Distinguishing between “no evidence of effect” and “evidence of no effect” in randomised controlled trials and other comparisons

If a court fails to convict a defendant because of incomplete evidence, does that establish his innocence beyond doubt? Not necessarily. Indeed in Scotland, if sufficient uncertainty remains, the court can give a verdict of “not proven” instead of “not guilty”. If a randomised controlled trial (RCT) fails to show a significant difference between the treatment and the control group, does that prove that the treatment has no useful clinical effect? Again, not necessarily. The treatment may work, but the trial may have been unable to prove it. Despite this, many such “negative” trials, including many published in this journal, may wrongly be taken as evidence that the treatment is not clinically useful.

For example, in an RCT of women at risk of preterm delivery that was not published as a full report, respiratory distress syndrome (RDS) occurred in three of 23 babies born to the treated group and three of 22 babies born to the untreated group. The difference is not significant (2p > 0.9). Had this been the first and only study of this treatment, many people might have decided that it was not effective and thus lost interest. In fact, reviews of this, and at least 14 other trials, eventually showed that the treatment—antenatal steroids—is highly effective because it reduced RDS and neonatal mortality in over 3500 preterm infants by about half. Note that the results of the single trial were quite consistent with this finding. The correct conclusion from that single trial is not that antenatal steroids do not work, but that the trial lacked sufficient power to detect anything but the most spectacular treatment effect. About half of all RCTs reported in Archives of Disease in Childhood between 1982 and 1996 recruited fewer than 40 children in total. Trials as small as this lack the power to detect moderate treatment effects and carry a significant risk of false negative results.

This is easier to see if trial data are presented with a point estimate of the effect, such as a relative risk or an odds ratio, and a measure of precision, such as a confidence interval (CI). If a treatment truly has no effect, the probability of a poor outcome should be the same for treated and untreated patients, so the relative risk and the odds ratio will each tend to be about 1. In the example just cited, the odds ratio for RDS in no RDS between treated and untreated groups is 0.95 (3/20 divided by 3/19), and the 95% CI around it ranges between 0.17 (a reduction of 83%) and 5.21 (an increase of 421%). So, although the odds ratio is close to 1, this particular trial rules out neither a substantially beneficial nor a substantially harmful effect because the CI is wide. An overview of all 15 trials gives an odds ratio for the effect of antenatal steroids on RDS of 0.53, with a much narrower 95% CI (0.44 to 0.63). In other words, it suggests that treatment with antenatal steroids is likely to reduce the odds of RDS by between 37% and 56%, an unequivocally substantial benefit, which is highly significant.

When should readers conclude that a treatment really is not clinically useful? Again, a CI is helpful, and surprisingly large numbers may be needed. In the fourth international study of infarct survival (ISIS-4), 58 050 patients with suspected myocardial infarction were randomly allocated to intravenous magnesium sulphate or placebo. There were 2216 deaths and 26 795 survivors in the treated group and 2103 deaths and 26 936 survivors in the placebo group, a difference that gives an odds ratio for increased mortality with magnesium of 1.06, with a 95% CI of 1.00 to 1.13 (2p = 0.07). In other words, magnesium, at least as it was given in this particular study, was not effective because it was unlikely to reduce mortality (and may even have increased it by up to 13%). Similarly, readers can only reliably conclude that two active treatments are equivalent—or that any difference between them is too small to be clinically important—when the sample is large enough.

How can researchers design RCTs powerful enough to show that no clinically important differences exist between treatment and placebo or between two active treatments? This requires prior estimation of appropriate sample sizes, which may require consultation with a statistician, but can easily be done for dichotomous outcomes (for example, survival or death) using software such as Epi Info. This software package allows calculation of relative risks, odds ratios, and 95% CI, and can be downloaded free of charge from the internet (http://www.soton.ac.uk/~medstats/epiinfo/). Calculating sample sizes when the outcome is a continuous variable (for example, blood pressure or length of stay) is more complicated and will almost certainly require consultation with a statistician. It may be added that the “null hypothesis”, to the effect that a treatment difference is exactly equal to 0 or a relative risk or an odds ratio exactly equal to 1, is often neither plausible nor interesting. Far more important is the question whether the size of the treatment effect is large enough to be of clinical interest, or small enough to be ignored. A conventional significance test (p value) cannot provide this information; only a range that covers the true value of the treatment difference with known confidence can do so.

