



Effective Health Care Program

Pulmonary Arterial Hypertension: Screening, Management, and Treatment

Executive Summary

Background

Epidemiology and Etiology of Pulmonary Arterial Hypertension

Pulmonary arterial hypertension (PAH), a subcategory of pulmonary hypertension (PH), is a rare and progressive disease whose prevalence is estimated to be between 15 and 50 cases per 1 million adults.¹ While the pathophysiology is not well understood, both genetic and environmental factors have been found to contribute to changes in the pulmonary vasculature, causing increased pulmonary vascular resistance. This increased resistance, if unrelieved, progresses to right ventricular pressure overload, dysfunction, and ultimately right heart failure and premature death.² The causes of PAH are numerous and are listed in Table A, taken from the Fourth World Symposium on PAH (2008).³ Before the availability of disease-specific therapy in the mid-1980s, the median life expectancy at the time of diagnosis was 2.8 years.^{1,4}

Screening and Diagnosis

There are two separate populations for which screening for PAH needs to be considered. First, there are patients with symptoms that raise the suspicion of PAH. The symptoms of PAH can be

Effective Health Care Program

The Effective Health Care Program was initiated in 2005 to provide valid evidence about the comparative effectiveness of different medical interventions. The object is to help consumers, health care providers, and others in making informed choices among treatment alternatives. Through its Comparative Effectiveness Reviews, the program supports systematic appraisals of existing scientific evidence regarding treatments for high-priority health conditions. It also promotes and generates new scientific evidence by identifying gaps in existing scientific evidence and supporting new research. The program puts special emphasis on translating findings into a variety of useful formats for different stakeholders, including consumers.

The full report and this summary are available at www.effectivehealthcare.ahrq.gov/reports/final.cfm.

insidious and nonspecific and may include shortness of breath, fatigue, weakness, chest pain, syncope, leg swelling, and abdominal distention. Symptoms that are present at rest suggest advanced disease.¹



Agency for Healthcare Research and Quality
Advancing Excellence in Health Care • www.ahrq.gov

Effective
Health Care

Table A. Updated clinical classification of pulmonary hypertension (Dana Point, 2008)^a

1.	Pulmonary arterial hypertension (PAH)
1.1	Idiopathic PAH
1.2	Heritable
1.2.1	BMPR2
1.2.2	ALK1, endoglin (with or without hereditary hemorrhagic telangiectasia)
1.2.3	Unknown
1.3	Drug and toxin-induced
1.4	Associated with:
1.4.1	Connective tissue disease
1.4.2	HIV infection
1.4.3	Portal hypertension
1.4.4	Congenital heart diseases
1.4.5	Schistosomiasis
1.4.6	Chronic hemolytic anemia
1.5	Persistent pulmonary hypertension of the newborn
1[*].	Pulmonary veno-occlusive disease (PVOD) and/or pulmonary capillary hemangiomatosis (PCH)
2.	Pulmonary hypertension owing to left heart disease
2.1	Systolic dysfunction
2.2	Diastolic dysfunction
2.3	Valvular disease
3.	Pulmonary hypertension owing to lung diseases and/or hypoxemia
3.1	Chronic obstructive pulmonary disease
3.2	Interstitial lung disease
3.3	Other pulmonary diseases with mixed restrictive and obstructive pattern
3.4	Sleep-disordered breathing
3.5	Alveolar hypoventilation disorders
3.6	Chronic exposure to high altitude
3.7	Developmental abnormalities
4.	Chronic thromboembolic pulmonary hypertension (CTEPH)
5.	Pulmonary hypertension with unclear multifactorial mechanisms
5.1	Hematologic disorders: myeloproliferative disorders, splenectomy
5.2	Systemic disorders: sarcoidosis, pulmonary Langerhans cell histiocytosis: lymphangioliomyomatosis, neurofibromatosis, vasculitis
5.3	Metabolic disorders: glycogen storage disease, Gaucher disease, thyroid disorders
5.4	Others: tumoral obstruction, fibrosing mediastinitis, chronic renal failure on dialysis

ALK1 = activin receptor-like kinase type 1; BMPR2 = bone morphogenetic protein receptor type 2; HIV = human immunodeficiency virus

^aFourth World Symposium on PAH in Dana Point, CA (2008).

Table reprinted from the Journal of the American College of Cardiology, Vol 54, No. 1, Suppl S, Simonneau G, Robbins IM, Beghetti M, et al., Updated Clinical Classification of Pulmonary Hypertension, Pages No. S43-54, Copyright 2009, with permission from Elsevier.³

Since these symptoms are nonspecific, screening may be necessary to help the physician decide whether the patient should undergo a diagnostic workup for PAH, or whether other conditions should be considered. The other population is patients with medical conditions that put them at risk for PAH. In these patients screening tests may be used to identify patients with asymptomatic elevation of pulmonary artery pressures, who might be more closely monitored for the development of symptoms or progressive disease or offered a diagnostic workup for PAH and possibly treatment for early disease.

Once screening indicates the possibility of PAH, diagnostic tests are necessary to confirm the presence of elevated right-sided heart pressures and to exclude valvular, primary myocardial, chronic lung disease, thromboembolic disease, and miscellaneous other causes of pulmonary hypertension (PH). The reference standard for diagnosing PAH is right heart catheterization (RHC), which is invasive but generally safe. In a retrospective and prospective study by Hoeper et al.,⁵ the rate of serious complications in patients undergoing RHCs for evaluation of pulmonary hypertension was 1.1 percent and included bleeding, vasovagal reactions, systemic hypotension, arterial injury, hypertensive crisis, pneumothorax, and cardiac arrhythmias. The procedure-related mortality was 0.055 percent.⁵

RHC not only confirms the diagnosis of PAH but also provides prognostic hemodynamic information (mean right atrial pressure [mRAP], pulmonary vascular resistance)⁶ to direct treatment decisions. A small subset of patients with PAH, when challenged with a short-acting pulmonary vasodilator, will experience a drop in mean pulmonary artery pressure of at least 10 mmHg (20%) to below 40 mmHg while maintaining cardiac output; this predicts a favorable long-term response to calcium channel blockers.¹

Since PAH is a progressive disease, regular reassessment is needed to monitor response to treatment and adjust prognosis. In addition to the assessment of clinical symptoms, RHC has traditionally been the means by which patients' clinical course is monitored; however, transthoracic echocardiography has emerged as a possible alternative monitoring mechanism because of its availability, safety, and relatively low cost. The number of echocardiographic modalities has increased substantially, providing unique insights into the structure and function of the right heart in patients with pulmonary hypertension.⁷ However, this test has not been definitively validated as a substitute for RHC in patients with PAH. Finally, the role of biomarkers has not been fully established in the management and prognosis of PAH. Defining whether

biomarkers alone or biomarkers plus echocardiography might be superior to echocardiography alone for informing treatment decisions is a necessary first step in establishing a noninvasive, multifaceted approach to the management of PAH.

