Rate of normal lung function decline in ageing adults: a systematic review of prospective cohort studies

Elizabeth T Thomas,1,6 Michelle Guppy,1,3 Sharon E Straus,4,5 Katy J L Bell,1,6 Paul Glasziou1

ABSTRACT

Objective To conduct a systematic review investigating the normal age-related changes in lung function in adults without known lung disease.

Design Systematic review.

Data sources MEDLINE, Embase and Cumulative Index to Nursing and Allied Health Literature (CINAHL) were searched for eligible studies from inception to February 12, 2019, supplemented by manual searches of reference lists and clinical trial registries.

Eligibility criteria We planned to include prospective cohort studies and randomised controlled trials (control arms) that measured changes in lung function over time in asymptomatic adults without known respiratory disease.

Data extraction and synthesis Two authors independently determined the eligibility of studies, extracted data and assessed the risk of bias of included studies using the modified Newcastle–Ottawa Scale.

Results From 4385 records screened, we identified 16 cohort studies with 31 099 participants. All included studies demonstrated decline in lung function—forced expiratory volume in 1 s (FEV1), forced vital capacity (FVC) and peak expiratory flow rate (PEFR) with age. In studies with longer follow-up (>10 years), rates of FEV1 decline ranged from 17.7 to 46.4 mL/year (median 22.4 mL/year). Over all, men had faster absolute rates of decline (median 43.5 mL/year) compared with women (median 30.5 mL/year). Differences in relative FEV1 change, however, were not observed between men and women. FEV1/FVC change was reported in only one study, declining by 0.29% per year. An age-specific analysis suggested the rate of FEV1 function decline may accelerate with each decade of age.

Conclusions Lung function—FEV1, FVC and PEFR—decline with age in individuals without known lung disease. The definition of chronic airway disease may need to be reconsidered to allow for normal ageing and ensure that people likely to benefit from interventions are identified rather than healthy people who may be harmed by potential overdiagnosis and overtreatment. The first step would be to apply age, sex and ethnicity-adjusted FEV1/FVC thresholds to the disease definition of chronic obstructive pulmonary disease.

PROSPERO registration number CRD42018087066.

INTRODUCTION

In 2016, the WHO estimated that chronic obstructive pulmonary disease (COPD) affected 251 million people worldwide, with its prevalence continuing to rise with an ageing population. Current guidelines in UK, Australasia, Europe and the USA recommend that COPD is diagnosed if an individual has symptoms such as dyspnoea or sputum production, if they have known risk factors such as smoking or biomass fuel exposure and if they demonstrate post-bronchodilator airflow limitation on spirometry. Airflow limitation on spirometry is defined when the ratio of forced expiratory volume in 1 s (FEV1) to forced vital capacity (FVC) is less than 70% after bronchodilator administration. However, this arbitrary diagnostic threshold has attracted criticism as it does not adjust for age or sex.

Ageing is invariably accompanied by changes in lung function due to factors such as loss of lung elasticity, weakened muscles of respiration, and decreased surface area for alveolar gas exchange. Several published cross-sectional studies and longitudinal studies report that lung function parameters such as FEV1 and FVC decline with age. The 2018 update of the Global Initiative for Chronic Obstructive Lung Disease (GOLD)
criterion continues to suggest the use of the fixed ratio rather than an FEV₁ or FVC that lies outside of the lower limit of normal (LLN) range. While the fixed ratio threshold may be simple for clinicians to use, it does not consider that lung function measurements may change with age and vary with gender and ethnicity. Many laboratory tests already have different reference range values for different ages and electronic spirometry machines do the same. The GOLD criteria acknowledge that this arbitrary fixed threshold may overdiagnose normal healthy older people with disease as healthy. The GOLD criteria acknowledge that this arbitrary fixed threshold may overdiagnose normal healthy older adults as diseased and underdiagnose some younger adults as diseased and undiagnose some younger people with disease as healthy.

Longitudinal studies need to be identified so that normal changes in lung function can be calculated for different ages. Monitoring change could be used in practice to complement a single time point measurement to identify people who are not within the expected normal range. We aimed to perform a systematic review of prospective cohort studies and randomised controlled trials that examined changes in lung function with age in asymptomatic individuals with no known lung disease who have never smoked. This knowledge would enable further work to develop age-specific, sex-specific and ethnicity-specific estimates that may be especially useful in a primary care setting. This implies that people are only diagnosed with COPD if their spirometry measurements fall outside of the normal range for their age, sex and ethnicity, rather than on the basis of a fixed value.

METHODS

Protocol registration
The protocol for this review was drafted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement and the Meta-analyses of Observational Studies in Epidemiology (MOOSE) reporting guidelines. It was registered on PROSPERO and is available from http://www.crd.york.ac.uk/PROSPERO/display_record.php?ID=CRD42018087066, see online supplementary file 1.

