Background
In the United States, cough is the most common complaint for which patients seek medical attention and is the second most common reason for a general medical examination, accounting for more than 26 million office visits annually. Cough often results from an acute, self-limited, viral upper respiratory tract infection; however, there are multiple causes of cough beyond this, including both respiratory tract and nonrespiratory tract-related etiologies. Cough that lasts more than 4 weeks in children younger than 14 years of age or more than 8 weeks in adolescents and adults 14 years of age and older is considered to be chronic by the American College of Chest Physicians (ACCP). Cough serves a potentially beneficial purpose by clearing the airways of excessive mucus, irritants, or abnormal substances such as edema fluid or pus. But while cough may serve a useful function, it can also lead to a variety of problems, including exhaustion (57%), feeling self-conscious (55%), insomnia (45%), changes in lifestyle (45%), musculoskeletal pain (45%), hoarseness (43%), excessive perspiration (42%), and urinary incontinence (39%). These problems are more likely to be prominent in the setting of chronic versus acute cough. As a consequence, chronic cough is responsible for up to 38 percent of pulmonary outpatient visits.
To effectively assess cough and monitor response to treatment, it is essential to have valid measurement tools. Currently there are many different tools used to assess cough frequency and severity, including quality-of-life questionnaires, visual analog scales, electronic recordings, and human counts. It is important to determine whether the tools currently in use accurately assess cough and response to treatment. While no universally accepted gold standard exists for comparison, data regarding the validity, consistency, reliability, and responsiveness of these tools are needed. The purpose of this review is to evaluate the effectiveness of instruments to evaluate cough and the comparative effectiveness of treatments for the symptom of cough in patients with either unexplained or refractory chronic cough.

In patients with no identifiable cause of cough (unexplained or idiopathic) or no response to specific treatment (unresponsive, refractory, or intractable), chronic cough poses a particularly challenging problem. The differential diagnosis for chronic cough has a different list of etiologies compared with acute cough. Treatment for chronic cough contrasts with acute cough in that acute cough treatment may focus on curing the underlying etiology (e.g., bacterial bronchitis or pneumonia) or suppressing symptoms for the short period of time needed for the etiology to resolve spontaneously (e.g., viral etiologies). Cough becomes chronic if it persists, often due to an underlying etiology that is difficult to diagnose or treat. Therefore, treatments for cough may have differential effectiveness depending on whether the cough is acute versus chronic. Side effects of medication may also become more salient in the setting of chronic cough given that treatment duration is longer, allowing more opportunity for side effects to occur. Chronic cough also differs from acute cough in that quality of life may be affected more severely and in different ways than with acute cough. Recent studies from the United Kingdom, United States, and Japan evaluating patients with chronic cough have estimated that up to 46 percent of patients have idiopathic cough despite a thorough diagnostic investigation.7

The management of nonspecific acute or chronic cough in young children can be especially difficult because of the risks associated with pharmacotherapy. In 2008, manufacturers voluntarily removed over-the-counter infant (<2 years of age) cough and cold products (e.g., those containing ephedrine, pseudoephedrine, phenylephrine, diphenhydramine, brompheniramine, or chlorpheniramine) because of many reports of serious adverse events. Later that year, manufacturers relabeled cough and cold products to warn against use in children <4 years of age.8 This position is supported by the American Academy of Pediatrics.

The diagnosis and management of cough has been the subject of several guideline efforts, two aimed at assessment of cough in adults,9,10 and one focused on children.11 Guidelines from ACCP, last updated in 2006, are the most comprehensive resource and will be the subject of a future update.10 Identifying the underlying etiology is the most important step in the successful management of chronic cough.10 If, however, no cause can be identified, or if treatment of the underlying etiology fails to resolve the cough, then the cough may be treated symptomatically. In the majority of cases, symptomatic treatment consists of antitussive therapy to decrease cough frequency and severity. Antitussive treatments vary in mechanism of action. Nonspecific antitussives such as dextromethorphan and codeine appear to act in the brain stem to reduce the cough reflex. Other nonspecific antitussives, such as benzonatate, act to anesthetize respiratory passages and thus reduce the stimulus to cough. Other agents aim to decrease the volume of respiratory tract secretions and thus the need to cough. These latter antitussive agents are also used to treat certain common underlying etiologies and include antihistamines, corticosteroids, antibiotics, decongestants, and mast cell stabilizers. Nonpharmacological antitussives are few but may include, for example, honey. Recently, speech therapy interventions have been used to treat chronic cough in patients suspected of upper airway hypersensitivity.12

In a limited number of situations where cough provides a useful function (such as in bronchiectasis, pneumonia, or atelectasis), protussive therapy may be used in an attempt to increase cough effectiveness without increasing its frequency. Protussive treatments aim to change the characteristics of mucus in such a way that it can be cleared more effectively by mucociliary action or cough. Such effective clearing can subsequently lessen the severity and frequency of a patient’s cough. Protussive pharmacological agents include expectorants, mucolytics, and mucus-modifying agents. Examples of these include guaifenesin, hypertonic saline, and acetyleysteine. In addition, physical maneuvers such as chest physical therapy, flutter valves, or pneumatic jackets may be used, especially in patients with respiratory muscle weakness.

**Scope and Key Questions**

This comparative effectiveness review (CER) was funded by the Agency for Healthcare Research and Quality (AHRQ) and is designed to evaluate the comparative...
effectiveness of measurement tools for assessing cough and of symptomatic treatments for chronic cough.

With input from our Key Informants, we constructed Key Questions (KQs) using the general approach of specifying the population of interest, the interventions, comparators, outcomes, timing of outcomes, and settings (PICOTS). The KQs considered in this review were:

**KQ 1:** In adults and adolescents (≥14 years of age) and children (<14 years of age), what is the comparative diagnostic accuracy, therapeutic efficacy, and patient outcome efficacy of instruments used to assess cough?

**KQ 2:** In adults and adolescents (≥14 years of age) and children (<14 years of age), what are the comparative safety and effectiveness of nonspecific (or symptomatic) therapies to treat patients with chronic cough?

a. In patients with unexplained chronic cough

b. In patients with refractory cough with a known underlying etiology

Figure A depicts the KQs within the context of the PICOTS.