Many investigators report in their tables of results two columns of means or percentages for the control and treated arms of the trial. In the former case, standard deviations, standard errors, or confidence limits for each column are commonly included. In fact, the quantities of interest to the reader are the differences between the two columns (or odds ratios for percentages), and these should always be shown with their standard errors or confidence limits. This is especially important when the data involve pairing or matching of treated and control subjects, as in crossover studies, because then the precision of the difference cannot be derived from the individual standard deviations.

The presentation of trial results has important implications for readers, authors, editors, referees, and patients. Wrongly discounting treatments as...
Isolated cough: probably not asthma

Children with cough, unassociated with wheeze or other evidence of systemic disease, are commonly seen in paediatric practice. In the distant past, the association between cough and wheeze was underrecognised but in recent years, as highlighted by McKenzie, children with cough alone have increasingly been diagnosed as having asthma.1

Problems relating to clinical studies of cough

Studies that include cough as an outcome need to be interpreted in light of inherent methodological problems. First, although the history of cough is a commonly used variable in epidemiology,2 the reporting of cough has poor repeatability. The chance corrected agreement (Cohen’s $\kappa$) between two occasions when cough was reported, is acceptably low (0.14–0.19).4 In contrast, wheeze related questions have excellent $\kappa$ values of 0.76–0.88.4 Second, cough, especially nocturnal cough, is unreliable reported compared with objective measures.5 6 In adults, the reporting and scoring of cough is dependent on psychological factors. Third, although cough scores using diary cards are widely used as an outcome measure, until recently no validated subjective cough scoring system was available.7 Compared to objective measurements, diary cards have been shown to be inaccurate for nocturnal cough8 and for metered dose inhaler use. Moreover, diary cards in studies can be parent completed9 or child completed, and these are not interchangeable.10 The accuracy of findings based entirely on diary cards is questionable.

Finally, a difficulty in the evaluation of the benefits of a therapeutic trial for cough is that cough usually resolves spontaneously (period effect). The widely quoted studies on “cough variant asthma (CVA)” were based on anecdotal evidence or were not placebo controlled.10 11 In trials that included a placebo arm, cough medications were no more effective than placebo.12 13 Guidelines that recommend a therapeutic trial of asthma medications for cough (and thus diagnose CVA) often fail to alert clinicians to this point.14

Traditional clinical model of cough and asthma

Based on the simple observations that some children with wheeze also cough, and that bronchoprovocation agents often stimulate both bronchoconstriction and cough in susceptible individuals, have led to the assumption that cough is part and parcel of the same mechanism as bronchoprovocation—the traditional clinical model. It has partly contributed to the establishment of common epidemiological and clinical practices.

In epidemiology studies, cough as a sole symptom has been used by some as synonymous with the diagnosis of asthma. Children with cough alone were categorised as “asthmatic”.2 15 16 Furthermore, the use of the symptom of cough alone as representative of asthma has partly contributed to the increased prevalence of asthma.17 18 In a paper that provided a breakdown of the method of how asthma was diagnosed, cough alone accounted for the largest rise in asthma prevalence.18 Between 1990 and 1993, the change in the cumulative prevalence of wheeze was 0.6% and that of troublesome cough was 3%.19