Search strategy and inclusion criteria
We conducted electronic searches of MEDLINE, EMBASE and Cumulative Index to Nursing and Allied Health Literature (CINAHL) databases from inception through to February 2019, using the search strategy specified in online supplementary file 2. This was developed with an information specialist. Electronic searches were complemented by manual searching through reference lists of studies that were identified for potential inclusion as well as backward and forward searching. We also searched the WHO Clinical Trials registry and ClinicalTrials.gov registries using the key words “normal ageing”, “lung function decline”, “FEV₁ decline”, “FVC decline” and “lung decline”.

We included cohort studies and also planned to include the control arms of randomised controlled trials that measured the decline of lung function in an ageing population. The inclusion criteria were:

- Longitudinal studies that followed some or all of the adult participants past the age of 65 years.
- Three or more measurements of lung function undertaken.
- Studies with a follow-up period of 3 years or longer.
- Studies that measure lung function (ie, FEV₁, FVC, peak expiratory flow rate (PEFR)).

We excluded studies if the participants did not meet the prespecified age criteria; if the population of interest were reported to include smokers or those with risk factors such as occupational inhalation; if participants were reported to have respiratory symptoms such as wheeze, dyspnoea or chronic cough or if the study included participants with known respiratory disease such as asthma or COPD.

Study selection and data extraction
Two authors (ETT, MG) independently screened the titles and abstracts of studies identified in the initial search for eligibility. Prior to commencing screening, a small subset of 50 titles were screened by the two reviewers as a calibration exercise to check for >80% agreement. Similarly, after screening, a calibration exercise was conducted for screening the full texts of the studies and targeting >80% agreement. The remaining full texts were retrieved and reviewed independently by the authors to determine eligibility for inclusion. Non-English publications were translated using Google Translate or with the assistance of a translator. Disagreements were resolved by consensus through discussion or with a third reviewer (PG). If there were multiple reports of the same study, the most recent publication with longest length of follow-up was selected for inclusion, and if the two studies had a similar length of follow-up then the study with the largest sample size was included. Two authors independently extracted data from the studies. The Excel data extraction form was piloted using 10 studies prior to data extraction as a calibration exercise to check for adequate agreement (>80%) between the reviewers. Any disagreements were resolved by consensus or with a third reviewer. Extracted measures included study setting, year and duration, participant eligibility criteria, sample size, participants demographics (ethnicity, gender, baseline age), any known risk factors or exposures, baseline lung function, lung function measurements, number and frequency of measurements, average length of follow-up and loss to follow-up. We also aimed to report the proportion of the cohort that subsequently developed symptoms or disease during follow-up.

We assessed risk of bias of included studies using the six items of the Newcastle–Ottawa Scale (NOS) for assessing quality of included cohort studies. Disagreements were resolved by discussion or a third reviewer.

Assessed factors included:

- Representativeness of the exposed cohort (eg, low risk: random selection; high risk: non-random selection for example, volunteer sampling).
Ascertainment of exposure—age (eg, low risk: from medical records; high risk: self-reported).

Demonstration that the outcome of interest was not present at the start of study (eg, low risk: participants were excluded on the basis of demonstrated airflow limitation; high risk: if participants were not screened).

Assessment of outcome (eg, low risk: spirometry; high risk: subjective measure of lung function).

Adequate duration of follow-up (eg, low risk: ≥3 years of follow-up; high risk: <3 years of follow-up).

Adequate follow-up of cohorts (eg, low risk: <20% attrition, loss to follow-up explained; high risk: >20% attrition, unexplained loss to follow-up).

Studies were assessed as good quality if they had low risk of bias in all six domains, moderate quality if they had low risk of bias in four or five domains and low quality if they had low risk of bias for three or fewer domains.

Statistical analysis
For each study cohort, we extracted the annual decline rates for each lung function measure. If these were not reported, we calculated crude decline rates for all reported lung function measure by subtracting the final measure from the initial measure and dividing the result by the duration of follow-up. If these data were not available, we determined crude rates of decline from the graphs provided or contacted the study authors for original data.

The data were first analysed descriptively using graphs to determine whether it was appropriate to pool the data. For continuous outcomes, the mean difference (MD) (or standardised MD if studies used different measuring scales) and SD were calculated. The data were reported as an annual decline (unit/year). The overall rates of decline and corresponding 95% CIs were presented in a forest plot. We planned to perform a meta-analysis to pool the estimates of decline.

We presented the data by functional parameter (FEV₁, FEV₀.75, FVC, PEFR) and planned to compare annual decline rates by sex and ethnicity in absolute and relative terms, where data were available. We also extracted and presented age-specific decline rates by decade of age if studies reported these data. We planned to separately analyse the data of those who developed disease during follow-up. We also planned to examine for birth cohort effects if the data were available. Sensitivity analyses were planned for study duration greater than 10 years.

Patient and public involvement
Patients were not involved in the design, data extraction or data analysis of this review.

RESULTS

Study characteristics
From searches of MEDLINE, Embase and CINAHL performed on 12 February 2019, we identified 4331 records. An additional 54 records were identified from clinical trial registries and reference list searches. From these, we retrieved 143 papers for full text review; 115 of these did not meet our selection criteria and a further 12 were removed as duplicates. In total, 16 studies were included in the systematic review (with one study contributing two data sets) (figure 1). The studies included 31 099 participants and were conducted between 1959 and 2014 ranging from 5 to 30 years in duration (table 1).