Role of Echocardiography

The role of echocardiography in the diagnosis and management of patients with PAH has evolved over time, and has been proposed for screening, assessing prognosis and evaluating response to treatment. Screening high-risk individuals for PAH generally begins with a transthoracic echocardiogram.⁸ Echocardiography can estimate the right ventricular systolic pressure and identify other signs of PH including increased right-sided chamber size and wall thickness. Most often, the peak velocity of the tricuspid regurgitant (TR) jet is measured by Doppler and—along with an estimate of right atrial pressure (RAP) based on inspiratory collapse and size of the inferior vena cava—TR jet is used to estimate the systolic pulmonary artery pressure (sPAP). However, a significant proportion of patients have no measureable TR jet. Estimates are often inaccurate compared with RHC; up to 60 percent of echocardiography estimates were more than 10 mmHg off from RHC measurement in one large multicenter registry of PAH patients.⁹

Furthermore, sPAP is dependent on right ventricle (RV) systolic function and stroke volume. In later stages of PH, RV function deteriorates, which can lessen the degree of sPAP elevation and lead to an underestimate of pulmonary vascular resistance (PVR). More recent echocardiographic-based methods have focused on evaluating RV systolic function. Therefore, although transthoracic echocardiography is the standard screening test for PAH, it is less than completely accurate and there is uncertainty as to which echocardiographic measurements are most useful.

Several studies have investigated the use of echocardiography in establishing prognosis in PAH. In a study of patients with systemic sclerosis (n=155), 3-year survival rates were lower in 47 patients with right ventricular systolic pressure (RVSP) ≥ 36 mmHg as calculated by Doppler echocardiography compared with patients with RVSP < 36 mmHg (67% vs. 86%, $p < 0.01$).¹⁰ Another study of patients with PAH (n=80) using echocardiography to calculate right ventricular free wall strain found that patients with strain worse than -12.5 percent were associated with increased 6-month disease progression and increased mortality at 1 year (unadjusted hazard ratio 6.2).¹¹ Uncertainty remains

regarding which echocardiographic measure(s) have prognostic value, although tricuspid annular plane systolic excursion (TAPSE) and pericardial effusion have been proposed.¹²

Traditionally, RHC assessment of hemodynamics is recommended to demonstrate treatment response;¹² echocardiography has seldom been studied in this role.

Role of Biomarkers

Because of the limitations of echocardiography, the potential role of biomarkers in screening for and managing of PAH has been the subject of increasing interest over the last decade. Brain natriuretic peptide (BNP) and N-terminal proBNP (NT-proBNP) are two biological substances found in the blood that have been studied as a screening test in patients at risk for PAH and which have been shown to correlate well with the presence of disease.^{13,14} Other biomarkers currently under investigation include atrial natriuretic peptide, endothelin-1, uric acid, troponin T, nitric oxide, asymmetric dimethylarginine, cyclic guanosine monophosphate, D-dimer, and serotonin. Several of these biomarkers have been shown to correlate with prognosis and mortality, either alone or in conjunction with other traditional measurements such as the 6-minute walk distance (6MWD) test, functional class assessment, and pulmonary hemodynamics.¹⁵ Select biomarkers may even be superior to traditional testing. Patients with idiopathic and familial PAH were shown to exhibit dysregulation over a broad range of inflammatory cytokines, such as tumor necrosis factor- α and interleukin-6, when compared with healthy controls, findings which correlated better with prognosis than 6MWD and pulmonary hemodynamics.¹⁶ It remains uncertain to what extent the correlations and case-control comparisons offer valid prognostic information for individual patients and can be used to make better management decisions.

Treatment Strategies

Medications

There has been rapid development and approval of vasodilator medications for PAH over the past three decades. Currently, there are four main classes of medications used to treat PAH:¹⁷

- Calcium channel blockers:
 - Amlodipine
 - Diltiazem
 - Nifedipine

- Prostacyclin analogues:
 - Epoprostenol
 - Iloprost
 - Treprostinil
- Endothelin receptor antagonists:
 - Bosentan
 - Ambrisentan
- Phosphodiesterase type 5 inhibitors:
 - Sildenafil
 - Tadalafil

These PAH medications have been shown to improve dyspnea, 6MWD, pulmonary hemodynamics, and functional class. Calcium channel blockers are associated with long-term (>1 year) improvements in hemodynamics and functional status in most of those patients who show acute vasoreactivity testing response; however, acute vasoreactivity is seen in a minority of patients tested.¹⁸ The limited usefulness of calcium antagonists—as well as the poor prognosis and diminished quality of life associated with PAH—reinforces the need for new drug therapies and improved delivery of current medications. Limited data suggest that epoprostenol and bosentan may provide a survival benefit; however, this end point has not been studied consistently between the medications.¹⁹ The three medications most recently approved by the U.S. Food and Drug Administration for PAH are: (1) inhaled treprostinil, a new delivery system for this prostacyclin analogue, (2) tadalafil, a new phosphodiesterase type 5 inhibitor, and (3) ambrisentan, an endothelin receptor antagonist. With the exception of tadalafil, these new medications were discussed in the Expert Consensus Document on Pulmonary Hypertension released in 2009 by the American College of Cardiology Foundation and the American Heart Association.¹⁹ Since then, however, numerous studies have been published regarding the safety and efficacy of these new medications. Also, more data have been published on the older medications for PAH. These new data may clarify any effect on mortality and gauge the comparative effectiveness of these drugs.

Additionally, combination drug therapy (using multiple drugs with different mechanisms of action) is an important area of research and may be the most promising way to improve clinical outcomes although at higher cost.² Combination therapy was addressed in the 2009 ACCF/AHA publication, and several studies have since been published on this topic. In order to

optimize PAH care, newer information regarding the latest drugs and combination therapies should be systematically reviewed.¹⁷

Scope and Key Questions

This comparative effectiveness review was funded by the Agency for Healthcare Research and Quality (AHRQ). It was designed to evaluate the comparative validity, reliability, and feasibility of echocardiography and biomarker testing for the diagnosis and management of PAH in addition to clarifying whether the use of echocardiography and biomarkers affects decisionmaking and clinical outcomes. We also wanted to address which medications are effective for treating PAH and how the newer medications compare with older ones and with each other. Further, there was a need for clarity about whether combination therapy is more effective than monotherapy and what effect monotherapy or combination therapy has on intermediate-term and long-term outcomes.