Kelly and colleagues showed that doctors were twice as likely to diagnose asthma on the basis of cough alone in 1990 compared with 1991 (22.6% and 10%, respectively).18 Of the children with no symptoms in the previous 12 months, 52% had a cough at some time that had been diagnosed as asthma.19 Although the prevalence of respiratory symptoms


documentation will deprive patients of better care. Wrongly accepting treatments as effective exposes patients to needless risk and wastes resources. We can all help to address these problems by expecting, and routinely including, CI or other measures of the precision of estimates of outcome in trial summaries and reports, and stating whether and how the sample size was calculated in advance.10 These measures have been recommended in the CONSORT statement,11 which Archives of Disease in Childhood has endorsed (see editors’ note in reference 6). We can also design and support larger trials with the power to detect realistically moderate, rather than over optimistically large, effects of treatment.6 12 Increasingly, such trials will require multi-centre collaboration and should be simple so that busy centres can contribute without taking on too great a burden of extra work.

We thank Richard Peto and the anonymous referee for helpful comments. The Perinatal Epidemiology Group is part of the Medical Research Council Health Services Research Collaboration.

Isolated cough: probably not asthma

Children with cough, unassociated with wheeze or other evidence of systemic disease, are commonly seen in paediatric practice. In the distant past, the association between cough and wheeze was underrecognised but in recent years, as highlighted by McKenzie, children with cough alone have increasingly been diagnosed as having asthma.1

Problems relating to clinical studies of cough

Studies that include cough as an outcome need to be interpreted in light of inherent methodological problems. First, although the history of cough is a commonly used variable in epidemiology,2 the reporting of cough has poor repeatability. The chance corrected agreement (Cohen’s $\kappa$) between two occasions when cough was reported, is acceptably low (0.14–0.19).4 In contrast, wheeze related questions have excellent $\kappa$ values of 0.76–0.88.4 Second, cough, especially nocturnal cough, is unreliable reported compared with objective measures.5 6 In adults, the reporting and scoring of cough is dependent on psychological factors. Third, although cough scores using diary cards are widely used as an outcome measure, until recently no validated subjective cough scoring system was available.7 Compared to objective measurements, diary cards have been shown to be inaccurate for nocturnal cough8 and for metered dose inhaler use. Moreover, diary cards in studies can be parent completed9 or child completed, and these are not interchangeable.10 The accuracy of findings based entirely on diary cards is questionable.

Finally, a difficulty in the evaluation of the benefits of a therapeutic trial for cough is that cough usually resolves spontaneously (period effect). The widely quoted studies on “cough variant asthma (CVA)” were based on anecdotal evidence or were not placebo controlled.10 11 In trials that included a placebo arm, cough medications were no more effective than placebo.12 13 Guidelines that recommend a therapeutic trial of asthma medications for cough (and thus diagnose CVA) often fail to alert clinicians to this point.14

Traditional clinical model of cough and asthma

Based on the simple observations that some children with wheeze also cough, and that bronchoprovocation agents often stimulate both bronchoconstriction and cough in susceptible individuals, have led to the assumption that cough is part and parcel of the same mechanism as bronchoprovocation—the traditional clinical model. It has partly contributed to the establishment of common epidemiological and clinical practices.

In epidemiology studies, cough as a sole symptom has been used by some as synonymous with the diagnosis of asthma. Children with cough alone were categorised as “asthmatic”.2 15 16 Furthermore, the use of the symptom of cough alone as representative of asthma has partly contributed to the increased prevalence of asthma.17 18 In a paper that provided a breakdown of the method of how asthma was diagnosed, cough alone accounted for the largest rise in asthma prevalence.18 Between 1990 and 1993, the change in the cumulative prevalence of wheeze was 0.6% and that of troublesome cough was 3%.19

Kelly and colleagues showed that doctors were twice as likely to diagnose asthma on the basis of cough alone in 1990 compared with 1991 (22.6% and 10%, respectively).18 Of the children with no symptoms in the previous 12 months, 52% had a cough at some time that had been diagnosed as asthma.19 Although the prevalence of respiratory symptoms...
was unchanged, the number of children who were diagnosed with asthma significantly increased.