**Methods**

The methods for this CER follow those suggested in the AHRQ Methods Guide for Effectiveness and Comparative Effectiveness Reviews (hereafter referred to as the Methods Guide) and Methods Guide for Medical Test Reviews (hereafter referred to as the Medical Test Guide).

**Input From Stakeholders**

During the topic refinement stage, we solicited input from Key Informants representing clinicians (adult and pediatric

---

**Figure A. Analytic framework**

- **Patients With Cough**
  - Adults and adolescents (≥14 years of age)
  - Children (<14 years of age)

- **Underlying Etiologies**
  - Asthma
  - GERD
  - Upper airway cough syndrome
  - Tobacco use
  - ACE inhibitor user
  - Pulmonary infection
  - Bronchitis
  - Cystic fibrosis
  - Others
  - Unexplained

- **Patients with unexplained or refractory chronic cough**
  - Antitussive therapies
  - Protussive therapies
  - Nonpharmacologic therapies

- **Adverse Events**
  - Sleep disturbance
  - Allergic reaction
  - Drowsiness
  - Headache
  - Chest pain
  - Dizziness
  - Rash
  - Others

- **Final Outcomes**
  - Cough symptoms
  - Cough severity
  - Cough frequency
  - Complications related to cough
  - Functional status
  - HRQOL
  - Health care utilization and costs

- **Diagnostic Accuracy**
  - Sensitivity, specificity, PPV, NPV, reliability, responsiveness, feasibility, validity

- **Therapeutic Efficacy**
  - Change in clinical practice
  - Impact on patient and provider decisionmaking

- **Patient Outcome Efficacy**
  - Acceptability
  - HRQOL
  - Chest pain, depression, anxiety

---

ACE = angiotensin-converting enzyme; GERD = gastroesophageal reflux disease; HRQOL = health-related quality of life; KQ = Key Question; NPV = negative predictive value; PPV = positive predictive value
pulmonology, otolaryngology, school nursing, respiratory medicine, primary care), patients, scientific experts, and payers, to help define the KQs. The KQs were then posted for public comment in September 2011 for 4 weeks, and the comments received were considered in the development of the research protocol. We next convened the Technical Expert Panel (TEP) comprising clinical, content, and methodological experts to provide input in defining populations, interventions, comparisons, and outcomes, and in identifying particular studies or databases to search. The Key Informants and members of the TEP were required to disclose any financial conflicts of interest greater than $10,000 and any other relevant business or professional conflicts. Any potential conflicts of interest were balanced or mitigated. Neither Key Informants nor members of the TEP performed analysis of any kind, nor did any of them contribute to the writing of this report.

**Literature Search Strategy**

To identify the relevant published literature, we searched PubMed®, Embase®, and the Cochrane Database of Systematic Reviews (CDSR; last search date for all three sources June 4, 2012). Where possible, we used existing validated search filters (such as the Clinical Queries Filters in PubMed). An experienced search librarian guided all searches. We supplemented the electronic searches with a manual search of references from a set of key primary and systematic review articles. All citations were imported into an electronic database (EndNote® X4; Thomson Reuters, Philadelphia, PA).

We used several approaches to identify relevant grey literature, including a request for scientific information packets submitted to drug and device manufacturers and a search of U.S. Food and Drug Administration (FDA) device registration studies and new drug applications. We also searched study registries and conference abstracts for relevant articles from completed studies. Grey literature databases searched included ClinicalTrials.gov (July 18, 2012); the World Health Organization (WHO) International Clinical Trials Registry Platform Search Portal (July 18, 2012); and ProQuest COS Conference Papers Index (January 18, 2012).

**Inclusion and Exclusion Criteria**

Criteria used to screen articles for inclusion/exclusion at both the title-and-abstract and full-text screening stages are detailed in Table 2 of the main report. For KQ 1, the search focused on English-language evaluative studies that compared qualitative and/or quantitative instruments used to assess cough in patients (inpatients or outpatients) with cough of any duration and considering the following outcomes: diagnostic accuracy (e.g., sensitivity, validity, reliability, among others); therapeutic efficacy (e.g., impact on patient or provider decisionmaking); and patient outcome efficacy (e.g., acceptability, quality of life). For KQ 2, the search focused on English-language, prospective (randomized controlled trial [RCT] or cohort studies), comparative assessments of pharmacological and nonpharmacological therapies aimed at treating the symptom of cough in patients with chronic cough, in particular, patients with unexplained chronic cough or refractory cough of known etiology. We accepted as chronic any cough described as such, or that exceeded 8 weeks in adults and adolescents or 4 weeks in children ≤14 years of age. Because determination of whether an individual’s chronic cough was truly unexplained or refractory was often difficult or impossible given available descriptions in the published article, we did not exclude articles based on diagnostic evaluation or empiric therapeutic trials, but rather described such information in an attempt to infer to what extent study populations could be considered unexplained or refractory according to current criteria. Articles were excluded if the therapy was directed at an underlying etiology rather than the symptom of cough, if cough resulted from invasive respiratory tract instrumentation, or if the intervention tested was not available in the United States. The following outcomes were considered: cough symptoms and severity, complications related to coughing, functional status, health-related quality of life, health care utilization and costs, and adverse effects of therapy.

**Study Selection**

Using the prespecified inclusion and exclusion criteria, titles and abstracts were reviewed independently by two investigators for potential relevance to the KQs. Articles included by either reviewer underwent full-text screening. At the full-text review stage, paired researchers independently reviewed the articles and indicated a decision to “include” or “exclude” the article for data abstraction. When the two reviewers arrived at different decisions about whether to include or exclude an article, they reconciled the difference through review and discussion, or through a third-party arbitrator if needed. Full-text articles meeting our eligibility criteria were included for data abstraction. Relevant review articles, meta-analyses, and methods articles were flagged for manual searching of references and cross-referencing against the library of citations identified through electronic database searching. All screening decisions were made and
Data Extraction

The research team created data abstraction forms and evidence table templates for each KQ. Based on clinical and methodological expertise, a pair of investigators was assigned to abstract data from each eligible article. One investigator abstracted the data, and the second reviewed the completed abstraction form alongside the original article to check for accuracy and completeness. Disagreements were resolved by consensus or by obtaining a third reviewer’s opinion if consensus could not be reached.

