external validity of randomised controlled trials in asthma: to whom do the results of the trials apply?

Justin Travers, Suzanne Marsh, Mathew Williams, Mark Weatherall, Brent Caldwell, Philippa Shirtcliffe, Sarah Aldington, Richard Beasley

Background: Asthma is a heterogeneous disease with a wide range of clinical phenotypes, not all of which may be encompassed in the subjects included in randomised controlled trials (RCTs). This makes it difficult for clinicians to know to what extent the evidence derived from RCTs applies to a given patient.

Aim: To calculate the proportion of individuals with asthma who would have been eligible for the major asthma RCTs from the data of a random community survey of respiratory health.

Methods: A postal survey was sent to 3500 randomly selected individuals aged 25–75 years. Respondents were invited to complete a detailed respiratory questionnaire and pulmonary function testing. Participants with current asthma were assessed against the eligibility criteria of the 17 major asthma RCTs cited in the Global Initiative for Asthma (GINA) guidelines.

Findings: A total of 749 participants completed the full survey, of whom 179 had current asthma. A median 4% of participants with current asthma (range 0–36%) met the eligibility criteria for the included RCTs. A median 6% (range 0–43%) of participants with current asthma on treatment met the eligibility criteria.

Interpretation: This study shows that the major asthma RCTs on which the GINA guidelines are based may have limited external validity as they have been performed on highly selected patient populations. Most of the participants with current asthma on treatment in the community would not have been eligible for these RCTs.

In recent years, clinical decision making has been directed away from the doctor's clinical experience towards a paradigm based on the evidence from randomised controlled trials (RCTs). The results of large RCTs have been translated into guidelines containing evidence-graded recommendations, which the clinicians are encouraged to accept as the basis of good clinical practice. However, RCTs are only able to guide clinical decision making when trials are well designed, have clinically relevant outcome measures and the subjects in the trial are representative of the range of real-life patients managed by the doctor. This last requirement is not always met, and it is recognised that older adults, women and ethnic minorities may be under-represented in RCTs. Design considerations often lead to RCTs that are performed in highly selected patient populations, such as those with the most typical features of a disease or those most likely to respond to the intervention being studied. This may be necessary to initially assess the efficacy of an intervention—for example, in a phase II clinical trial. However, when these same design considerations are applied to phase III and IV clinical trials, they may result in the exclusion of many subjects in whom treatment may be potentially useful, thereby restricting the generalisability of the trial results.

Asthma is a heterogeneous disease with a wide range of clinical phenotypes, hence, it is not surprising that there is evidence that many individuals with asthma are not eligible for RCTs owing to highly selective inclusion and exclusion criteria. This makes it difficult for clinicians to know to what extent the evidence for the safety and effectiveness of an intervention applies to a given individual. The proportion of individuals with asthma who are eligible for the major asthma RCTs from which the clinical evidence is derived is not known. In this population-based survey of respiratory health, we determined the proportion of individuals with asthma who would have met the eligibility criteria for the RCTs that form the basis of asthma consensus guidelines.

Methods

Participants

Data of the participants were obtained from the results of the Wellington Respiratory Survey, a detailed survey of respiratory health performed between 2002 and 2005 in Wellington, New Zealand. Participants were recruited from a postal questionnaire sent to 3500 individuals randomly selected from the electoral register. Random selection was performed so that equal numbers of questionnaires were sent to men and women in each of the five-decade age groups between the ages of 25 and 75 years. Participants who responded to the postal questionnaire were invited to undertake the full survey that included a detailed, interviewer administered questionnaire, pulmonary function tests, chest CT scan, skin-prick tests to common allergens, blood tests for eosinophil count and serum immunoglobulin E and a 1 week peak flow diary.

Pulmonary function testing

Pulmonary function tests were carried out using two Jaeger Master Screen Body volume constant plethysmography units with a pneumotachograph (Masterlab 4.5 and 4.6 Erich-Jaeger, Wurzberg, Germany) as described previously. Static and dynamic lung volumes were measured before and after giving 400 μg of salbutamol via a spacer device. Peak flow readings were recorded by participants twice daily over a 1 week period following instructions in the use of a Breath Alert peak flow meter (Medical Developments International, Melbourne, Australia). Participants were not tested within 3 weeks of an upper or lower respiratory tract infection. The survey was approved by the Wellington Ethics Committee, and written informed consent was obtained from each participant.