Overall age-related lung function decline
A meta-analysis was not performed due to substantial heterogeneity across the included studies, and a narrative synthesis was undertaken instead. Thirteen studies reported changes in FEV₁ as an outcome. All studies demonstrated a decline with age, with overall rates of decline from each study ranging from 9.9 to 56.0 mL/year (median 29.2 mL/year). Seven of these studies examined the differences in rates of decline between males and females, showing greater absolute FEV₁ decline in males (median 43.5 mL/year) than in females (median 30.5 mL/year) (table 2, figure 2). Relative rates of FEV₁ decline were calculated for men in eight studies and for women in six studies that reported baseline FEV₁ values. There was no statistically significant difference between men and women’s relative change of FEV₁ from baseline (p=0.7). FEV₀.75 decline was reported in one study. This study provided two data sets (follow-up after 15 years, 30 years) provided in table 2.

Five studies reported changes in FVC, with rate of decline estimates ranging from 14.1 mL/year in the
<table>
<thead>
<tr>
<th>Source ID</th>
<th>Cohort</th>
<th>Study duration (years)</th>
<th>Study centres</th>
<th>Study setting</th>
<th>Study period</th>
<th>Sample size</th>
<th>Mean age of sample (years, SD)</th>
<th>% of males</th>
<th>Outcome</th>
<th>Time points of measurement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ahmadi-Abhari, 2014²⁰</td>
<td>EPIC-Norfolk</td>
<td>13</td>
<td>1</td>
<td>England</td>
<td>1993–2011</td>
<td>8062</td>
<td>58.5* (9.2)</td>
<td>45</td>
<td>FEV₁, FVC</td>
<td>3 (0, 4, 13 years)</td>
</tr>
<tr>
<td>Bartholomew, 1998¹¹</td>
<td>Busselton Population Health Surveys</td>
<td>6</td>
<td>1</td>
<td>Australia</td>
<td>1966–1981</td>
<td>1499</td>
<td>41.6 (16.1)</td>
<td>29.7</td>
<td>FEV₁, FVC</td>
<td>3 (0, 3, 6 years)</td>
</tr>
<tr>
<td>Burchfiel, 1995²²</td>
<td>Kuakini Honolulu Heart Program</td>
<td>6</td>
<td>1</td>
<td>USA</td>
<td>1965–1975</td>
<td>1248</td>
<td>54.6†</td>
<td>100</td>
<td>FEV₁</td>
<td>Mean 5.2</td>
</tr>
<tr>
<td>Burrows, 1986²³</td>
<td>Tucson Epidemiological study of obstructive Lung Disease (TESOLD)</td>
<td>9.6</td>
<td>1</td>
<td>USA</td>
<td>1972–1983</td>
<td>466</td>
<td>48.3 (19.1)</td>
<td>33.9</td>
<td>FEV₁</td>
<td>3 (0, 2, 6 years)</td>
</tr>
<tr>
<td>Griffith, 2001²⁴</td>
<td>Cardiovascular Health Study</td>
<td>7</td>
<td>4</td>
<td>USA</td>
<td>1989–1997</td>
<td>5242</td>
<td>73.5/72.7 (5.5)/(5.2)</td>
<td>42.4</td>
<td>FEV₁, FVC</td>
<td>3 (0, 4, 7 years)</td>
</tr>
<tr>
<td>Luoto, 2018²⁷</td>
<td>Good Aging in Skåne</td>
<td>13.5</td>
<td>1</td>
<td>Sweden</td>
<td>2001–2014</td>
<td>387</td>
<td>70.6* (10.6)</td>
<td>44.2*</td>
<td>FEV₁, FVC</td>
<td>Aged &lt;80: every 6 years Aged 80 or over: every 3 years</td>
</tr>
<tr>
<td>Maseklo, 2006²⁸</td>
<td>MacArthur Successful Aging study</td>
<td>7</td>
<td>3</td>
<td>USA</td>
<td>1988–1995</td>
<td>544</td>
<td>74</td>
<td>31.8</td>
<td>PEFR</td>
<td>3 (0, 3, 7 years)</td>
</tr>
<tr>
<td>Pearson, 1998²⁹</td>
<td>Baltimore Longitudinal Study of Aging</td>
<td>Males: 11.5 Females: 5.7</td>
<td>1</td>
<td>USA</td>
<td>1962–1991</td>
<td>173</td>
<td>42.4</td>
<td>52.6</td>
<td>FEV₁</td>
<td>4.6/3 (every 2 years)</td>
</tr>
<tr>
<td>Pelkonen, 2001³⁰</td>
<td>Seven Countries Study</td>
<td>30</td>
<td>2</td>
<td>Finland</td>
<td>1959–1989</td>
<td>200</td>
<td>47.6 (30 years) 49.4 (15 years)</td>
<td>100</td>
<td>FEV₁, FVC</td>
<td>6 (0, 5, 10, 15, 20, 25, 30 years)</td>
</tr>
<tr>
<td>Proctor, 2006³¹</td>
<td>Origins of Variance in the Old-Old (OCTO-Twin)</td>
<td>8</td>
<td>1</td>
<td>Sweden</td>
<td>1991–2003</td>
<td>579</td>
<td>83.2 (2.8)</td>
<td>33.0</td>
<td>PEFR</td>
<td>5 (0, 2, 4, 6, 8 years)</td>
</tr>
<tr>
<td>Sherman, 1992³²</td>
<td>Six Cities study of Air Pollution and Health</td>
<td>12</td>
<td>6</td>
<td>USA</td>
<td>1974–1989</td>
<td>1486</td>
<td>47.2/48.2* (12.3)/(12.5)</td>
<td>32.0</td>
<td>FEV₁, FVC</td>
<td>4 (0, 3, 6, 12 years)</td>
</tr>
<tr>
<td>Wang, 2004³⁴</td>
<td>-</td>
<td>5</td>
<td>1</td>
<td>USA</td>
<td>1985–1992</td>
<td>71</td>
<td>37†* (19-65)</td>
<td>100</td>
<td>FEV₁</td>
<td>3–11; every 6 months</td>
</tr>
</tbody>
</table>
youngest cohort\textsuperscript{32} (median age 36.2 years) to 65.6 mL/year in the older cohort\textsuperscript{24} (mean age 73.0 years). In studies that measured both FEV\textsubscript{1} and FVC over time, there was a greater decline in FEV\textsubscript{1} than FVC in two studies, and greater decline in FVC than FEV\textsubscript{1} in three studies. These measures are average estimates across study participants and do not enable calculation of individuals’ FEV\textsubscript{1}/FVC ratios. In the one study where individuals’ FEV\textsubscript{1}/FVC ratios were reported as an outcome,\textsuperscript{26} there was a decline by 0.29\% per year.