The Key Questions (KQs) considered in this comparative effectiveness review were:

KQ 1: For patients with suspected pulmonary arterial hypertension (PAH) and asymptomatic patients at high risk for PAH, what are the comparative effectiveness and safety of echocardiography versus echocardiography plus biomarkers as screening modalities before right heart catheterization to establish the diagnosis of PAH (i.e., what is their comparative diagnostic accuracy efficacy)?

KQ 2: For patients with PAH, what are the comparative effectiveness and safety of (a) echocardiography versus biomarkers and (b) echocardiography versus echocardiography plus biomarkers in managing PAH and on intermediate-term (≤ 90 days) and long-term (> 90 days) patient outcomes?

KQ 3: For patients with PAH, what are the comparative effectiveness and safety of monotherapy or combination therapy for PAH using calcium channel blockers, prostanoids, endothelin receptor antagonists, or phosphodiesterase inhibitors on intermediate-term and long-term patient outcomes?

Figures A and B show the analytic framework for this comparative effectiveness review.

Methods

The methods for this comparative effectiveness review follow those suggested in the AHRQ “Methods Guide

for Effectiveness and Comparative Effectiveness Reviews” (available at www.effectivehealthcare.ahrq.gov/methodsguide.cfm; hereafter referred to as the Methods Guide).²⁰

Input From Stakeholders

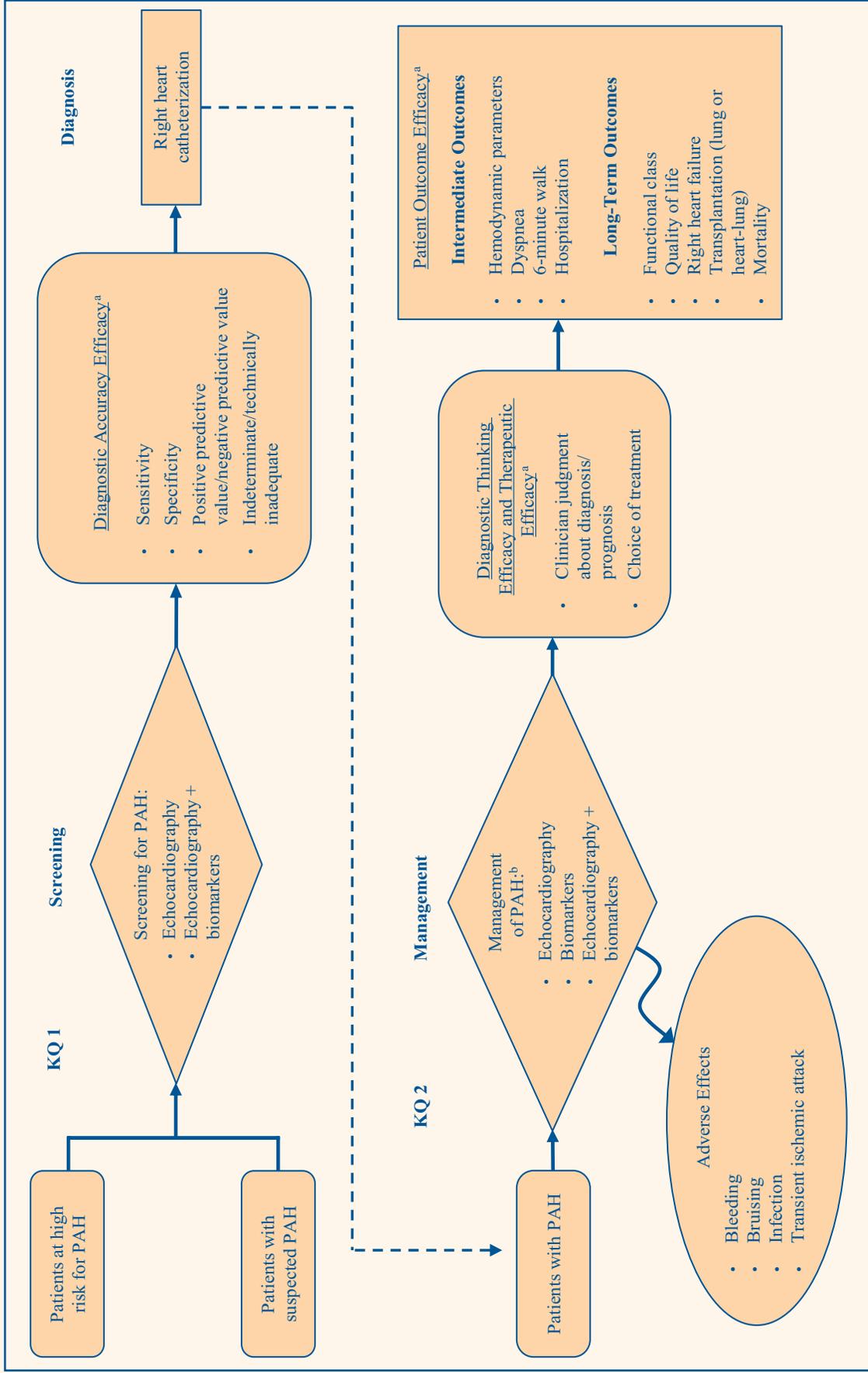
During the topic refinement stage, we solicited input from Key Informants representing clinicians (in pulmonology, cardiology, and pathology), patients, scientific experts, and Federal agency officials, to help define the KQs. The KQs were then posted for public comment for 30 days, and the comments received were considered in the development of the research protocol. We next convened a Technical Expert Panel (TEP), comprising clinical, content, and methodological experts, to provide input in defining populations, interventions, comparisons, or outcomes as well as 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 of interest. Any potential conflicts of interest were balanced or mitigated. Neither Key Informants nor members of the TEP did analysis of any kind or contribute to the writing of the report.