Abbreviations: COPD, chronic obstructive pulmonary disease; FEV₁, forced expiratory volume in 1 s; GINA, Global Initiative for Asthma; RCT, randomised controlled trial
Asthma RCTs were identified in a systematic manner. To identify RCTs, they reported the use of drugs for asthma in the previous 12 months, showed an increase in the forced expiratory volume in 1 s (FEV₁) ≥15% compared with baseline after bronchodilator administration, and/or documented diurnal peak flow variation of ≥20% in any of the first 7 days of recordings.

Symptoms of asthma were wheeze, shortness of breath and wheeze at night or chest tightness at night. A participant with current asthma was identified as having ‘current asthma on treatment’ if they reported asthma diagnosed by a doctor and either symptoms or use of drugs for asthma in the previous 12 months.

### Identification of current asthma
Participants were identified as having “current asthma” if they:
- reported asthma diagnosed by a doctor and either symptoms of asthma or use of drugs for asthma in the previous 12 months,
- showed an increase in the forced expiratory volume in 1 s (FEV₁) ≥15% compared with baseline after bronchodilator administration, and/or
- documented diurnal peak flow variation of ≥20% in any of the first 7 days of recordings.

### Identification of RCTs
Asthma RCTs were identified in a systematic manner. To qualify as a trial forming the basis of consensus guidelines, an RCT had to be cited as a reference accompanying a level A or B evidence-graded treatment recommendation in the Global Initiative for Asthma (GINA) Workshop Guideline ‘Global strategy for asthma management and prevention: 2005 Update’. Trials had to be RCTs of drug treatment for asthma in adults, with at least 400 participants randomised, and have been published in the past 30 years. References were assessed independently by two reviewers (JT, BC). Inclusion and exclusion criteria were obtained from the full text of all qualifying trials.

### Analysis
The proportion of participants with current asthma who met the eligibility criteria for each of the identified RCTs was determined. Whenever we were unable to determine whether a participant met a particular RCT eligibility criterion from our survey data, the participant was considered to have met that criterion. For example, an RCT may have a criterion that participants be free from exacerbations in the previous 2 months, when our survey recorded only that participants were exacerbation free in the past 3 weeks. In this case, all participants who were exacerbation free for 3 weeks were considered to meet this criterion and remain potentially eligible.

The sponsor had no involvement in the study design, collection, analysis interpretation of data, the writing of the report or the decision to submit for publication.

### RESULTS
A total of 2319 participants responded to the postal survey, representing a response rate of 78%. Of these respondents, 749 participants completed the detailed questionnaire and satisfactory pulmonary function testing, and formed the study group (fig 1). Compared with the 1570 survey respondents who were not included in the study group, the study group had a higher rate of asthma diagnosed by the doctor (23.1% v 17.3%), more men (53.7% v 44.3%) and ex-smokers (41.4% v 35.3%). There were no significant differences in the prevalence of chronic bronchitis or emphysema diagnosed by a doctor.

Of the 749 participants in the study group, 179 (24%) met our criteria for current asthma and 127 (17%) met our criteria for current asthma on treatment. Among the 179 participants with current asthma, 67 also met the criteria for chronic obstructive pulmonary disease (COPD), defined as an FEV₁/forced vital capacity ratio <0.7 postbronchodilator. Of these 67 participants, 29 had a tobacco cigarette history of >10 pack-years.

