Table 4).

**Discussion**

Although many trials and cohort studies report on the important outcome, ie, exacerbation frequency, the association between lung function and exacerbation frequency is less often investigated. The current study systematically reviewed the information contained in the literature and combined it into an estimate of exacerbation frequency as a function of FEV₁% predicted. The coefficient for lung function showed borderline significance for total exacerbations using the event-based definition ($P = 0.053$), and was insignificant for total exacerbations using a symptom-based definition and severe exacerbations. Based on the estimated equation, the final estimates of the total exacerbation frequency per GOLD severity stage using the event-based definition were 0.82 for mild, 1.17 for moderate, 1.61 for severe, and 2.10 for very severe COPD. In spite of the overlapping uncertainty intervals, these estimates are useful for health economic/modeling purposes, as long as they are accompanied by an appropriate uncertainty probabilistic sensitivity analysis. In this way, the 95% confidence intervals vary substantially per GOLD stage, which would be ignored using a single exacerbation frequency for all GOLD stages.
In accordance with the general finding that using the symptom-based definition results in higher estimates of the total exacerbation frequency, we found slightly higher estimates for mild, moderate, and severe COPD using the symptom-based definition compared with the event-based definition. However, this difference was not significant, and seemed to get smaller with increasing severity of COPD. We also did not see an effect of follow-up duration. The mean follow-up in the studies in this review was 14 months, ranging from three to 36 months.

The study had a couple of limitations and strengths. A reason why the relationship between lung function and

---

**Figure 1** Left graphs: Logarithm of the annual total or severe exacerbation frequency plotted against the mean forced expiratory volume in one second (FEV1)% predicted of the study, line = estimated relationship obtained from the log-linear regression. Right graphs: Annual total or severe exacerbation frequency plotted against the mean FEV1% predicted of the study, line = relationship based on the retransformed exacerbation rates using the smearing factor.
exacerbation frequency in our study was relatively weak may be our use of published data. Regression on study summary estimates, as done in this study, has substantially less power than regression on patient-level data. It is likely that variation in lung function across studies is lower than variation in lung function across patient-level data within studies. By plotting the mean exacerbation frequency against the mean FEV₁ % predicted of a particular study, the within-study variation was not accounted for. Thus, a limitation of our study was that the heterogeneity in mean lung function between the studies in our review was relatively limited, especially for severe exacerbations. The majority of studies had a mean FEV₁ % predicted between 35% and 60%, and studies with a very low (<30%) and a very high mean FEV₁ % predicted (>80%) were scarce or completely lacking. However, using a systematic review, the current study reflects the full evidence present in the current literature. This is preferable to using a single patient-level study, which may be biased towards the specific population under study.

Another limitation may be that most of the data were obtained from patients participating in clinical trials each using specific inclusion criteria. We included data from 28 clinical trials including 6780 patients and nine cohort studies including 2211 patients. Trial populations may be biased towards a lower exacerbation frequency because they include clinically stable patients with no other major comorbidities and who are motivated to participate in a trial. However, an overestimation could also be possible, because a large number of trials included only patients with at least one or two exacerbations in the year before inclusion. The cohort studies included in our review used similar inclusion criteria as the trials, and therefore probably included similar patient populations. No systematic difference in exacerbation rate was found between the cohort studies and trials. How these compare with the COPD population seen in daily practice is difficult to determine. One indication may be found in large retrospective database analyses. These studies used event-based definitions and usually found lower exacerbation frequencies than our study, which gives us confidence that we did not underestimate exacerbation frequencies.

Exacerbations depend on the season, and are more likely to occur in winter. Therefore, according to current recommendations, studies need to have a follow-up of at least 12 months or recruitment should be spread throughout the year to give reliable estimates of exacerbation frequency. A strength of our study is that the majority of trials (89%) had a follow-up of at least six months and 65% had a follow-up of at least one year. Conversion of exacerbation rates from studies with a follow-up duration of less than 12 months to annual rates may have overestimated or underestimated the exacerbation frequency. However, we did not find a significant difference between studies with a follow-up duration shorter and longer than 12 months.