Literature Search Strategy

To identify the relevant published literature, we searched PubMed[®], Embase[®], and the Cochrane Database of Systematic Reviews. We limited the search to English-language studies conducted from 1995 to the present for KQs 1 and 2, and 1990 to the present for KQ 3; prior to 1990, newer drug treatments were not available and prior to 1995 older echocardiographic and biomarker testing technology was less applicable. We supplemented the electronic searches with a manual search of citations from a set of key primary and review articles. All citations were imported into an electronic database (EndNote[®] X4; Thomson Reuters, Philadelphia, PA).

We also searched the gray literature of study registries and conference abstracts for relevant articles from completed studies. Gray literature databases included ClinicalTrials.gov; metaRegister of Controlled Trials; ClinicalStudyResults.org; World Health Organization International Clinical Trials Registry Platform Search Portal; and ProQuest COS Conference Papers Index. Scientific information packets were requested from the manufacturers of medications and devices and reviewed for relevant articles from completed studies not previously identified in the literature searches.

Figure A. Analytic framework for KQs 1 and 2

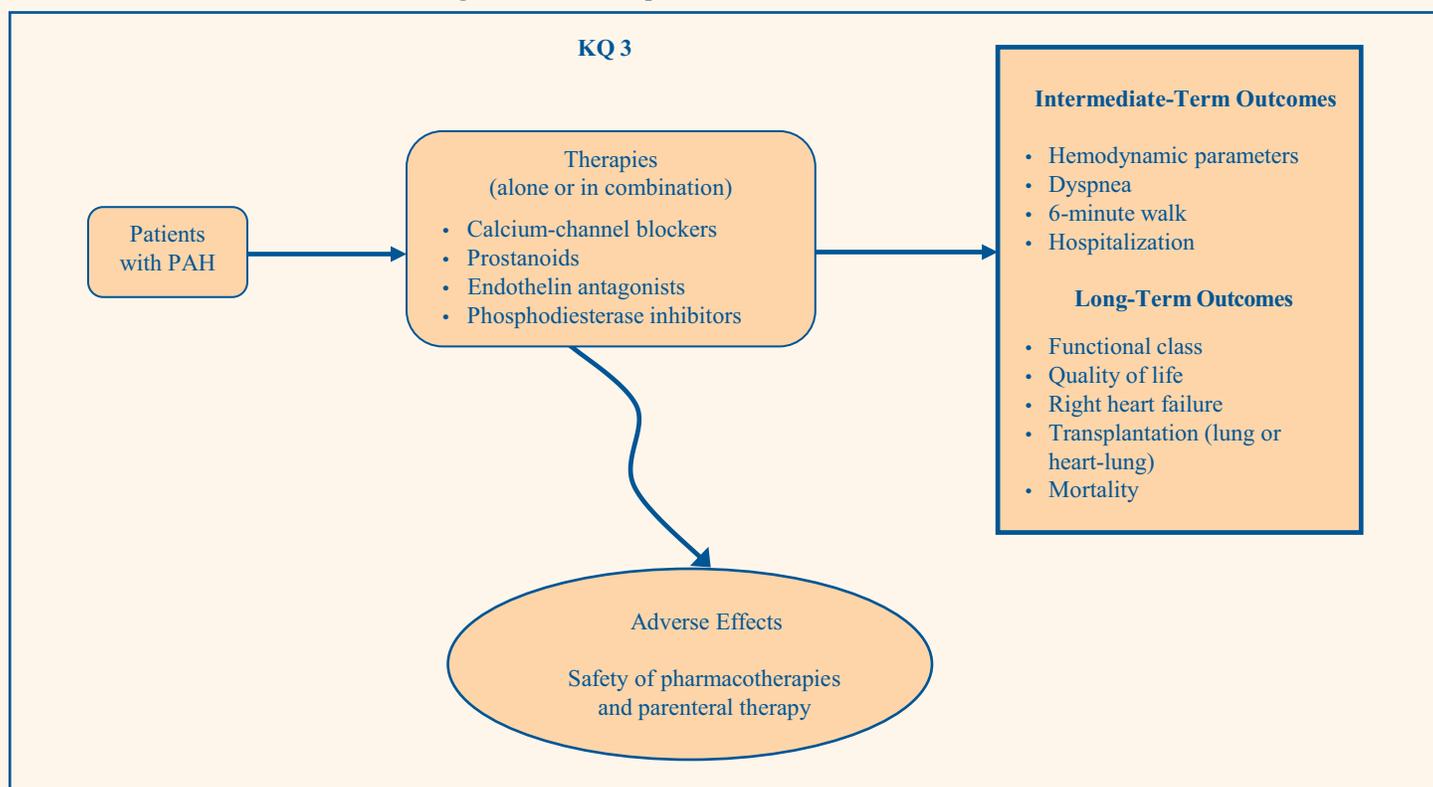


KQ = Key Question; PAH = pulmonary arterial hypertension

^aFryback DG, Thornbury JR. The efficacy of diagnostic imaging. Med Decis Making 1991;11(2):88-94.

^bIn conjunction with routine clinical assessment (functional class, dyspnea, 6-minute walk).

Figure B. Analytic framework for KQ 3



KQ = Key Question; PAH = pulmonary arterial hypertension

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 the full report. For KQ 1, the search focused on studies that reported the accuracy of echocardiography, biomarkers, or the combination of these tests for diagnosis of PAH in patients suspected of having PAH or in asymptomatic patients at high risk for PAH. For KQ 2, the search focused on English-language studies describing data on how echocardiographic or biomarker testing among patients with PAH was related to diagnostic thinking efficacy and therapeutic efficacy (clinician judgment about diagnosis or prognosis or choice of treatment) and patient outcome efficacy (prognosis related to intermediate and long term outcomes, including hemodynamic parameters, dyspnea, 6MWD, functional status, and mortality). For KQ 3, the search focused on the effect of pharmacotherapy with prostanoids (epoprostenol, treprostinil, iloprost), endothelin receptor antagonists (bosentan, ambrisentan) or phosphodiesterase inhibitors (sildenafil, tadalafil) on intermediate-term and long-term outcomes as well as adverse effects in patients with PAH. For KQ 3, we chose not to use composite endpoints such as time to clinical worsening (TTCW) due to weighting issues and lack of comparability among studies.

Study Selection

Using the prespecified inclusion and exclusion criteria, two reviewers independently examined titles and abstracts for potential relevance to the KQs. Articles included by any reviewer underwent full-text screening. At the full-text screening stage, two independent reviewers read each article to determine if it met eligibility criteria. Disagreements were resolved by discussion or by a third-party arbitrator, if needed. Relevant systematic review articles, meta-analyses, and methods articles were flagged for hand-searching and cross-referencing against the library of citations identified through electronic database searching. All screening decisions were made and tracked in a Distiller SR database (Evidence Partners, Inc., Manotick, ON, Canada).