We designed the data abstraction forms to collect the data required to evaluate the specified eligibility criteria for inclusion in this review, as well as demographic and other data needed for determining outcomes (intermediate, final, and adverse events outcomes). We gave particular attention to describing the details of the treatment, patient characteristics, and study design that were related to outcomes. In addition, we described comparators carefully, as treatment standards may have changed during the study period. The safety outcomes were framed to help identify adverse events from drug therapies and nonpharmacological therapies. Data necessary for assessing quality and applicability were also abstracted. Before the data abstraction form templates were used, they were pilot-tested with a sample of included articles and revised as necessary.

Quality Assessment of Individual Studies

We evaluated the quality of individual studies using the approach described in the Methods Guide.13 To assess quality, we used the strategy to (1) classify the study design, (2) apply predefined criteria for quality and critical appraisal, and (3) arrive at a summary judgment of the study’s quality. Criteria of interest for all studies included similarity of groups at baseline, extent to which outcomes were described, blinding of subjects and providers, blinded assessment of the outcome(s), intention-to-treat analysis, differential loss to followup between the compared groups or overall high loss to followup, and conflicts of interest. Criteria specific to RCTs included methods of randomization and allocation concealment. For observational studies, additional elements such as methods for selection of participants, measurement of interventions/exposures, addressing any design-specific issues, and controlling confounding were considered. We used the summary ratings of good, fair, or poor based on the study’s adherence to well-accepted standard methodologies and adequate reporting.

For studies of diagnostic tests (KQ 1), we used the QUality Assessment tool for Diagnostic Accuracy Studies (QUADAS)-215 to assess quality in four key domains: patient selection, index test(s), reference standard, and flow and timing. The questions in each domain are rated in terms of risk of bias and concerns regarding applicability, with associated signaling questions to help with these bias and applicability judgments.

Data Synthesis

We began our data synthesis by summarizing key features of the included studies for each KQ.

For KQ 1 we considered the three dimensions of (1) cough frequency, (2) cough severity (which might include quantity and characteristics of sputum, difficulty of expectoration, dyspnea, between cough sensations, or pain), and (3) cough-specific quality of life (QOL). We then sought to measure the validity, reliability, and responsiveness of various instruments used to assess each of these dimensions. For cough frequency, we evaluated validity by concurrence with measures of other constructs (e.g., cough severity, cough-specific QOL, tussigenic challenge (or cough reflex sensitivity), and exhaled nitrous oxide), and we assessed reliability using intermethod reliability (e.g., manual cough counts vs. electronic recording device cough counts) and test-retest reliability. Although we consider cough severity and cough-specific QOL to be separate dimensions of cough, most of the standardized questionnaires included in this report measured aspects of both of these dimensions. Therefore, for the purpose of this report, we considered instruments that measured both severity and QOL together to be "severity/QOL" instruments. Within this report, we did not identify any validated instruments that focused purely on cough severity. For these severity/QOL instruments, we evaluated validity by looking at concurrence with measures of other constructs including cough frequency, quality of life, and tussigenic challenge findings. We assessed reliability by test-retest reliability, as well as internal consistency. We evaluated responsiveness of both frequency and severity/QOL measures by reporting data on changes in these measures over time associated with treatment (or no treatment) of cough symptoms or the underlying etiology of cough.

For KQ 2, we determined the feasibility of completing a quantitative synthesis (i.e., meta-analysis). Feasibility depended on the volume of relevant literature, conceptual homogeneity of the studies, and completeness of the
reporting of results. We considered meta-analysis for comparisons where at least three studies reported the same outcome. We considered measures of cough frequency, regardless of the scale used, to be similar enough to combine using effect sizes (standardized mean differences); similarly, measures of cough severity that used different measurement scales were considered similar enough to combine using effect sizes.

When a meta-analysis was appropriate, we used random-effects models to quantitatively synthesize the available evidence using Comprehensive Meta-Analysis software (Version 2; Biostat, Englewood, NJ). We tested for heterogeneity using graphical displays and test statistics (Q and I^2 statistics). We present summary estimates, standard errors, and confidence intervals in our data synthesis.

We supplemented the meta-analysis of direct comparisons with a mixed treatment meta-analysis that incorporated data from placebo comparisons and head-to-head comparisons, including multi-armed trials (i.e., trials that included more than one comparison). The general strategy for analysis was to construct a random-effects model that was comparable to the standard random-effects models used in the meta-analysis of effect sizes. This model, which was fitted using SAS® PROC NLMIXED (2009; SAS Institute Inc., Cary, NC), estimated the effect sizes (relative to placebo) for each treatment.

**Strength of the Body of Evidence**

We rated the strength of evidence for each KQ and outcome using the general approach described in the Methods Guide. In brief, the approach requires assessment of four domains: risk of bias, consistency, directness, and precision. Additional domains were used when appropriate: coherence, dose-response association, impact of plausible residual confounders, strength of association (magnitude of effect), and publication bias. These domains were considered qualitatively, and a summary rating of “high,” “moderate,” or “low” strength of evidence was assigned after discussion by two reviewers. In some cases, high, moderate, or low ratings were impossible or imprudent to make, for example, when no evidence was available or when evidence on the outcome was too weak, sparse, or inconsistent to permit any conclusion to be drawn. In these situations, a grade of “insufficient” was assigned.

**Applicability**

We assessed applicability across our KQs using the method described in the Methods Guide. In brief, this method uses the PICOTS format as a way to organize information relevant to applicability. The most important issue with respect to applicability is whether the outcomes are different across studies that recruit different populations (e.g., age groups, exclusions for comorbidities) or use different methods to implement the interventions of interest; that is, important characteristics are those that affect baseline (control-group) rates of events, intervention-group rates of events, or both. We used checklists to guide the assessment of applicability. We used these data to evaluate the applicability to clinical practice, paying special attention to study eligibility criteria, demographic features of the enrolled population in comparison with the target population, characteristics of the intervention used in comparison with care models currently in use, and clinical relevance and timing of the outcome measures. We summarized issues of applicability qualitatively.