Table 2: Estimates regression coefficients, covariance, and smearing factors for the relation between FEV₁ % predicted and annual exacerbation rate described as: annual exacerbation frequency = \( \frac{a}{b} \exp(a + b \cdot \text{FEV}_1 \% \text{ predicted}) \)

<table>
<thead>
<tr>
<th>Intercept: a</th>
<th>Total exacerbations: Event-based definition</th>
<th>Total exacerbations: Symptom-based definition</th>
<th>Severe exacerbations</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.181 (0.351), ( P = 0.004 )</td>
<td>0.981 (0.364), ( P = 0.01 )</td>
<td>-1.043 (0.904), ( P = 0.27 )</td>
<td></td>
</tr>
<tr>
<td>Coefficient FEV₁ % predicted: b</td>
<td>-0.014 (0.007), ( P = 0.053 )</td>
<td>-0.009 (0.007), ( P = 0.19 )</td>
<td>-0.013 (0.020), ( P = 0.51 )</td>
</tr>
<tr>
<td>Covariance intercept and coefficient</td>
<td>-0.00227</td>
<td>-0.00227</td>
<td>-0.0176</td>
</tr>
<tr>
<td>Smearing factor: ( \varphi )</td>
<td>0.893</td>
<td>0.960</td>
<td>1.072</td>
</tr>
</tbody>
</table>

Note: Values are mean (standard error of the mean), \( P \) value.

Abbreviation: FEV₁, forced expiratory volume in one second.

Table 3: Estimated annual exacerbation frequency per GOLD stage based on the regression equations (95% uncertainty interval)

<table>
<thead>
<tr>
<th>GOLD stage</th>
<th>Mean FEV₁ % predicted</th>
<th>Total exacerbations: Event-based definition</th>
<th>Total exacerbations: Symptom-based definition</th>
<th>Severe exacerbations</th>
</tr>
</thead>
<tbody>
<tr>
<td>I: Mild COPD (FEV₁ % pred ≥ 80%)</td>
<td>90</td>
<td>0.82 (0.46–1.49)</td>
<td>1.15 (0.67–2.07)</td>
<td>0.11 (0.02–0.56)</td>
</tr>
<tr>
<td>II: Moderate COPD (50% ≤ FEV₁ % pred &lt; 80%)</td>
<td>65</td>
<td>1.17 (0.93–1.50)</td>
<td>1.44 (1.14–1.87)</td>
<td>0.16 (0.07–0.33)</td>
</tr>
<tr>
<td>III: Severe COPD (30% ≤ FEV₁ % pred &lt; 50%)</td>
<td>42</td>
<td>1.61 (1.51–1.74)</td>
<td>1.76 (1.70–1.88)</td>
<td>0.22 (0.20–0.23)</td>
</tr>
<tr>
<td>IV: Very severe COPD (FEV₁ % pred &lt; 30%)</td>
<td>23</td>
<td>2.10 (1.51–2.94)</td>
<td>2.09 (1.57–2.82)</td>
<td>0.28 (0.14–0.63)</td>
</tr>
</tbody>
</table>

Abbreviations: FEV₁, forced expiratory volume in one second; COPD, chronic obstructive pulmonary disease; GOLD, global initiative for chronic obstructive lung disease; pred, predicted.
To validate the exacerbation frequencies found in our study, they may be compared with the limited patient-level data on the exacerbation frequency specified by subgroup of lung function. The cohort study of Andersson et al, which was included in the review, was the only study providing estimates for four COPD severity stages, using almost the same cutoff points for the stages as the GOLD classification. The study used an event-based definition for exacerbations, and found an annual exacerbation frequency of 0.67 for mild, 0.70 for moderate, 1.06 for severe, and 2.56 for very severe COPD, which was somewhat lower than our estimates, except for very severe COPD. Vestbo et al reported on the exacerbation frequencies in several cohort studies and placebo arms of trials in relation to the FEV$_1$% predicted, and also found exacerbation frequencies below 1.0 for patients with an FEV$_1$% predicted above 50%. The average values for exacerbations for patients with an FEV$_1$% predicted between 40% and 50% ranged between 1.0 and 1.5, which was comparable with our results. Burge et al showed the number of exacerbations per year in the placebo arm of the ISOLDE (Inhaled Steroids in Obstructive Lung Disease) trial using an event-based definition and specified the frequency for three lung function categories, ie, <1.25, 1.25–1.54, and >1.54 liters (about comparable with <45%, 44%–55%, and >55% predicted). Below 45% predicted a mean of 2.6 exacerbations was found, while above >55%, the average value was about 1.2. From the aforementioned studies, the general picture seems to be that, above 50% predicted, the total annual exacerbation frequency is around or slightly below 1.0, while below 40%–45% predicted, the exacerbation rate increases significantly to about two or more exacerbations per year. The results of our study showed the same picture.

In conclusion, the current study provides an estimate of the association between annual exacerbation frequency and FEV$_1$% predicted in COPD, based on aggregated summary data from individual studies. Results were in line with the few studies reporting on this relationship using patient-level data. The resulting GOLD stage-specific exacerbation frequencies show overlapping uncertainty intervals, and hence any analysis based on these rates should be accompanied by a proper sensitivity analysis.

Acknowledgment
This study was supported financially by the Netherlands Asthma Foundation.

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

References