Data Extraction

The investigative team created data abstraction forms and evidence table templates. Based on clinical and methodological expertise, two investigators were assigned to the research questions to abstract data from the eligible articles. One investigator abstracted the data, and the second overread the article and the accompanying abstraction to check for accuracy and completeness.

Disagreements were resolved by consensus or by obtaining a third reviewer's opinion if consensus was not reached between the first two investigators.

Quality Assessment of Individual Studies

We evaluated the quality of individual studies using the approach described in the Methods Guide.²⁰ To assess methodological quality, we employed the Methods Guide strategy to: (a) apply predefined criteria for quality and critical appraisal and (b) arrive at a summary judgment of the study's quality. To indicate the summary judgment of the quality of the individual studies, we used the summary ratings of good, fair, or poor. For studies of diagnostic tests (KQ 1 and KQ 2), we used QUADAS-2,²¹ a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. QUADAS-2 describes risk of bias in four key domains: patient selection, index test(s), reference standard, and flow and timing; each domain is rated as having high, low, or unclear risk of bias. For studies of pharmacotherapies, we used the Cochrane Risk of Bias tool, which evaluates random sequence generation, allocation concealment, blinding of participant and personnel, blinding of outcome assessment, completeness of outcome data, completeness of outcome reporting, and other indications that the studies are unbiased.

Two raters independently evaluated each study and resolved differences by consensus; if they could not reach consensus, they rated the item as unclear, and the rationale for each differing assessment was described. They described results for individual domains. If the distribution of ratings permitted, they examined methodological domains for association with the effects obtained in meta-analysis.

To indicate the summary judgment of the quality of the individual studies, we used the summary ratings of good, fair, or poor based on the studies' adherence to well accepted standard methodologies and the adequacy of their reporting.

Data Synthesis

Quantitative synthesis (i.e., meta-analysis) was done when we found multiple studies of similar design, population, intervention, comparator, and outcome that reported sufficient data for analysis. When a meta-analysis was appropriate, we used random-effects models to quantitatively synthesize the available evidence. We use meta-analyses both to quantify and to attempt to explain between-study variation as well as to calculate summary estimates. When a meta-analysis was not appropriate we

described the reasons, presented data in tabular form, and summarized studies either individually or qualitatively.

For sensitivity and specificity data, we used a binomial model to calculate summary estimates of sensitivity and specificity and associated confidence intervals and summary receiver operating characteristic (ROC) curve using SAS statistical software. Sensitivity analyses were conducted using summary ROC meta-analysis using the diagnostic odds ratio with dr-ROC software (Diagnostic Research Design and Reporting; Glenside, PA). For meta-analysis of correlation coefficients and hazard ratios for observational studies, we used a random effects model implemented in SAS (SAS Institute Inc.; Cary, NC). For treatment effects meta-analysis, we used a random effects model meta-analysis implemented in Comprehensive Meta-Analysis Software (Version 2.2.064, Biostat; Englewood, NJ). We tested for heterogeneity using graphical displays and test statistics (Q and I² statistics), while recognizing that the ability of statistical methods to detect heterogeneity may be limited, particularly when the number of studies is small. We present summary estimates and confidence intervals in our data synthesis.

Strength of the Body of Evidence

The strength of evidence for each KQ was assessed using the approach described in the Methods Guide.²² In brief, the approach requires assessment of four domains: risk of bias, consistency, directness, and precision. Additional domains were used when appropriate: 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. A grade of insufficient was assigned when no evidence was available or when evidence on the outcome was too weak, sparse, or inconsistent to permit any conclusion to be drawn.

Diagnostic evaluation studies (KQs 1 and 2) are generally indirect, as the link between the test intervention and outcome is mediated by prognosis, management, and the effectiveness of treatments. As a rule of thumb, we considered correlation coefficients greater than 0.7 as strong association, 0.40 to 0.69 as moderate, and less than 0.40 as weak. In our summary strength of evidence assessments for KQs 1 and 2, lack of directness was weighed less heavily and risk of bias most heavily. Thus, we allowed high strength of evidence levels despite the lack of directness among these studies.

Applicability

We assessed applicability across our KQs using the PICOTS format as described in the Methods Guide.^{20,23} We used these data to evaluate the applicability to clinical practice, paying special attention to study eligibility criteria, demographic features of the enrolled population (such as age, ethnicity, and sex) in comparison with the target population, the version or characteristics of the intervention used in comparison with therapies currently in use (such as specific components of treatments considered to be supportive therapy), and the clinical relevance and timing of the outcome measures. We summarized issues of applicability qualitatively. Because applicability issues may differ for different users, we reported across a range of potential applicability issues.

In assessing the applicability of diagnostic evaluation studies, we were particularly concerned about the prevalence of PAH versus PH in the study populations compared, the spectrum of underlying type of PAH, and the assessment of adverse events associated with testing. In assessing PAH drug trials, we were particularly concerned with whether the researchers had assessed the severity of illness; the use of run-in periods; attrition before randomization; the use of surrogate or combined outcome measures; short study duration; the reporting of adverse events, in particular including those related to administration or monitoring of treatment; whether the sample size was sufficient to assess minimally important differences from a patient perspective; and the use of intention-to-treat-analysis.

Results

Figure C depicts the flow of articles through the literature search and screening process. Searches of PubMed®, the Cochrane Database of Systematic Reviews, and Embase® yielded 8,256 citations, 1,626 of which were duplicate citations. Manual searching identified 46 additional citations, for a total of 6,676 citations. After applying inclusion/exclusion criteria at the title-and-abstract level, 1,324 full-text articles were retrieved and screened. Of these, 1,127 were excluded at the full-text screening stage, leaving 197 articles (representing 186 studies) for data abstraction. (Article counts by KQ do not add to 197 because some studies were included for multiple KQs.)