**Results**

Figure B depicts the flow of articles through the literature search and screening process. Searches of PubMed®, Embase®, and CDSR yielded 21,860 citations, 6,504 of which were duplicate citations. Manual searching identified 75 additional citations, for a total of 15,431 citations. After applying inclusion/exclusion criteria at the title-and-abstract level, 833 full-text articles were retrieved and screened. Of these, 718 were excluded at the full-text screening stage, leaving 115 articles for data abstraction. Overall, we included 121 studies represented by these 115 publications: 78 studies were relevant to KQ 1, 48 to KQ 2 (5 studies were relevant to both KQs). Studies were conducted in Europe (54%); the United States or Canada (23%); Australia or New Zealand (11%); Asia (8%); and other locations (8%). Nineteen studies in KQ 1 (23%) and 3 studies in KQ 2 (6%) included children. Forty-five studies (37%) were published before 2000. No additional information was found through our grey literature search.

**Key Question 1. Instruments Used To Assess Cough**

Key points from the Results chapter are:
- Electronic recording devices are accurate for assessing cough frequency, but they show variable correlation with instruments that measure other dimensions of cough.
- The Leicester Cough Questionnaire (LCQ) and the Cough-specific Quality of Life Questionnaire (CQLQ)
are the most widely studied cough-specific quality-of-life questionnaires in adult populations. Both have demonstrated validity and reliability, with emerging evidence available on responsiveness.

- There is moderate strength of evidence to support the validity and responsiveness of the Parent Cough-specific Quality of Life Questionnaire (PC-QOL) in assessing the severity/QOL of cough among children.

- Emerging data support the responsiveness of recording devices, cough-related questionnaires, and tussigenic challenge tests, but further research is needed to accurately estimate the minimally important difference (MID) of these assessment instruments.

- Although diaries and visual analog scales are based on face validity, assess a wide variety of different cough outcomes, and are widely used both in research and
practice, there is little data to validate their accuracy in assessing cough, and what data exist show inconsistent correlations with other cough measurement tools. These tools are usually simple and easy to use, but more data are needed to determine their reliability and validity in assessing cough frequency or severity/QOL.

• While all of the included studies evaluated aspects of the comparative diagnostic accuracy of the various cough measurement tools, none evaluated the comparative therapeutic efficacy or patient outcome efficacy of these tools.

Cough can be assessed along several dimensions, the most of important of which may be frequency, severity, and cough-specific QOL. Cough frequency is objective and relatively easy to measure but may not necessarily correlate with severity or cough-specific QOL, whereas cough severity and cough-specific QOL may be closely interrelated. Most of the standardized questionnaires included in this report measured aspects of both of these latter dimensions. Therefore, for the purpose of this report, we considered instruments that measured both severity and QOL together to be “severity/QOL” instruments. In this CER we evaluate the available data that support the validity and reliability of instruments to measure one of two dimensions of cough: (1) cough frequency; or (2) the severity/QOL impact of cough (including assessments of the impact of cough on sleep, work, general well-being, health-related quality of life, etc.). We also evaluate the available data that support these instruments’ ability to measure potentially meaningful clinical change over time (responsiveness).

To be eligible for inclusion in this report, a study had to either (1) compare a cough frequency or severity/QOL assessment instrument with one or more cough assessment, health-related quality of life, or clinical change instrument; or (2) report data on changes in the instrument score over time in response to treatment for cough or the underlying etiology of the cough. For the purposes of this report, we consider tussigenic challenge tests and exhaled nitric oxide tests as severity/QOL assessments.

A total of 78 studies met the inclusion criteria for this KQ. Of these, 67 (86%) were judged to have a low risk of bias and 11 (14%) were judged to have a high risk of bias. In most cases, the funding source was not reported or was unclear. Seven studies were RCTs, and the remaining 71 were observational studies. A total of 5,927 participants were included across studies; sample sizes of individual studies ranged from 1 to 671 subjects. Thirty-three studies (42%) enrolled patients with chronic cough of mixed, unknown, or unspecified etiology; 18 (23%) enrolled patients with acute cough or cough of unspecified duration, and 27 (35%) focused on specific clinical conditions such as chronic bronchitis, asthma, or lung cancer. Fifty-nine studies included adults and adolescents (≥14 years of age), 15 included only children (<14 years of age), and 4 included adults, adolescents, and children.

Table A summarizes the findings of our review and the strength of evidence for the available outcomes of validity, internal consistency, reliability, and responsiveness for the main instruments. Details about the specific components of these ratings (risk of bias, consistency, directness, and precision) are available in the main report. We did not identify any studies evaluating the comparative therapeutic efficacy or patient outcome efficacy of these tools; therefore, the current evidence base is insufficient for us to draw any conclusions about these outcomes.

**Key Question 2. Nonspecific Therapies for Chronic Cough**

Key points from the Results chapter are:

• A wide variety of pharmaceutical agents have been used to treat the symptom of chronic cough, including opioid, anesthetic, and nonopioid/nonanesthetic antitussives; expectorant and mucolytic protussives; antihistamines; antibiotics; inhaled corticosteroids; and inhaled anticholinergics.

• Patients with unexplained or refractory chronic cough are not well defined as a population in the evidence base, restricting the applicability of many studies.

• Of the agents reviewed, the opioid and certain nonopioid/nonanesthetic antitussives most frequently demonstrated efficacy for managing the symptom of chronic cough in adults.

• There were several important quality limitations in the literature, including (1) too few good-quality studies focusing on chronic cough; (2) relatively short durations of followup (3) a diversity of outcomes measured across studies, which limited between-study comparisons; and (4) when similar outcomes were assessed across studies, the instruments used were diverse and inconsistent, making comparison and interpretation difficult.