### Table 1 Characteristics of participants with current asthma (n = 179)

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Mean (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>55.2 (13.6)</td>
</tr>
<tr>
<td>FEV₁ pre-bronchodilator as a percentage of predicted</td>
<td>77.8 (21.8)</td>
</tr>
<tr>
<td>Women (%)</td>
<td>87 (49)</td>
</tr>
<tr>
<td>Smoker (%)</td>
<td>54 (30)</td>
</tr>
<tr>
<td>Asthma diagnosed by doctor (%)</td>
<td>137 (77)</td>
</tr>
<tr>
<td>Current symptoms (%)</td>
<td>122 (68)</td>
</tr>
<tr>
<td>Any drug for asthma used in the previous 12 months (%)</td>
<td>127 (71)</td>
</tr>
<tr>
<td>Inhaled corticosteroid use in the previous 12 months (%)</td>
<td>93 (52)</td>
</tr>
<tr>
<td>Inhaled short-acting β-agonist use in the previous 12 months (%)</td>
<td>114 (64)</td>
</tr>
<tr>
<td>COPD (%)</td>
<td>67 (37)</td>
</tr>
<tr>
<td>Peak flow variability ≥20% (%)</td>
<td>61 (34)</td>
</tr>
<tr>
<td>Bronchodilator reversibility ≥15% (%)</td>
<td>43 (24)</td>
</tr>
</tbody>
</table>

COPD, chronic obstructive pulmonary disease; FEV₁, forced expiratory volume in 1 s.

*Current or ex-smoker with >10 pack-years of exposure to cigarette smoke.

Wheeze, shortness of breath and wheeze at night or chest tightness at night in the past year.

Defined as a FEV₁/forced vital capacity ratio <0.7 after bronchodilator.
Table 1 presents the characteristics of the participants with current asthma.

There were 215 individual references in the relevant chapter of the GINA guidelines, from which 17 qualifying RCTs were identified and included in the analysis (fig 2). Table 2 gives the characteristics of the included RCTs. The number of participants screened was stated in 8 of the 17 RCTs included in the analysis.

Inclusion criteria used in all 17 RCTs were a diagnosis of asthma, age greater than a lower age limit and bronchodilator reversibility. Other inclusion criteria were a specified FEV₁ in 16, specified inhaled corticosteroid use in 12, specific symptoms or use of rescue drugs in 9, age less than an upper age limit in 7, peak flow variability in 4 and other inclusion criteria in 4 RCTs. Exclusion criteria used were upper age limit in 7, peak flow variability in 4 and other symptoms or use of rescue drugs in 9, age less than an age range in 16, specified inhaled corticosteroid use in 12, specified current asthma.

**DISCUSSION**

This study shows that the major asthma RCTs have been performed on highly selected patient populations, with a median of only 4% of participants with current asthma in our community being eligible. Similarly, a median of only 6% of participants with current asthma on treatment would have been eligible for these RCTs. These findings suggest that although the treatment recommendations of major asthma guidelines have a strong scientific evidence base, they may be limited with respect to the external validity of the RCTs from which they are derived.

We made no attempt to perform a complete search of the asthma literature, and probably RCTs have been reported elsewhere that are more inclusive of participants with current asthma than those examined. However, the RCTs selected here are those that provide the evidence for the treatment recommendations of the GINA guidelines, and hence have a direct effect on asthma management worldwide. Using actual RCTs rather than a hypothetical “typical” RCT has allowed a realistic estimate of the degree of selectivity of existing asthma trials. It also allows comparison of the degree of selectivity between different drugs for asthma or drug indications—for example, there was only one qualifying RCT comparing theophylline with another agent, for which none of the 179 participants with current asthma in our study were eligible.