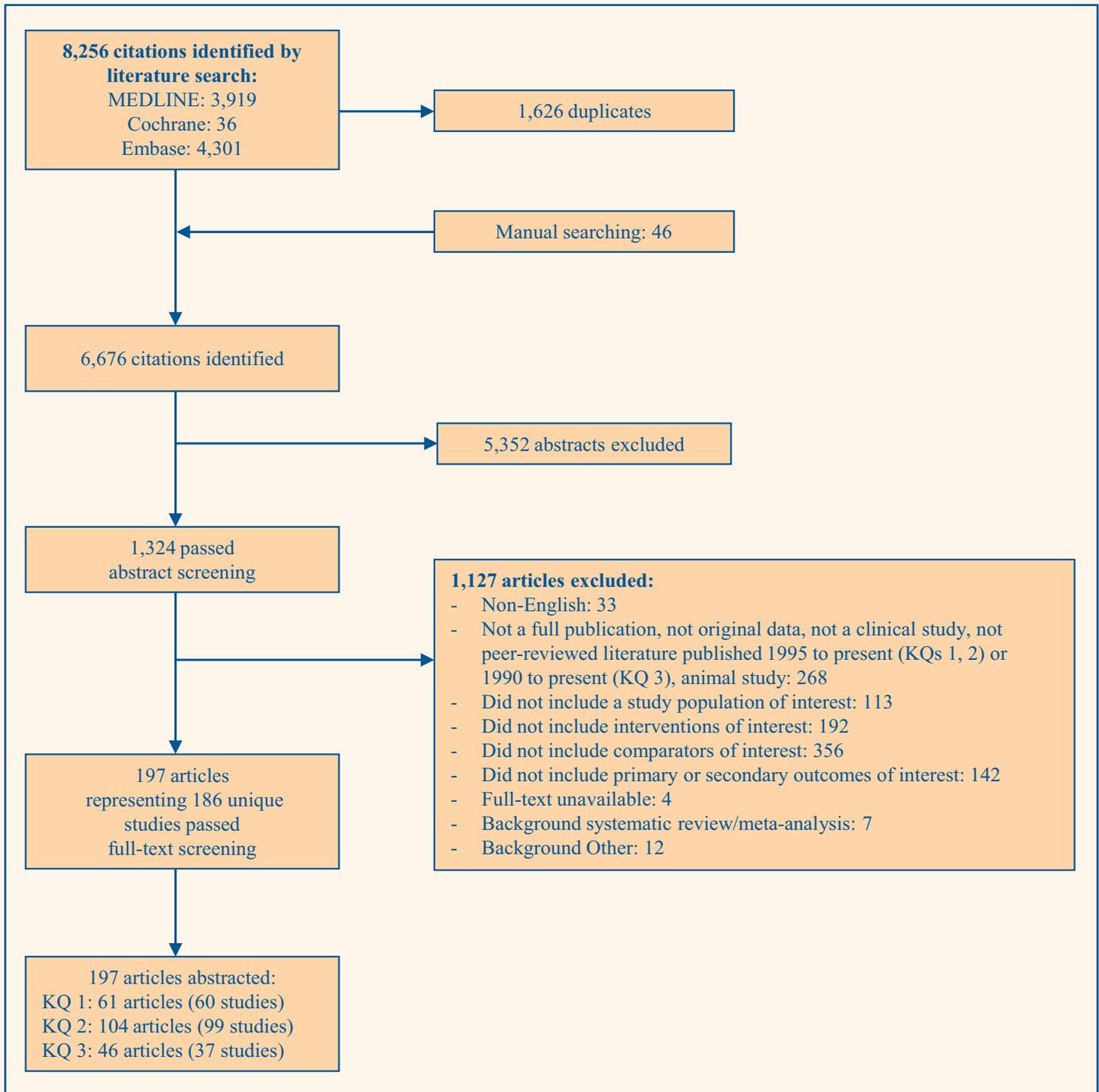
KQ 1: Screening for Pulmonary Arterial Hypertension

Key Points from the Results chapter are:

- For patients suspected of having PAH with elevated sPAP by echo, additional testing with the biomarker NT-proBNP may identify more patients who do not have PAH, compared with echo sPAP alone (based on one good-quality prospective cohort study) (low strength of evidence).
- For patients suspected of PAH, echocardiographic estimation of RVSP (or TG) by TRV, sPAP by TRV and RAP, and PVR by (TRV/VTI_{RVOT}) shows reasonably good accuracy, compared with RHC (moderate strength of evidence).
- Both for asymptomatic patients at high risk for PAH and for symptomatic patients suspected of PAH, natriuretic peptide testing (with either BNP or NT-proBNP) shows highly variable sensitivity and specificity estimates (not simultaneously high) for pulmonary hypertension (PH) or PAH diagnosis (low strength of evidence) and moderate correlation with hemodynamic measures by RHC (moderate strength of evidence).
- There were no studies of the safety of biomarker and echocardiography testing, nor were there any studies of combined echocardiographic and biomarker screening of asymptomatic patients at high risk for PAH (insufficient strength of evidence).

We identified one good-quality study involving 372 patients that compared echocardiography with echocardiography plus biomarkers in patients with suspected PAH, most of whom were symptomatic. There were no other studies that directly compared combinations of echocardiographic and biomarker testing. In order to draw inferences about the comparative effectiveness of other tests, we reviewed the diagnostic accuracy of independent echocardiographic or biomarker testing compared with RHC. By evaluating the relative diagnostic performance of these tests versus a reference standard of RHC, one can impute the comparative effectiveness via indirect comparisons. We identified 60 unique studies involving a total of 7,096 patients that describe the effectiveness of echocardiography or biomarkers in patients with suspected PAH, or in asymptomatic patients

Figure C. Literature flow diagram



KQ = Key Question

Note: Some studies were included for multiple KQs.

at high risk for PAH, as screening modalities before RHC to establish the diagnosis of PAH. Symptom status of study populations consisted of asymptomatic (3 studies; 481 patients), symptomatic (41 studies; 4,394 patients), mixed (8 studies; 1,186 patients), and symptoms not described (8 studies; 1,035 patients). Table B summarizes the findings of our review and the strength of evidence ratings for the available outcomes of sensitivity, specificity, correlation coefficients, and adverse effects of biomarker and echocardiographic tests. Details about the specific components of these ratings (risk of bias, consistency, directness, and precision) are available in the full report. Among biomarker studies, natriuretic peptide (BNP, NT-proBNP) was the only biomarker reported in more than one study; therefore it is the only biomarker for which we generated a strength of evidence table. Limited data on cyclic GMP, asymmetric dimethylarginine (ADMA) and endothelin-1 were reported in one study each. Likewise, the echocardiographic estimates of sPAP and PVR (TRV/VTI_{RVOT}) were the only echocardiographic parameters reported in a sufficient number of studies to support strength of evidence rating. Limited data on FAC, RA size, RIMP, RV size, tricuspid lateral annular systolic velocity (S'), and TAPSE are described in the full report.