Clinically, the advocacy of the diagnosis and treatment of the symptom of cough alone as asthma is also based on the traditional clinical model. Several groups have advocated the use of inhaled or oral steroids in children with cough alone. In one protocol children with cough were presumed to have asthma and some children were given repeated courses of oral steroids.

Cough is used as a marker of asthma stability in several asthma management guidelines and for assessing asthma severity. Guidelines for asthma management in children that state simply that cough or wheeze/dyspnoea are symptoms of asthma instability and therefore an indication to increase medications, are open to misinterpretation. In a functional severity score for asthma in children, a child with cough alone without other evidence of airway obstruction can be categorised as the second most severe group. In asthma guidelines the inclusion of the symptom of cough alone without qualification can lead to overtreatment. Although there is no doubt that children with asthma can present with cough and, in some, cough is a major symptom in asthma exacerbations, there is no evidence that the same mechanism is responsible for cough and bronchoconstriction. Indeed the converse is true—the alternative model.

Alternative model
There is good laboratory and clinical evidence that the pathways for cough and bronchoconstriction are distinctly different. In the laboratory, bronchoconstriction and cough pathways can be separately inhibited. Medications (lidocaine, oral codeine) that inhibit cough have no effect on bronchoconstriction. Conversely, medications (cromoglycate, atropine) that inhibit the pharmacologically induced bronchoconstriction have no effect on the cough response in adults. Atropine and cromoglycate inhibit bronchoconstriction but have no effect on the cough response, while inhaled lidocaine abolishes cough but does not influence bronchoconstriction.

Fujimura and colleagues examined the effect of airway calibre changes on cough receptor sensitivity (CRS) and found that methacholine and procaterol affected airway calibre, however neither affected CRS in adults with airway hyperresponsiveness (AHR). Similarly in children with stable asthma no correlation was found between CRS and airway calibre.

Clinically, three groups have described a lack of correlation between cough severity and asthma severity. During the recovery phase of an acute asthma exacerbation, cough frequency objectively measured had no correlation with airway calibre measured by diurnal peak expiratory flow (PEF) or its variability. In adults with stable asthma, cough frequency measured over 24 hours did not relate to asthma severity defined by FEV₁ (forced expiratory volume in one second) or diurnal variation in PEF. Nocturnal cough measured objectively in non-hospitalised children with asthma had no correlation with AHR, PEF variability, degree of morning dip in PEF, mean overnight saturation or treatment of asthma.

In children with asthma, there was no relation between CRS and airway calibre (FEV₁) in both the acute and interval phase of asthma. In contrast, in the recovery phase of acute severe asthma, improvements in FEV₁ correlated with the reduction in sputum eosinophils and eosinophil cationic protein.

Thus, while the trigger for wheeze and tachypnoea may be similar to that for cough, the pathways are distinctly different. Treatment with increasing doses of medications used for asthma of cough alone without any evidence of airway obstruction is questionable. Why the cough pathway is so easily stimulated in some is not known. The development of heightened CRS may account for the tendency of some patients to have troublesome cough.

Cough receptor sensitivity and cough
Cough can be mechanically or chemically stimulated endogenously (by airway secretions, mediators of inflammation) or exogenously by a variety of irritants (aspirated foreign materials, smoke, etc.), and tussive agents (capsaicin, citric acid). Cough sensitivity and pattern depends on the site and type of stimulation. The more proximal airways (larynx to trachea) are extremely sensitive to mechanical stimulation. The more distal airways are more chemosensitive and less mechanosensitive.

In both children and adults, heightened CRS occurs in those with a dry cough but not in those with productive cough. Prospective studies have shown an association between a decrease in symptoms and CRS of adults whose cough was secondary to angiotensin converting enzyme inhibitor use and upper respiratory tract infection.

In children, a temporal association between increased CRS and cough has also been demonstrated. Some children with recurrent cough had increased CRS and others had AHR during the coughing period, which returned to control values when cough free. Children with asthma who have cough as a major symptom of their acute episodes had a transient increase in CRS during the acute asthma exacerbation. In contrast, children who do not usually cough during their asthma exacerbations had no change in CRS during the acute and non-acute asthma periods. CRS of children with asthma during the non-acute phase were similar to controls.