• Data on nonpharmacological therapies for chronic cough were sparse.
### Table A. Summary of strength of evidence (SOE) and effect estimate for KQ 1

<table>
<thead>
<tr>
<th>Instrument (Dimension[s] Assessed)</th>
<th>Validity (Correlation With Other Measures of Cough)</th>
<th>Reliability</th>
<th>Responsiveness</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Leicester Cough Questionnaire (LCQ) (Severity/QOL)</strong></td>
<td>Moderate SOE</td>
<td>High SOE</td>
<td>Moderate SOE</td>
</tr>
<tr>
<td>15 studies; 1,058 subjects</td>
<td>4 studies; 430 subjects</td>
<td>2 studies; 256 subjects</td>
<td>8 studies; 659 subjects</td>
</tr>
<tr>
<td>Range of $r = 0.26–0.93$</td>
<td>Range of $r = 0.77–0.93$</td>
<td>Range of $r = 0.86–0.92$</td>
<td>Range of ES = 0.84–19.5</td>
</tr>
<tr>
<td><strong>Cough-specific Quality of Life Questionnaire (CQLQ) and Adverse Cough Outcome Survey (ACOS)</strong></td>
<td>Moderate SOE</td>
<td>Insufficient SOE</td>
<td>Insufficient SOE</td>
</tr>
<tr>
<td>5 studies; 336 subjects</td>
<td>1 study; 184 subjects</td>
<td>1 study; 52 subjects</td>
<td>7 studies; 460 subjects</td>
</tr>
<tr>
<td>Range of $r = 0.24–0.56$</td>
<td>Range of $r = 0.63–0.92$</td>
<td>Range of $r = 0.75–0.93$</td>
<td>Range of ES = 0.32–0.41</td>
</tr>
<tr>
<td><strong>Parent Cough-specific Quality of Life questionnaire (PC-QOL) (Severity/QOL)</strong></td>
<td>Moderate SOE</td>
<td>Moderate SOE</td>
<td>Moderate SOE</td>
</tr>
<tr>
<td>4 studies; 593 subjects</td>
<td>3 studies; 247 subjects</td>
<td>3 studies; 247 subjects</td>
<td>3 studies; 247 subjects</td>
</tr>
<tr>
<td>Range of $r = 0.01–0.70$</td>
<td>Range of $r = 0.56–0.91$</td>
<td>Range of ES = 0.32–0.41</td>
<td>Range of ES = 0.32–0.41</td>
</tr>
<tr>
<td><strong>Electronic recording devices (Frequency)</strong></td>
<td>High SOE</td>
<td>NA</td>
<td>Insufficient SOE</td>
</tr>
<tr>
<td>17 studies; 546 subjects</td>
<td></td>
<td>1 study; 67 subjects</td>
<td>1 study; 21 subjects</td>
</tr>
<tr>
<td>Range of $r = 0.89–0.99$</td>
<td>Detected change with treatment</td>
<td></td>
<td>Sensitivity of 0.81–0.95 for detecting clinically important change</td>
</tr>
<tr>
<td><strong>Visual analog scales (Severity/QOL)</strong></td>
<td>Insufficient SOE</td>
<td>NA</td>
<td>Insufficient SOE</td>
</tr>
<tr>
<td>9 studies; 410 subjects</td>
<td></td>
<td>1 study</td>
<td>1 study</td>
</tr>
<tr>
<td>No summary measure</td>
<td></td>
<td>Sensitivity of 0.81–0.95</td>
<td></td>
</tr>
</tbody>
</table>

ACOS = Adverse Cough Outcome Survey; CQLQ = Cough-specific Quality of Life Questionnaire; ES = effect size; KQ = Key Question; LCQ = Leicester Cough Questionnaire; MID = minimal important difference; NA = not applicable; PC-QOL = Parent Cough-specific Quality-of-Life questionnaire; $r$ = correlation coefficient; SOE = strength of evidence

1. All strength of evidence ratings of “Insufficient” or “NA” (not applicable) are shaded.
2. The ACOS has been revised and replaced by the CQLQ.

- Studies evaluating management of unidentifed or refractory chronic cough in children are extremely limited.
- All preparations appeared to be well-tolerated, but side effects and adverse events were uncommonly reported; underreporting side effects and adverse events could limit the assessment of effectiveness of these drugs.

Sixty-seven comparisons from 48 studies evaluated therapies in patients with chronic cough and met our inclusion criteria. The 48 studies were described in 42 publications. Thirty-three of the 48 studies were parallel-group RCTs, and 12 were randomized crossover studies. The range of years of publication was 1953 to 2012; 76 percent of the articles were published before 2000. Only three studies were performed in children. A total of 2,923 participants were included across trials; sample sizes were relatively small, ranging from 8 to 214 participants. Duration of followup was relatively short in most studies, ranging from 1 hour to 115 days. Thirty-three (33) studies (69 percent) had a followup duration of 2 weeks or less. The majority of studies were rated fair in quality ($n=29$, or 60%); 11 studies were good in quality, and 8 were poor in quality. Fair- and poor-quality studies had the following limitations: limited description of study entry criteria, randomization, and patient population; incomplete followup; less valid statistical analyses (not intention-to-treat, post hoc subgroup analyses); and/or inadequate reporting of methods and findings.

A variety of agents were studied and could be broadly categorized into antitussives, protussives,
and nonantitussive/nonprotussive agents. Antitussives were further categorized as opiates, anesthetics, nonpharmacological, or “other” types. Protussives included expectorants, mucolytics, and nonpharmacological therapies. Nonantitussive/nonprotussive pharmacotherapies included antihistamines, antibiotics, anticholinergics, and bronchodilators. Figure C represents the various categories of agents and the comparisons among these agents represented in the included studies. The 48 studies represented 67 different comparisons within or between treatment classes and included studies of 59 individual agents. There were 39 comparisons (58%) with placebo. The most common class comparisons were between other antitussives and placebo (12 comparisons, 18%), followed by comparisons between antitussive opiates and placebo (11 comparisons, 16%) and comparisons between antitussive opiates and other antitussives (10 comparisons, 15%). Fourteen different class comparisons were evaluated by only one or two studies. Only two studies evaluated nonpharmacological interventions.

The heterogeneity of the included studies in terms of the interventions and comparators, combined with the lack of three or more studies reporting the same outcome where there were multiple comparisons, precluded us from performing meta-analyses on almost all outcomes. Even when similar outcomes were assessed across studies, the instruments used were diverse and inconsistent, making comparison and interpretation difficult. Therefore the evidence from head to head trials is insufficient to draw conclusions about relative benefit.