PEFR was reported as an outcome in two studies,\textsuperscript{27,30} which showed decline rates ranging from \(-6.6\) L/min/year in females to \(-11.5\) L/min/year in males.

**Age-specific lung function decline by decade of age**

The age-specific rates of FEV\textsubscript{1} change by decade of age were extracted or calculated from three studies.\textsuperscript{22,23,28,35} In all but one study, estimates of decline increased from the fourth (age 30–40 years) to eighth decades of life (table 3). One study could not be included in this comparative analysis as they included smokers and reported decline rates at the end of study follow-up (rather than baseline age).\textsuperscript{35} This study reported that the rates of relative decline increase from the seventh (\(-1.7\%\) /year) through to the tenth decade (\(-3.1\%\) /year), though absolute rates of decline varied. Another study also reported that the rate of decline may be non-linear in multiple regression models of FEV\textsubscript{1} and FVC decline (where age squared was also a statistically significant variable).\textsuperscript{34}

Two studies examined lung function change within age brackets that did not conform to our decade-specific analysis. Bartholomew and Knuiman\textsuperscript{21} reported greater decline rates in never smokers aged above 45 years (females: \(-30.7\) mL/year, males \(-45.8\) mL/year) compared with those aged below 45 years (females: \(-24.3\) mL/year, males: \(-36.8\) mL/year). Lange \textit{et al}\textsuperscript{25} compared decline rates in both male and female non-smokers in 20-year age groups. Females aged 60–79 years had the greatest decline rates (\(-31.7\pm2.1\) mL/year) compared with the 40–59 age group (\(-17.7\pm1.4\) mL/year) and the 20–39 age group which reported an increase of \(5.0\pm2.7\) mL/year. Similarly, males aged 60–79 years had the greatest decline rates (\(-37.1\pm3.7\) mL/year) compared with the 40–59 year age group (\(-24.2\pm2.6\) mL/year) and the 20–39 year age group (\(-4.6\pm4.2\) mL/year).