**Table 2** Characteristics of included randomised controlled trials

<table>
<thead>
<tr>
<th>Reference</th>
<th>Year</th>
<th>Participants screened (n)</th>
<th>Participants randomised (n)</th>
<th>Age range (years)</th>
<th>Interventions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Greening et al ²</td>
<td>1994</td>
<td>Not stated</td>
<td>429</td>
<td>&gt;18</td>
<td>Salmeterol/BDP higher-dose BDP</td>
</tr>
<tr>
<td>Woolcock et al ²⁴</td>
<td>1996</td>
<td>990</td>
<td>738</td>
<td>&gt;17</td>
<td>Salmeterol/BDP (2 different salmeterol doses) v higher-dose BDP</td>
</tr>
<tr>
<td>Pauwels et al ²⁵</td>
<td>1997</td>
<td>1114</td>
<td>852</td>
<td>18–70</td>
<td>Formoterol/budesonide v budesonide (2 different budesonide doses in each group)</td>
</tr>
<tr>
<td>Busse et al ²⁶</td>
<td>1998</td>
<td>Not stated</td>
<td>473</td>
<td>18–70</td>
<td>Budesonide (4 different doses) v placebo</td>
</tr>
<tr>
<td>Reed et al ²⁶</td>
<td>1998</td>
<td>Not stated</td>
<td>747</td>
<td>6–65</td>
<td>BDP v theophylline</td>
</tr>
<tr>
<td>Wenzel et al ²⁶</td>
<td>1998</td>
<td>Not stated</td>
<td>539</td>
<td>&gt;12</td>
<td>Salmeterol v albuterter</td>
</tr>
<tr>
<td>Lavoielette et al ²⁶</td>
<td>1999</td>
<td>Not stated</td>
<td>642</td>
<td>&gt;15</td>
<td>Montelukast/BDP v montelukast/BDP v placebo</td>
</tr>
<tr>
<td>Bleeker et al ²⁶</td>
<td>2000</td>
<td>592</td>
<td>451</td>
<td>&gt;15</td>
<td>Fluticasone v zafirlukast</td>
</tr>
<tr>
<td>Nelson et al ²⁶</td>
<td>2000</td>
<td>Not stated</td>
<td>447</td>
<td>&gt;15</td>
<td>Salmeterol/Fluticasone v montelukast/Fluticasone</td>
</tr>
<tr>
<td>Busse et al ²⁶</td>
<td>2001</td>
<td>Not stated</td>
<td>525</td>
<td>12–75</td>
<td>Omalizumab v placebo</td>
</tr>
<tr>
<td>Fish et al ²⁶</td>
<td>2001</td>
<td>Not stated</td>
<td>948</td>
<td>&gt;15</td>
<td>Formoterol v montelukast</td>
</tr>
<tr>
<td>O’Byrne et al ²⁶</td>
<td>2001</td>
<td>2525</td>
<td>1970</td>
<td>&gt;12</td>
<td>Formoterol/budesonide v budesonide v placebo and formoterol/budesonide v budesonide (2 different budesonide doses in each group)</td>
</tr>
<tr>
<td>Lalloo et al ²⁶</td>
<td>2003</td>
<td>494</td>
<td>467</td>
<td>&gt;15</td>
<td>Formoterol/budesonide v higher-dose budesonide</td>
</tr>
<tr>
<td>Pauwels et al ²⁶</td>
<td>2003</td>
<td>Not stated</td>
<td>7241</td>
<td>5–66</td>
<td>Budesonide v placebo</td>
</tr>
<tr>
<td>Price et al ²⁶</td>
<td>2003</td>
<td>1192</td>
<td>889</td>
<td>15–75</td>
<td>Montelukast/budesonide v higher-dose budesonide</td>
</tr>
<tr>
<td>Ringdal et al ²⁶</td>
<td>2003</td>
<td>1168</td>
<td>806</td>
<td>&gt;15</td>
<td>Salmeterol/Fluticasone v montelukast/Fluticasone</td>
</tr>
<tr>
<td>Vaquerizo et al ²⁶</td>
<td>2003</td>
<td>846</td>
<td>639</td>
<td>18–70</td>
<td>Montelukast/budesonide v budesonide</td>
</tr>
</tbody>
</table>

BDP, beclomethasone dipropionate.
We were not always able to determine from our survey data whether a subject with asthma met a particular criterion, as many trials used criteria for exacerbations, symptom scores and measures of drug use that we were not able to duplicate. Where this occurred, participants were deemed to remain eligible by the criterion we could not assess. Hence, our estimates of the proportion of participants with asthma eligible for a given trial are maximum values, and the true degree of selectivity of these RCTs is probably greater than we have shown.