Figure C. Overview of intervention class comparisons
We were, however, able to evaluate the relative effects on cough severity for four classes of treatments for chronic cough: antitussive opiates, antitussive dextromethorphan, antitussive moguisteine, and protussive mucolytics. This analysis included 11 studies and 700 patients. Most of the studies compared the treatment with placebo, but one compared opiates with dextromethorphan and placebo. Because each study used a different measure of severity, we converted all results to effect sizes (standardized mean differences). Relative to placebo, the effect of dextromethorphan on cough severity was 0.54 (95% confidence interval [CI], 0.27 to 0.80; \( p=0.0008 \)), the effect of opiates was 0.63 (95% CI, 0.40 to 0.86; \( p<0.0001 \)), the effect of moguisteine was 0.62 (95% CI, 0.04 to 1.16, \( p=0.0366 \)), and the effect of mucolytics was 0.14 (95% CI -0.20 to 0.49; \( p=0.384 \)). The studies showed significant heterogeneity (\( p=0.0023 \)). The effects of dextromethorphan, moguisteine, and opiates compared with placebo on cough severity support a benefit of these therapies, but the evidence is insufficient to determine relative benefit among these therapies.

We performed a similar meta-analysis for cough frequency, including 7 studies and 396 patients. Relative to placebo, the effect of dextromethorphan on cough frequency was 0.40 (95% CI, 0.18 to 0.85; \( p=0.0248 \)), the effect of codeine was 0.57 (95% CI, 0.36 to 0.91; \( p=0.0260 \)), and the effect of moguisteine was 0.60 (95% CI, 0.31 to 1.17, \( p=0.1117 \)). Again, the studies showed significant heterogeneity (\( p=0.0231 \)). The effects of dextromethorphan and codeine compared with placebo on cough frequency support a benefit of these therapies, although the estimates are too imprecise to determine if one is superior to another. The effect of moguisteine was too imprecise to draw conclusions about its efficacy.

Tables B and C summarize the strength of evidence for the most commonly used classes of therapies and evaluated outcomes. Details about the specific components of these

### Table B. Summary of strength of evidence (SOE) and effect estimate for KQ 2—active treatment comparisons

<table>
<thead>
<tr>
<th>Treatment Comparison</th>
<th>Cough Severity</th>
<th>Cough Frequency</th>
<th>Adverse Effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Antitussive (anesthetic) vs. antitussive (opiate)</td>
<td>Insufficient SOE</td>
<td>1 study; 45 subjects</td>
<td>Insufficient SOE</td>
</tr>
<tr>
<td></td>
<td>Insufficient SOE</td>
<td>Imprecise results</td>
<td>Insufficient SOE</td>
</tr>
<tr>
<td>Antitussive (opiate) vs. antitussive (other)</td>
<td>Insufficient SOE</td>
<td>16 studies; 958 subjects</td>
<td>Insufficient SOE</td>
</tr>
<tr>
<td></td>
<td>Insufficient SOE</td>
<td>Opiates, dextromethorphan, and moguisteine had significant effect sizes vs. placebo in MTM (ranging from 0.54–0.63), but wide and overlapping CIs are too imprecise to (determine equivalence or noninferiority or) draw conclusions about relative effectiveness</td>
<td>Insufficient SOE</td>
</tr>
<tr>
<td>Protussive (mucolytic) vs. antitussive (other)</td>
<td>Insufficient SOE</td>
<td>4 studies; 274 subjects</td>
<td>Insufficient SOE</td>
</tr>
<tr>
<td></td>
<td>Insufficient SOE</td>
<td>Mucolytics had much smaller effect size vs. placebo, ( p=NS ), in MTM compared with dextromethorphan</td>
<td>Insufficient SOE</td>
</tr>
<tr>
<td>Protussive (mucolytic) vs. antitussive (opiate)</td>
<td>Insufficient SOE</td>
<td>4 studies; 274 subjects</td>
<td>Insufficient SOE</td>
</tr>
<tr>
<td></td>
<td>Insufficient SOE</td>
<td>Mucolytics had much smaller effect size vs. placebo, ( p=NS ), in MTM compared with codeine</td>
<td>Insufficient SOE</td>
</tr>
</tbody>
</table>

CI = confidence interval; ES = effect size; KQ = Key Question; MTM = mixed treatment meta-analysis; NS = not statistically significant; SOE = strength of evidence

*All strength of evidence ratings of “Insufficient” are shaded.*
ratings (risk of bias, consistency, directness, and precision) are available in the main report. Across outcomes and comparisons, although the included evidence was from RCTs with an overall low risk of bias, the findings were inconsistent; the evidence, when available, was indirect (i.e., based on mixed treatment meta-analysis); and the findings, when available, were imprecise. There was insufficient evidence to support conclusions about comparative effectiveness of the interventions for any of our key outcomes. Evidence for other comparisons was too sparse to construct such summary tables.

Discussion

Key Findings

We reviewed 78 studies involving 5,927 patients that evaluated instruments used to assess cough. Our findings suggest that selected cough-specific quality-of-life instruments are valid and reliable for assessing cough frequency, but they show variable correlation with other cough measurement tools. This may be because cough frequency is unidimensional, whereas the impact that cough may have on an individual’s functional status, quality of life, or sense of wellbeing may depend on many other factors. Multidimensional quality-of-life assessments such as the LCQ, CQLQ, and other cough-specific instruments may therefore be more useful than simple cough frequency in assessing meaningful impact of cough. Visual analog scales, although widely used both in research and practice, have little to no data to validate their accuracy in assessing cough, and inconsistent correlations with other cough measurement tools.