One good-quality study evaluated the diagnostic value of serum NT-proBNP in a noninvasive diagnostic decision algorithm that also used data from electrocardiography and echocardiography. Among 69 patients without RV strain on ECG, serum NT-proBNP level >80 pg/mL had 100 percent sensitivity and 24 percent specificity. Taken in combination with the decision algorithm, and in patients with echocardiographic estimates of sPAP ≥ 36 mmHg, the presence of either RV strain on ECG or serum NT-proBNP >80 pg/mL had a sensitivity of 100 percent and specificity of 19 percent for diagnosis of PAH based on the RHC reference standard. By using this decision algorithm to exclude precapillary PH, the investigators concluded that 9 percent of referred patients with elevated sPAP by echocardiography (≥ 36 mmHg) could avoid undergoing invasive RHC. After excluding patients with RV strain, serum NT-proBNP testing would have avoided RHC in 16 percent of patients.

Fourteen studies (4 good quality, 7 fair, and 3 poor) evaluated biomarkers in patients both with and without PAH. Most studies were of natriuretic peptide (serum NT-proBNP or BNP); we found one study each for urinary cGMP, ADMA, and plasma endothelin-1 (ET-1). Sensitivity and specificity estimates associated with natriuretic peptide among four studies that permitted their calculation were highly variable, presumably

reflecting differences in study populations because differences in test thresholds did not result in the expected direction of change in sensitivity and specificity. The remaining 10 studies reported statistically significant correlation coefficients between natriuretic peptide levels and hemodynamic measures CO, mPAP, PVR, and sPAP.

Nineteen studies (6 good, 10 fair, 3 poor) reported the diagnostic accuracy of echocardiographic estimates of pulmonary pressures based on TRV measurement, with or without estimate of RAP, compared with a reference standard diagnosis based on RHC. Summary estimates for sensitivity (0.90; 95% CI, 0.80 to 0.96) and specificity (0.87; 95% CI, 0.80 to 0.92) showed moderate heterogeneity ($I^2=61.9\%$). Studies with lower prevalence of PH (less than 15% of study subjects) showed greater homogeneity than studies with higher prevalence of PH (sensitivity 0.84 [95% CI, 0.72 to 0.91]; specificity 0.84 [95% CI, 0.72 to 0.91]). The 10 low-prevalence studies (sensitivity 0.91 [95% CI, 0.85 to 0.94]; specificity 0.91 [95% CI, 0.85 to 0.94]) included 4 studies of liver transplant patients (which had complete verification of test-negative subjects) and 6 studies that had high degrees of verification bias.

Seven studies (3 good, 3 fair, 1 poor) evaluated the echocardiographic estimation of PVR using TRV/VTI_{RVOT} against RHC diagnosis of elevated PVR. Three of these studies included patients with known PH. Two studies used a threshold for PVR much higher than that used for diagnosis (8 Wood units vs. 2 Wood units), with the goal of distinguishing more severe PAH; these studies also used a higher test threshold of 0.2 and 0.38 compared with 0.14 to 0.175. Sensitivity ranged from 57 to 94 percent, while specificity ranged from 57 to 100 percent. Because of clinical heterogeneity no meta-analysis was performed.

Six studies correlated TRV/VTI_{RVOT} with PVR by RHC. Correlation coefficients indicated strong correlation ranging from 0.73 to 0.84, with bias ranging from 0 to 6.1, and standard deviations ranging from 1.9 to 4.3 Wood units.

We found no studies describing the safety (or harms) of echocardiography or biomarker testing.

KQ 2: Management of PAH

Key points from the Results chapter are:

- No data are available regarding the comparative effectiveness of echocardiography versus biomarkers or echocardiography versus echocardiography plus biomarkers with respect to the management of PAH or patient outcomes (insufficient strength of evidence).

Table B. Summary of strength of evidence and effect estimates for echocardiography versus echocardiography plus biomarkers as screening modalities for PAH (KQ 1)^a

Test	Sensitivity	Specificity	Correlation With RHC
Echo sPAP with NT-proBNP vs. Echo sPAP in symptomatic patients	SOE = Insufficient (1 study, 121 patients) NT-proBNP >80 pg/mL has a low false negative rate compared with RHC reference standard; the serial testing study design did not allow for NT-proBNP testing to improve sensitivity beyond that of echo sPAP alone.	SOE = Low (1 study, 121 patients) NT-proBNP ≤80 pg/mL ruled out PAH in 9–16% of patients with elevated echo sPAP ≥36 mmHg.	SOE = Insufficient (No studies)
Echo sPAP with NT-proBNP vs. Echo sPAP in asymptomatic patients	SOE = Insufficient (No studies)	SOE = Insufficient (No studies)	SOE = Insufficient (No studies)
NT-proBNP compared with RHC	SOE = Low (3 studies, 198 patients) NT-proBNP has variable sensitivity (range, 56% to 100%) for diagnosing PAH; uncertain performance for ruling out PAH.	SOE = Low (3 studies, 198 patients) NT-proBNP has variable specificity (range, 24% to 95%); uncertain performance for ruling in PAH.	SOE = Moderate (3 studies, 176 patients) Correlation of NT-proBNP and RHC is only moderate (range, 0.43 to 0.72).
TRV/TG/sPAP compared with RHC	SOE = Moderate (19 studies, 2,459 patients) Echocardiographic estimate of sPAP showed variable sensitivity ranging from 58% to 100%, with lower prevalence studies finding higher sensitivity.	SOE = Moderate (19 studies, 2,459 patients) Echocardiographic estimate of sPAP showed variable specificity ranging from 50% to 98%, with lower prevalence studies finding higher specificity.	SOE = Moderate (23 studies, 4,217 patients) Echocardiographic estimates of sPAP showed moderate to strong correlation (range, 0.38 to 0.96) with RHC and were on average unbiased, but were limited by imprecision and by a significant minority of patients in whom TRV was not measurable.
TRV/VTI_{RVOT} compared with RHC	SOE = Moderate (6 studies, 196 patients) Echocardiographic estimate of PVR showed reasonably high sensitivity (range, 89% to 100%) for ruling in PAH.	SOE = Moderate (6 studies, 196 patients) Echocardiographic estimate of PVR showed variable specificity (range, 50% to 97%), with better specificity in lower prevalence studies (range, 94% to 97%).	SOE = High (6 studies, 196 patients) Showed strong correlation between echocardiographic estimates of PVR and PVR by RHC (range, 0.74 to 0.84).