**Overall rates of mortality/symptom/disease development**

Few studies reported these outcomes in an initially asymptomatic, non-smoking population. One study (Proctor \textit{et al})\textsuperscript{30} reported 85\% mortality rate in the elderly cohort (age range 79–96) over 8 years. Another study (Lange \textit{et al})\textsuperscript{25} reported that in their study of non-asthmatics, 364 (2\%) patients who did not report having asthma at the beginning of the study, later reported it in follow-up. However, this estimate included smokers. One study (Wang and Petsonk)\textsuperscript{33} performed their analyses on a highly screened population, meaning they excluded participants from all analyses who developed disease or...
<table>
<thead>
<tr>
<th>Source ID</th>
<th>Mean age of sample (years, SD)</th>
<th>Duration</th>
<th>Sample size</th>
<th>Mean absolute unit decline/year (SD)</th>
<th>Overall relative decline (%)</th>
<th>Variables reported to alter the rate of change</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>FEV_{1} (mL)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ahmadi, 2014</td>
<td>58.5*(9.2)</td>
<td>13</td>
<td>3621 4441</td>
<td>−17.7 (78.6)</td>
<td></td>
<td>Smoking; CRP categories</td>
</tr>
<tr>
<td>Bartholomew, 1998</td>
<td>41.6 (16.1)</td>
<td>6</td>
<td>445 1054</td>
<td>−43.5 (100.4)</td>
<td>−30.5 (144.8)</td>
<td>1.1 1.2</td>
</tr>
<tr>
<td>Burchfiel, 1995</td>
<td>54.6†</td>
<td>6</td>
<td>1248</td>
<td>−21.6‡</td>
<td></td>
<td>0.7</td>
</tr>
<tr>
<td>Burrows, 1986</td>
<td>48.3 (19.1)</td>
<td>9.6</td>
<td>158 308</td>
<td>−10.3‡ (6.3)</td>
<td>−9.1‡ (5.7)</td>
<td></td>
</tr>
<tr>
<td>Griffith, 2001</td>
<td>73.0* (5.3)</td>
<td>7</td>
<td>1976* 2604*</td>
<td>−52.3 (3.1)§</td>
<td>−47.0 (2.8)§</td>
<td>1.9 1.7</td>
</tr>
<tr>
<td>Liao, 2015</td>
<td>51.7†</td>
<td>15</td>
<td>1592 2713</td>
<td>−23.5 (10.4)</td>
<td>−18.3 (10.0)</td>
<td>0.8 0.8</td>
</tr>
<tr>
<td>Liou, 2015</td>
<td>47.4* (10.6)</td>
<td>17</td>
<td>207¶ 336¶</td>
<td>−25.8 (14.0)*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Luoto, 2018</td>
<td>70.6* (10.6)</td>
<td>13</td>
<td>171¶ 216¶</td>
<td>−46.4‡ (47.7)</td>
<td></td>
<td>2.2‡</td>
</tr>
<tr>
<td>Pearson, 1998</td>
<td>42.4 11.5/5.7</td>
<td>91 82</td>
<td></td>
<td>−43.5</td>
<td>−35.1</td>
<td>1.0 1.3</td>
</tr>
<tr>
<td>Sherman, 1992</td>
<td>47.9 (12.4)</td>
<td>12</td>
<td>475 1011</td>
<td>−32.8 (23.5)</td>
<td>−27.5 (20.4)</td>
<td>1.0 1.1</td>
</tr>
<tr>
<td>Triebner, 2017</td>
<td>36.2** (19–65)</td>
<td>19.7**</td>
<td>648</td>
<td></td>
<td>−22.4 (96.4)</td>
<td></td>
</tr>
<tr>
<td>Wang, 2004</td>
<td>37** (19–65)</td>
<td>12</td>
<td>475 1011</td>
<td>−32.8 (23.5)</td>
<td>−27.5 (20.4)</td>
<td>1.0 1.1</td>
</tr>
<tr>
<td>Xu, 1995</td>
<td>42.4† (11.9)</td>
<td>24</td>
<td>1418 4875</td>
<td>−28.3 (138.9)</td>
<td>−16.0 (135.5)</td>
<td>0.7 0.5</td>
</tr>
<tr>
<td><strong>FVC (mL)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ahmadi, 2014</td>
<td>58.5* (9.2)</td>
<td>13</td>
<td>3621 4441</td>
<td>−31.1 (118.1)</td>
<td></td>
<td>Smoking; CRP categories</td>
</tr>
<tr>
<td>Bartholomew, 1998</td>
<td>41.6 (16.1)</td>
<td>6</td>
<td>445 1054</td>
<td>−47.2 (104.0)</td>
<td>−36.0 (154.5)</td>
<td>1.0 1.1</td>
</tr>
<tr>
<td>Griffith, 2001</td>
<td>73.0* (5.3)</td>
<td>7</td>
<td>1976* 2604*</td>
<td>−78.4§ (4.2)</td>
<td>−65.6§ (3.8)</td>
<td>2.9 2.4</td>
</tr>
<tr>
<td>Luoto, 2018</td>
<td>70.6* (10.6)</td>
<td>13</td>
<td>171¶ 216¶</td>
<td>−43.7‡ (67.2)</td>
<td></td>
<td>1.7</td>
</tr>
<tr>
<td>Triebner, 2017</td>
<td>36.2** (19–65)</td>
<td>19.7**</td>
<td>648</td>
<td></td>
<td>−14.1 (42.8)</td>
<td></td>
</tr>
<tr>
<td><strong>FEV_{1}/FVC</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Liao, 2015</td>
<td>47.4* (10.6)</td>
<td>17</td>
<td>207* 336*</td>
<td>−0.0029 (0.0023)*</td>
<td></td>
<td>Smoking, less versus more likely dust exposure</td>
</tr>
</tbody>
</table>

Continued
Symptoms during study follow-up. No studies reported the rates of lung function change in those who developed disease during the course of the study compared with those who did not.

Sensitivity analyses
Heterogeneity in study duration was explored in figure 3. After removing studies with a follow-up of less than 10 years, the median rate of decline of FEV$_1$ was 22.4 mL/year (figure 4).