We reviewed 48 studies involving 2,923 patients that evaluated nonspecific (or symptomatic) therapies to treat patients with chronic cough. Our review found that a wide variety of pharmaceutical agents have been used to treat the symptom of chronic cough. The opioid and certain nonopioid/nonanesthetic antitussives demonstrated the most promise for managing the symptom of chronic cough. In particular, codeine (with dose response and placebo-controlled data) and dextromethorphan have reasonably good data for reducing cough frequency and severity. However, due to inconsistency and imprecision of results, and small numbers of head-to-head comparisons, the overall strength of evidence is insufficient to draw firm conclusions about the comparative effectiveness of these

<table>
<thead>
<tr>
<th>Treatment Comparison</th>
<th>Cough Severity</th>
<th>Cough Frequency</th>
<th>Adverse Effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Codeine/opiates—Antitussive (opiates) vs. placebo</td>
<td>Low SOE 11 studies; 396 subjectsb 0.63 (95% CI, 0.40 to 0.86; p&lt;0.0001), from MTM</td>
<td>Low SOE 7 studies; 700 subjectsb 0.57 (95% CI, 0.36 to 0.91; p=0.0260), from MTM</td>
<td>Insufficient SOE Imprecise results</td>
</tr>
<tr>
<td>Dextromethorphan—Antitussive (other) vs. placebo</td>
<td>Low SOE 11 studies; 396 subjectsb 0.54 (95% CI, 0.27 to 0.80; p=0.0008), from MTM</td>
<td>Low SOE 7 studies; 700 subjectsb 0.40 (95% CI, 0.18 to 0.85; p=0.0248), from MTM</td>
<td>Insufficient SOE No summary measure</td>
</tr>
<tr>
<td>Protussive (mucolytic) vs. placebo</td>
<td>Insufficient SOE 11 studies; 396 subjectsb 0.14 (95% CI -0.20 to 0.49; p=0.384) from MTM</td>
<td>Insufficient SOE No summary measure</td>
<td>Insufficient SOE No summary measure</td>
</tr>
<tr>
<td>Moguisteine—Antitussive (other) vs. placebo</td>
<td>Low SOE 11 studies; 396 subjectsb 0.62 (95% CI, 0.04 to 1.16, p=0.0366), from MTM</td>
<td>Insufficient SOE 7 studies; 700 subjectsb 0.60 (95% CI, 0.31 to 1.17, p=0.1117), from MTM</td>
<td>Insufficient SOE No summary measure</td>
</tr>
</tbody>
</table>

CI = confidence interval; KQ = Key Question; MTM = mixed treatment meta-analysis; SOE = strength of evidence

*All strength of evidence ratings of “Insufficient” are shaded.

Total number of studies/patients from mixed treatment meta-analysis.
agents. Finally, the evidence exploring the effectiveness of treatments in patients with truly unexplained cough was minimal. We considered the vast majority of study populations to have unresponsive chronic cough. Only three studies, including one of morphine, were clearly in patients with unexplained cough and required subjects to have gone through a diagnostic evaluation to exclude most causes of cough. Interestingly, therapy in each of these studies was associated with a reduction in cough severity, suggesting that chronic unexplained cough can respond to nonspecific therapies aimed at the symptom and not the underlying etiology.

Unfortunately, we identified only one study of a currently available (in the United States) treatment (amoxicillin clavulanate) in children with chronic cough, but the study’s applicability was limited in terms of its sample size and the description of the diagnostic evaluation of cough. Given the lack of studies on treatment of chronic unexplained cough in children, it is not surprising that there were no data on harms in this population.

**Applicability**

It is reasonable to assume that the utility, performance, reliability, and validity of cough instruments may differ between children and adults, between acute and chronic cough conditions, and between underlying etiologies such as asthma, chronic bronchitis, acute rhinitis, lung cancer, and chronic refractory cough. More consistent reporting of patient characteristics such as age, underlying etiology, duration of symptoms and/or illness, overall medical comorbidity, and prior treatment would facilitate evaluations of various cough instruments in important subgroups. For our analysis of instruments for the assessment of cough (KQ 1), most of the studies were conducted in Europe (41 studies, 53%); 32 of these were conducted exclusively in the United Kingdom. Nineteen (24%) studies were conducted in the United States or Canada. Location of study was not, however, obviously related to design, patient, outcome, or analytical characteristics.

By restricting inclusion to trials of patients with unexplained or refractory cough, we improved the applicability of our findings to this population but also decreased the availability of evidence that could be reviewed. Expanding our evidence to include patients with acute cough would have substantially increased the evidence base but greatly reduced the applicability of the findings to the unexplained or refractory chronic cough population. Few studies directly reported assembling patients fitting our intended population of idiopathic or refractory chronic cough. More often patients were selected from persons with chronic cough (of variable duration) with a variety of diseases associated with cough. While we tried to apply criteria to improve applicability (e.g., excluding cystic fibrosis and bronchiectasis), the studies we ultimately included contained more diversity than we intended. In particular, studies with mixed etiologies of cough (including, e.g., patients with tuberculosis or lung cancer) and studies from different eras and geographic locations challenge the usefulness of data on treatment. The majority of studies took place in Europe, with 9 in the United Kingdom and 17 in other countries in Europe (total of 54%); only 9 (19%) took place within the United States or Canada.

For the studies focusing on the adult population, many of the drug treatment trials we identified included drugs that are not currently available in the United States (12 studies, 30 percent). The applicability of the included studies was also reduced given the age of much of the evidence, and therefore of the corresponding interventions and underlying clinical management of the patients. Publication dates ranged from 1953 to 2012, with 32 (76%) of the articles being published before 2000. Given the changes in both available therapies and the diagnosis and treatment of underlying etiologies, more recent studies of contemporary therapies are needed.

**Limitations of the Comparative Effectiveness Review Process**

Our findings have limitations related to the literature and our approach. Important limitations of the literature include: (1) few studies exploring the clinical population of interest (unexplained or refractory chronic cough) and in specific patient subgroups of interest (e.g., children, women, immunocompromised patients); (2) variable definitions of chronic cough; (3) diverse etiologies of cough that might respond differently to different therapies; (4) incomplete reporting of patient characteristics, study design, or outcomes; (5) small sample sizes and short duration of followup; (6) lack of gold standard outcomes to assess efficacy and tolerability; and (7) inconsistent reporting of comparative statistical analyses. In addition, most of the studies were comparatively old, and as such the evidence base suffers from age because of advances in clinical trial methodology, improved diagnostic evaluation of cough, and development of valid and reliable measures for cough and cough-specific quality of life.

Our review methods also had limitations. Our study was limited to English-language publications. In addition, even within patients with chronic cough, the target
population of patients with unexplained chronic cough or refractory chronic cough with a known underlying etiology was difficult to identify. Rarely was a thorough negative diagnostic evaluation performed to assemble a group with unexplained chronic cough; in the case of studies of patients with a known underlying etiology, seldom was previously tried therapy described well enough to determine whether patients were treatment-refractory. In general, we considered use of a symptomatic treatment in a population with a known underlying etiology to imply refractory cough unless patients were noted to be treatment-naïve; certain etiologies, however, were considered differently. For example, most studies of cough-variant asthma, a common cause of chronic cough in children, which is usually highly responsive to appropriate asthma management, were excluded.