Our definition of asthma did not exclude those with concomitant COPD, defined as a post-bronchodilator FEV1/forced vital capacity <0.7. About one third of the participants with asthma met these criteria, most of whom did not have a significant smoking history. As a result, it is probable that this group with concomitant COPD was predominantly made up of participants with asthma, who had developed an irreversible component to their airway obstruction. Importantly, they represented a group of participants that had mostly received a doctor’s diagnosis of asthma and had been prescribed asthma treatment based on trials that largely excluded them.

The most common reason that participants with current asthma were not able to meet RCT eligibility criteria was the need to show bronchodilator reversibility. Application of this criterion resulted in only about a quarter of participants with current asthma in our study being eligible for the RCTs included in the study. This proportion was greater than that observed from a similar study from Australia, in which only 7–18% of participants with current asthma showed bronchodilator reversibility depending on the criteria used. These observations are likely to be due to the widespread use of inhaled corticosteroid treatment, resulting in good asthma control and associated reduction in lung function variability in asthma. The use of bronchodilator reversibility criteria in asthma RCTs may be justified on the basis that it provides the greatest opportunity to determine the maximum efficacy of a therapeutic agent. It also identifies individuals who do not have optimal asthma control, and as a result could be considered suitable for the addition of another therapeutic agent. This has led to novel study designs in which participants reduce or withdraw their inhaled corticosteroid treatment to demonstrate unstable asthma and its associated lung function variability.

Other eligibility criteria, such as the requirement that participants be non-smokers, also tend to produce a more homogeneous study population. This has the advantage of reducing the likelihood of a participant having concomitant COPD and limiting the number of variables apart from the RCT intervention, again maximising the likelihood of demonstrating a therapeutic effect specific to asthma. However, the criterion that participants be non-smokers resulted in the exclusion of 30% of the participants with current asthma in our study, representing a major group in which drug efficacy would not have been assessed. The importance of this limitation is evident from the experience with inhaled corticosteroids in asthma, where the studies showing a reduced efficacy of inhaled corticosteroids in smokers were not undertaken until >25 years after the introduction of these agents.

In our study group, the prevalence of wheezing and asthma diagnosed by a doctor was relatively high, with rates of 28.8% and 23.1%, respectively. These findings reflect the high prevalence of asthma in the New Zealand population, consistent with previous surveys that have reported rates of wheezing between 26% and 30%. Owing to the high survey response rate and the similarity between survey responders and the study group that undertook the investigative procedures, any effect of non-response bias is likely to be small.

In summary, we conclude that the degree to which the results from asthma RCTs apply to individual patients cannot be assessed directly and the clinician cannot assume that their patient will respond to a drug in the same way as trial participants. As a result, clinicians should consider that the treatment recommendations of major asthma guidelines may be limited with respect to the external validity of the RCTs on which they are based. This does not mean that the results of these RCTs are not generalisable to the wider community of individuals with asthma, but rather that the degree of generalisability is uncertain. We encourage the inclusion of a wider range of participants with asthma in future clinical trials.

### Authors’ affiliations

Justin Travers, Suzanne Marsh, Mathew Williams, Brent Caldwell, Philippa Shirtcliffe, Sarah Aldington, Medical Research Institute of New Zealand, Wellington, New Zealand

Mark Weatherall, Wellington School of Medicine and Health Sciences, Wellington, New Zealand

Richard Beasley, University of Southampton, Southampton, UK

Funding: The Wellington Respiratory Survey was funded by GlaxoSmithKline UK.

Competing interests: Richard Beasley is a member of the GINA Assembly.

### REFERENCES


BNF for Children 2006, second annual edition

In a single resource:
- guidance on drug management of common childhood conditions
- hands-on information on prescribing, monitoring and administering medicines to children
- comprehensive guidance covering neonates to adolescents

For more information please go to bnf.org
External validity of randomised controlled trials in asthma: to whom do the results of the trials apply?

Justin Travers, Suzanne Marsh, Mathew Williams, Mark Weatherall, Brent Caldwell, Philippa Shirtcliffe, Sarah Aldington and Richard Beasley

Thorax 2007 62: 219-223 originally published online November 14, 2006
doi: 10.1136/thx.2006.066837