Predictors of the rate of decline in lung function in people without known lung disease

Smoking
Although smokers were not included in our main analysis, some studies did compare non-smokers and smokers which we report here. The decline rates were compared in non-smokers or former smokers with current smokers in nine studies.$^{21\text{--}27,29,31,35}$ In the seven studies measuring FEV$_1$, decline, current smokers were observed to have a faster rate of decline.$^{21\text{--}26,31,35}$ In the three studies measuring FVC, smoking increased FVC decline.$^{21,24,35}$ FEV$_1$/FVC decline was greater in smokers than non-smokers in one study$^{26}$ and FEV$_{0.75}$ in another study.$^{29}$

Body mass index (BMI)
Three studies reported the association of BMI with FEV$_1$ change. In the study by Bartholomew and Knuiman,$^{21}$ increased BMI significantly affected FEV$_1$ decline ($p=0.008$ for females; $p=0.007$ for males). However, an estimate for this association was not provided. In Triebner et al.,$^{32}$ obese individuals reported greater declines of FEV$_1$ (29 mL/year) and FVC (25 mL/year) compared with individuals with normal BMI (FEV$_1$ 22 mL/year, FVC 19 mL/year). In Luoto et al.,$^{35}$ having a BMI greater than 35 was significantly associated with a slower decline of FEV$_1$ (32 mL/year compared with 46 mL/year, $p=0.04$), but it did not significantly affect FVC decline.

Ethnicity
Griffith et al$^{24}$ was the only study that assessed ethnicity, specifically comparing African-American participants to white participants. We did not include the African-American cohort in our analysis as only two measurements were performed on this population. However, FEV$_1$ and FVC declines were greater in whites compared with African-Americans.

Systolic blood pressure
Griffith et al,$^{24}$ examined the correlation of systolic blood pressure greater than 160 mm Hg with FEV$_1$ and FVC decline and found that declines were on average 5.6 mL/year and 10.9 mL/year greater, respectively ($p<0.01$).

Dust exposure
Liao et al$^{28}$ explored the effects of dust exposure on FEV$_1$ and FEV$_1$/FVC decline. Participants with more dust exposure experienced a mean FEV$_1$ decline that was 4.5 mL/
year greater than participants with less dust exposure (p=0.007). Dust exposure did not significantly affect FEV1/FVC ratio decline, suggesting that FVC declined in parallel to FEV1.

Menopausal status

Triebner et al. reported that menopausal status affected the rate of decline, with rates of FEV1 decline on average 3.8 mL/year greater in perimenopausal women, and 5.2 mL/year greater in postmenopausal women. FVC decline was 10.2 mL/year greater in perimenopausal women and 12.5 mL/year greater in postmenopausal women, compared with premenopausal women.

Risk of bias

Risk of bias was determined using a modified version of the NOS (figures 5 and 6). No studies received low risk of bias in all domains, but four studies had a low risk of bias in all but one domain.

Table 3  Age-specific lung function decline by decade of age as reported in four cohorts

<table>
<thead>
<tr>
<th>Study ID</th>
<th>Sample size (n)</th>
<th>Baseline age (years)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>30–39</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Male</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1248</td>
</tr>
<tr>
<td>Burrows, 1985</td>
<td>22</td>
<td>2.83</td>
</tr>
<tr>
<td>Burrows, 1986</td>
<td>23</td>
<td>2.83</td>
</tr>
<tr>
<td>Pearson, 1995</td>
<td>28</td>
<td>2.83</td>
</tr>
</tbody>
</table>

The estimates from Burrows were derived from formulae modelling change in FEV1 with age. See online supplementary file 3 for calculations.

Figures 2 The rate of forced expiratory volume in 1 s (FEV1) decline in 13 study populations, grouped by sex.

Table 3  Age-specific lung function decline by decade of age as reported in four cohorts

<table>
<thead>
<tr>
<th>Absolute mean decade-specific FEV1 function decline rates (mL/year)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline age (years)</td>
</tr>
<tr>
<td>Study ID</td>
</tr>
<tr>
<td>----------</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Burrows, 1985</td>
</tr>
<tr>
<td>Burrows, 1986</td>
</tr>
<tr>
<td>Pearson, 1995</td>
</tr>
</tbody>
</table>

The estimates from Burrows were derived from formulae modelling change in FEV1 with age. See online supplementary file 3 for calculations.

*Estimates adjust for covariates including height and age.
†Includes participants aged 45–49 years.

FEV1, forced expiratory volume in 1 s.
spirometer to measure lung function which is a validated objective instrument. All studies (100%) had adequate duration of follow-up (3 years or longer). Eight studies (50%) had a high risk of bias for having high attrition rates in their studies (>20%).

**DISCUSSION**

**Statement of principal findings**

This systematic review of 16 prospective cohort studies conducted in 13 countries provides a summary of all the available evidence looking at lung function change with age. Lung function declines with age in normal, asymptomatic adults with higher rates of decline in absolute lung function parameters in men compared with women. However, the relative rates of decline from baseline between men and women do not differ significantly. The decline in absolute and relative lung function parameters may accelerate with age and is also exacerbated by smoking. We were unable to compare lung function decline rates of different ethnicities due to insufficient data. There was a paucity of longitudinal studies that reported changes in FEV₁/FVC rather than reporting the two parameters separately.