It is possible that our a priori definition of chronic cough in childhood (i.e., persisting at least 4 weeks if <14 years of age, or 8 weeks if 14 years or older) was too long and did not reflect care delivery. However, our decision to include studies that described their population as suffering with chronic cough regardless of time cut-off may have mitigated this problem. Focusing on nonspecific or symptomatic treatments to the exclusion of treatments aimed at specific causes of chronic cough proved more complicated to implement than we had anticipated. Certain therapies that we classified as specific (e.g., antihistamines and decongestants for upper airway cough syndrome) are sometimes referred to as nonspecific. Furthermore, some other specific treatments were occasionally tested as nonspecific treatments in populations that did not meet diagnostic criteria for conditions for which the specific treatment would be appropriate. Finally, we grouped antitussive and protussive drugs into subsets that sometimes included pharmacologically diverse agents or separate drugs with certain similarities.

**Research Gaps**

We found sufficient evidence to suggest that the LCQ and CQLQ (for adults) and the PC-QOL (for children) may be valid instruments for assessing severity/QOL of cough, and that electronic recording devices, in general, appear to be valid assessments of cough frequency compared with human cough counts. Unfortunately, however, the current evidence base is insufficient to provide conclusive findings related to the comparative effectiveness of available therapies for patients with unexplained or refractory chronic cough. There are, therefore, numerous areas of evidence gaps and areas for potential future research. We used the framework recommended by Robinson et al. to identify gaps in evidence and describe why these gaps exist. This approach considers PICOTS (population, intervention, comparator, outcomes, timing, and setting) to identify gaps and classifies gaps as due to (1) insufficient or imprecise information, (2) biased information; (3) inconsistency or unknown consistency, and (4) not the right information. Results are as follows:

**KQ 1—Instruments used to assess cough:**

- Evidence establishing the responsiveness, validity, reliability, and consistency of available assessment instruments other than the LCQ and CQLQ, and building on available evidence for the LCQ and CQLQ instruments
- Additional validation or measurement studies focusing on the pediatric population in addition to the limited studies that report on the PC-QOL
- Development and validation of child/patient-completed, cough-specific quality-of-life instruments (as opposed to parent/proxy instruments such as the PC-QOL)
- Feasibility of cough assessment instruments in usual care (outside of RCTs or validation studies)
  - Insufficient evidence currently exists; could be explored through observational studies
- Uncertainty about the effects of patient self-reporting, parent reporting, or provider reporting in use of cough assessment tools
  - Insufficient evidence currently exists; could be explored through observational studies
- Incomplete evidence regarding the minimally important difference of cough frequency or severity/QOL instruments
- Impact of measurement tools on therapeutic efficacy or patient outcome efficacy

**KQ 2—Nonspecific therapies for chronic cough:**

- Comparative effectiveness of pharmacological therapies in the adult population
  - Current evidence is both imprecise and inconsistent. Additional comparative RCTs of contemporary and available agents are needed.
- Comparative effectiveness of pharmacological therapies in the pediatric population
  - Current evidence is insufficient and does not reflect available therapies. Comparative RCTs of contemporary and available agents specific to the pediatric population are needed.
• Comparative effectiveness of nonpharmacological therapies in both adult and pediatric populations
  – Current evidence is insufficient. Comparative RCTs of contemporary and available agents specific in both adult and pediatric populations are needed.
  – Additional RCTs or potentially patient-level meta-analyses of existing and future RCTs focusing on subpopulations of interest including women, pregnant women, patients with specific underlying etiologies, immunocompromised patients, and patients with a history of substance abuse

• Comparative effectiveness of available therapies in impacting health utilization and costs
  – Insufficient evidence currently exists; could be explored through observational studies

• Comparative effectiveness of available therapies in impacting cough severity, frequency, and quality of life
  – Current evidence is both imprecise and inconsistent. Additional comparative RCTs using standardized instruments are needed.

Conclusions

There is no established gold standard for assessing either frequency or severity/QOL of cough, thereby making it difficult to quantitatively assess test accuracy for cough instruments. Validity of severity/QOL questionnaires was generally demonstrated in the published literature by correlation with other cough assessment instruments, whereas validity of cough recording devices was generally demonstrated using human cough counts as the reference standard. Reliability of questionnaires was generally demonstrated by test-retest correlation and by demonstrating internal consistency. Several instruments, including the LCQ, CQLQ, and the PC-QOL, show good internal consistency but variable correlation with other cough measurement tools. This suggests that these tools may be reliable but demonstrate variable validity. The lack of validated reference tests and the diverse number of instruments used among treatment evaluations also complicates comparisons across studies. We identified no evidence exploring the impact of cough assessment instruments on therapeutic efficacy or patient outcome efficacy.

A wide variety of pharmaceutical agents have been used to treat the symptom of chronic cough, including opioid, anesthetic, and nonopioid/nonanesthetic antitussives; expectorant and mucolytic protussives; antihistamines; antibiotics; inhaled corticosteroids; and inhaled anticholinergics. There were relatively few good-quality studies focusing on chronic cough using reliable outcome measurements over durations of followup pertinent to chronic cough. The opioid and certain nonopioid/nonanesthetic antitussives most frequently demonstrated efficacy for managing the symptom of chronic cough compared with placebo, but there were insufficient data to draw conclusions between therapies. Data on nonpharmacological therapies for chronic cough are extremely limited, as are data on the management of unidentified or refractory chronic cough in children.

Our systematic review highlights the clear need for further studies in patient populations with unexplained or refractory chronic cough as determined by current diagnostic and empiric treatment recommendations. Further, it shows the need for more systematic design and reporting of these studies and assessment of patient-centered outcomes.

References


Full Report


Errata

Table A in the Executive Summary and Tables 6, 11, and 12 in the full report have been updated to reflect the following changes:

1. **CQLQ**—corrected sample size and correlation coefficients for French 2002 paper for Internal Consistency.
3. **PC-QOL**—added data from Newcombe 2010 study for Repeatability.

The text and conclusions remain unchanged.