The pathophysiology as to why increased CRS develops in some is not known. Possible factors include increased calcitonin gene related peptide nerve density, release of tachykinins, persistent inflammation, and qualitative or quantitative surfactant abnormality. Perhaps sensitivity of the cough receptors can be viewed in a similar way to sensitivity of the bronchoconstrictive pathway in susceptible people under certain circumstances.

Cough: is it asthma?
Several epidemiological studies also support McKenzie's article that highlighted the problem of overdiagnosis of asthma based on the symptom of cough alone. In a prospective community study, 56% of children with recurrent cough aged 4–7 years, later became asymptomatic: 37% reported continuing cough and 7.2% developed wheeze. The proportion of children in the group who subsequently developed wheeze was similar to the asymptomatic group who developed wheeze on follow up (10%).

Without intervention, recurrent respiratory symptoms in 8–10 year olds generally improve. In the absence of wheeze or shortness of breath, cough was not indicative of hidden or atypical asthma in most children with persistent nocturnal cough. Wright and colleagues in their prospective study of infants followed up to 11 years old, showed that recurrent cough early in life resolved in most children. These children with recurrent cough without wheeze did not have AHR or atopy, and significantly differed from those with classic asthma with or without cough.

The only randomised controlled trial on inhaled salbutamol and beclomethasone in children with cough alone showed that these medications were no different from placebo in reducing cough frequency measured objectively with a cough meter. The study also showed that AHR could not predict the child's response to these asthma medications.
Finally, the use of the term asthma needs to be interpreted in light of current controversy of wheezing phenotypes in childhood.21-23

Proposed model
Given the evidence outlined above, a model relating cough, CRS, and AHR in children is proposed. In a child with asthma, AHR may or may not be present.41 A trigger leads to an inflammatory response of the airway characteristic of asthma, which is clinically manifested by wheeze and dyspnoea. In some children, the cough pathway is also stimulated and these children have increased CRS during an exacerbation of their acute asthma.*

Some children without previously diagnosed asthma may present with cough. If no evidence of airway obstruction is present, most children with cough alone do not have asthma but have increased CRS.36 A small group of children have AHR but these children are no more likely to respond to inhaled salbutamol or corticosteroids than children without AHR.35 In the human airway, cough and bronchoconstriction pathways exist and both or either can be stimulated alone or simultaneously if sufficient and appropriate challenge agents are given. In children with recurrent cough who do not have evidence of airway obstruction, while both pathways are stimulated by the same trigger (such as viral infection), one pathway may be more dominant. It is not known if asymptomatic children with AHR are more likely to have hypersensitive cough receptors. A large prospective community study is required to answer this question.

Conclusions
The appropriateness of the inclusion of the symptom of cough alone as a marker of asthma severity should be questioned. Laboratory and clinical evidence indicate that cough and bronchoconstriction pathways are distinctly different in the human airway and that indices of cough (cough frequency, CRS) have no relation to airway calibre. There is no doubt that children with asthma can present with cough, as wheeze can be underrecognised. However, recent epidemiological and clinical studies have shown that cough alone is a poor marker of asthma and indeed most children with cough do not have asthma.3,15 40 A short trial of asthma treatment may be indicated in some children, but when these are used it is important to stop treatment if it is not beneficial and not to increase the dose. If the cough is related to asthma, earlier uncontrolled trials that used theophylline suggest that the cough should relent within a week.62 26 There is no objective scientific evidence to use high or prolonged doses of inhaled corticosteroids in children with cough alone. Children with cough alone are more likely to have a temporal increase in their CRS rather than asthma.

Department of Respiratory Medicine, Mater Misericordiae Children’s Hospital, South Brisbane, Queensland 4101, Australia

ANNE B CHANG

References