**Figure 3** The rate of forced expiratory volume in 1 s (FEV₁) decline in 13 study populations by years of follow-up. The size of the circle corresponds to individual study sample size.

**Figure 4** Sensitivity analysis, excluding studies with less than 10 years of follow-up. The size of the circle corresponds to individual study sample size. FEV₁, forced expiratory volume in 1 s.
This systematic review examined all the available primary studies to allow an examination of the consistency of estimates of decline in FEV₁, FVC, FEV₁/FVC ratio and PEFR. This review particularly focused on older adults; this group is relatively understudied and yet more prone to overdiagnosis and misdiagnosis. While the majority of current prediction equations of lung function are based on cross-sectional studies, our review searched for longitudinal studies as they change in lung function may provide a complement to measurement at one time point in predicting future lung function. Our review included participants who were ageing normally, but may have had non-pulmonary comorbidities such as hypertension and diabetes mellitus. This enabled us to investigate a population that was more representative of a normal ageing population.

Our review has some limitations. We did not have access to unpublished individual participant data to allow the calculation of FEV₁/FVC for the majority of studies, where this were not reported. Five studies separately measured changes in both FEV₁ and FVC, however it is difficult to conclude whether the rate of decline in FEV₁ and FVC is proportional. Out of the five studies that reported both FEV₁ and FVC decline, two studies demonstrated that FEV₁ declines faster than FVC, but in the three remaining studies, the FVC declined at a faster rate (see table 2). Longitudinal studies that specifically measure the FEV₁/FVC would provide the most reliable measure of this decline. Individual patient data would also allow a more robust analysis of changes in lung function between individuals in the studies.

We were unable to pool the results due to significant heterogeneity across the populations. This review’s findings are also limited by the quality of the included studies, all of which were judged moderate or low quality. Since this review is based on limited populations, the findings may not be generalisable to all individuals, especially those of non-Caucasian ethnicities or from less economically developed countries where smoking and air pollution may be more prevalent for example. The review’s findings may underestimate lung function decline among asymptomatic people, as volunteer bias may be present with cohort studies where healthier individuals may be more likely to participate. Our study aimed to examine the rate of lung function change in the elderly group; however, the majority of included studies did not focus on this age group. COPD misdiagnosis particularly affects those older than 80 years of age; therefore, more studies are required in the elderly group.

Our review did not consider the extent of short-term within-person variation, or ‘noise’, in lung function measurements, which is likely to be considerable. Any observed change in measurement is a combination of the true change, or ‘signal’, and the random background ‘noise’. The clinical utility of monitoring lung function to decide whether or not COPD is present is in part determined by the ratio of signal to noise in
the measurements. Changes in measured lung function over a longer period of time may be more likely to indicate some true change rather than just background noise; therefore, we specified in our inclusion criteria that eligible studies should measure lung function on a minimum of three occasions.

We observed substantial heterogeneity across all the included studies and results. This may be due to inherent differences within the populations studied (including distribution of ages, proportion of men vs women and ethnicities) or the duration of follow-up, or that decline in normal healthy people may vary across individuals without causing disease. We explored differences in duration of follow-up as a potential source of heterogeneity in a sensitivity analysis excluding studies with less than 10 years of follow-up, but found that this did not change the median estimate substantially.

Quality of spirometry, as well as properly maintained and calibrated equipment, causing measurement error and contributing to the ‘noise’ in measurement discussed above, is likely to have contributed to variation in the results. Only nine of the included studies specifically reported that the spirometers used in their studies were calibrated and the measurements had to be acceptable and reproducible, following the American Thoracic Society guidelines on the standardisation of spirometry. Two studies used peak flow meters to measure PEFR. These instruments are well known to vary in consistency and accuracy. Maselko et al used a Mini-Wright meter, but the second study by Proctor et al did not specify which peak flow meter they used.

The majority of studies specified that they excluded patients with known disease or symptoms at the commencement of the study. However, most of the studies did not report whether any of the participants in their study sample developed symptoms or respiratory disease in the course of follow-up. Thus, undiagnosed COPD or other respiratory, cardiac, renal or other diseases that cause decline in lung function may have contributed to heterogeneity in the results.

Variation within the results may be also explained by the ‘horse-racing effect’, where an initially low FEV1 measurement may reflect a greater loss of function in the preceding years and hence predicts faster decline in subsequent years (just as the position of the horse in halfway through the race is related to its speed in the early part of the race and hence speed for the final part of the race). Regression to the mean, due to inclusion of people with randomly high (or low) measured lung function in the primary studies, may also have contributed to heterogeneity of the results. A simple way that primary studies may assess for a horse racing effect, while allowing for regression to the mean, is by constructing Bland–Altman plots of change versus mean FEV1 level (or substituting PEFR for mean FEV1 as these are highly correlated).

Comparison with previous research

To date, there have been no systematic reviews or meta-analyses examining the rate of lung function decline with age to assess the potential impact of the fixed threshold on COPD misdiagnosis. Cross-sectional studies have compared people diagnosed with COPD using a fixed threshold and the LLN definition, reporting that the GOLD criteria leads to misdiagnosis of COPD. A prospective cohort study found that the fixed threshold of the GOLD criteria overdiagnosed a large proportion of elderly people over the age of 70 years, and the LLN criteria tended to underdiagnose COPD, when compared with the reference standard which consisted of an expert panel who used all available diagnostic information including spirometry.

Meaning of the study: possible explanations and implications for clinicians and policy-makers

This review has found that lung function declines with age in all studied populations. The rate of decline appears to accelerate with age, and age-specific estimates of FEV1, FVC and FEV1/FVC ratio may be more appropriate for the diagnosis of COPD than the fixed threshold.
currently used across all ages. Currently, prediction equations for calculating mean lung function values as well as the lower limit of normal (LLN) for all ages are based on data from cross-sectional studies; however; it is argued that this is problematic as they do not factor in the important dimension of time.\textsuperscript{51, 52} Spirometers used in practice commonly derive their reference values from the National Health and Nutrition Examination Survey (NHANES), a cross-sectional study which was conducted in the USA between 1988 and 1994. Though the predicted values do reflect a decline in FEV\textsubscript{1} and FEV\textsubscript{1}/FVC with age, these decline rates may not be as reliable as the estimates from longitudinal studies included in our review. According to the NHANES III, the median rate of FEV\textsubscript{1} decline for a Caucasian male with an average height of 1.75 m aged between 30 and 80 is 32 mL/year and a female with an average height of 1.6 m has an FEV\textsubscript{1} that declines a median of 25 mL/year. Both of these estimates are lower than the median FEV\textsubscript{1} decline of the studies included in our review. Declines a median of 25 mL/year for a Caucasian male with an average height of 1.75 m aged between 30 and 80 is 32 mL/year and a female with an average height of 1.6 m has an FEV\textsubscript{1} that declines a median of 25 mL/year. Both of these estimates are lower than the median FEV\textsubscript{1} decline of the studies included in our review, which was 43.5 mL/year and 30.5 mL/year for men and women, respectively. Therefore, the predicted age-specific lung function used in spirometers may often mislabel people as having abnormal lung function when they are actually within normal limits.\textsuperscript{53} More reliable age-specific estimates and prediction equations are required.

Clinicians need to consider whether ‘abnormal’ spirometry results may in fact represent normal ageing. This is especially true for making a formal diagnosis of COPD. If a patient is symptomatic and has airflow obstruction as defined by GOLD criteria, it may be necessary to consider alternative diagnoses such as a dyspnoea of cardiac origin. One proposal for identifying individuals who are experiencing greater loss of lung function than expected is to develop ‘decline charts’ that predict FEV\textsubscript{1} or FEV\textsubscript{1}/FVC loss for different ages. This can allow clinicians to monitor lung function over time and assess whether individuals are tracking along expected decline curves. These would also need to account for noise in measurement.

Future research should focus on conducting long-term longitudinal studies in less-studied populations, with emphasis on older adults. These studies should examine the rates of decline in people who eventually become symptomatic or develop disease. This information can guide clinicians to predict what rate of lung function decline may be a prognostic indicator of COPD onset and progression. Further well-designed prospective studies that investigate changes in FEV\textsubscript{1}/FVC may allow for the development of algorithms that predict individuals’ expected lung function over time according to their sex, smoking history, age, BMI and ethnicity. The observed change in lung function parameters might then be compared with the expected change to help the clinician determine whether this is extreme enough to warrant diagnosis of disease.

**Acknowledgements** The authors thank Mr Justin Clark for his assistance with the literature search, Ms Mari Tashiro for her help with the translation of Japanese studies, Dr Mark Jones for his advice on statistical analysis and Dr Claudia Dobler for her specialist input on this review.

**Contributors** ETT was involved with devising the review methods, conducting electronic searches, screening of abstracts, data extraction, data analysis and interpretation and co-drafting of the review. MG was involved with devising the review methods, screening of abstracts, data extraction, data analysis and interpretation and co-drafting the review. KJLB was involved with devising the review methods, data analysis and interpretation and co-drafting the review. SES was involved with devising the review methods, data analysis and interpretation and co-drafting the review. PG was involved with devising the review methods, data analysis and interpretation and co-drafting the review.

**Funding** KJLB and PG have received funding from the Australian National Health and Medical Research Council (Centre for Research Excellence Grant No 1104136, Australia Fellowship No 527500 and Program Grant No 633003). The funders had no role in design and conduct of the study; collection, management, analysis and interpretation of the data; and preparation, review or approval of the manuscript.

**Competing interests** None declared.

**Patient consent for publication** Not required.

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

**Data sharing statement** No additional data are available.

**Author note** This manuscript was prepared in accordance with the PRISMA statement and the MOOSE reporting guidelines.

**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

**REFERENCES**