NT-proBNP = N-terminal pro-B-type natriuretic peptide; PVR = pulmonary vascular resistance; RHC = right heart catheterization; SOE = strength of evidence; sPAP = systolic pulmonary artery pressure; TRV = tricuspid regurgitant jet velocity; VTI_{RVOT} = velocity-time integral of right ventricular outflow tract

^aDarker background indicates insufficient strength of evidence.

- sPAP estimated by echocardiography shows good correlation with sPAP from RHC (low strength of evidence).
- BNP level shows moderate correlation with these RHC measures: mPAP (moderate strength of evidence), PVR (low strength of evidence), RAP (moderate strength of evidence), cardiac index (low strength of evidence), and clinical outcomes such as the 6MWD test (moderate strength of evidence).
- BNP level shows poor correlation with RHC pulmonary capillary wedge pressure (PCWP) (low strength of evidence).
- BNP level alone is not an accurate surrogate marker for disease severity (high strength of evidence).
- Increase in level of log-transformed BNP is a strong predictor of mortality (moderate strength of evidence).
- Presence of pericardial effusion is a strong predictor of mortality, although there was wide variability in results for this measure (moderate strength of evidence).
- Right atrial (RA) size correlates with increased risk of mortality (moderate strength of evidence).
- FAC is a poor predictor of mortality, but results are variable across studies (moderate strength of evidence).
- Serum uric acid level appears to predict mortality (low strength of evidence).
- TAPSE has inconsistent association with mortality (insufficient strength of evidence).
- We found no studies addressing diagnostic thinking efficacy, therapeutic efficacy, or harms (insufficient strength of evidence).

We identified 99 unique observational studies, involving a total of 8,655 patients, that evaluated the use of biomarkers or echocardiographic parameters in the management of PAH or as predictors of patient outcomes. Of these studies, 68 were rated good quality, 29 fair quality, and 2 poor quality. We did not find any studies that assessed the comparative effectiveness of echocardiography versus biomarkers, or echocardiography versus echocardiography plus biomarkers, as outlined in our original KQ. Instead, we focus on available studies that evaluated the ability of echocardiography or biomarkers to assess the severity of PAH, to predict events such as lung transplantation or death, or to assess a patient's response to therapy. By evaluating the independent association of biomarkers or echocardiography, one can impute the comparative effectiveness via indirect comparison. The most common

biomarker evaluated was BNP (59 studies), followed by uric acid (9), endothelin-1 (6), troponin T (4), nitric oxide (2), cGMP (2) and ANP (1). We found no studies assessing D-dimer or asymmetric dimethylarginine to evaluate their ability to assess severity of disease, response to therapy, or outcome.

Thirty-nine studies evaluated several echocardiographic parameters. These included sPAP (17 studies), RIMP/MPI/Tei (14), RA size (11), pericardial effusion (11), RV size (9), FAC (8), mPAP (8), TAPSE (6), TR jet (4), TRV/VTI_{RVOT} (3), RVEF (2), echocardiography-derived cardiac index (2), and RVSP (2).

For the comparators, we focused on RHC hemodynamics, 6MWD, and functional class (FC) as the reference standards for assessing severity of disease. Thirty-four studies used RHC as a reference test, 15 studies used 6MWD as a reference test, and 10 studies used FC as a reference test.

Thirty-nine studies looked at correlation between biomarkers and/or echocardiographic parameters and the comparators. Twenty-three studies evaluated hazard ratios (HR) for death, two studies evaluated HR for a composite outcome of death or lung transplant, and one study evaluated HR for lung transplant alone. Twenty-three studies evaluated changes in mean values in response to therapy, and four studies evaluated changes in median values in response to therapy. Eight studies assessed mean or median change from baseline in response to therapy.

In studies evaluating correlation of the above measures with RHC measures or a commonly used measure of disease severity (6MWD) studies were too underpowered to give reliable results. However, by combining studies looking at the same parameters and performing a meta-analysis we were able to increase the power for seven different comparisons: (1) BNP versus RHC-mPAP, (2) BNP versus RHC-PVR, (3) BNP versus RHC-CI, (4) BNP versus RHC-RAP, (5) BNP versus RHC-PCWP, (6) BNP versus 6MWD, and (7) echocardiography-derived sPAP versus RHC-sPAP. BNP showed moderate correlation with most RHC measures (mPAP, PVR, cardiac index, RAP) and clinical measures of disease severity (6MWD) and showed weak correlation with PCWP. Most effect estimates were precise (mPAP, PVR, cardiac index, RAP, 6MWD), but estimates for PCWP were imprecise, making it difficult to interpret the clinical importance of the findings for this measure. For the other measures, correlation with BNP was only moderate, indicating that BNP levels alone could not serve as an accurate surrogate marker for disease severity. Echocardiography-derived

sPAP showed strong correlation with RHC-sPAP, although there was a great deal of heterogeneity among these studies and only moderate strength of evidence to support the use of this measure.

In studies evaluating the ability of biomarkers or echocardiographic measures to predict mortality, we were able to perform a meta-analysis on six measures: BNP, pericardial effusion, RA size, FAC, uric acid and TAPSE. BNP level and pericardial effusion were strong predictors of mortality. RA size was also predictive of mortality. Data on uric acid suggested an association with mortality, while fractional area change (FAC) showed uncertain association with mortality.

The strength of evidence ratings for the most commonly reported biomarkers and echocardiographic parameters are

summarized in Table C (management of PAH) and Table D (prediction of patient outcomes).

KQ 3: Pharmacotherapy for Pulmonary Arterial Hypertension

Key Points from the Results chapter are:

- In patients who have been receiving monotherapy, combination therapy appears to be moderately more effective than continuation of monotherapy for improving 6-minute walk distance (6MWD), with a magnitude of effect that is approximately equal to the estimated minimal important difference (MID) for PAH, of 6MWD of 33 meters (low strength of evidence).

Table C. Summary of strength of evidence and effect estimates for the use of echocardiography or biomarkers in the management of PAH (KQ 2)

Comparison	Number of Studies (Patients)	Summary Correlation Coefficient Estimate (95% CI)	SOE and Findings
BNP compared with RHC-mPAP	14 (606)	0.39 (0.31 to 0.47)	SOE = Moderate Serum BNP level shows moderate correlation with mPAP.
BNP compared with RHC-PVR	13 (684)	0.46 (0.31 to 0.59)	SOE = Low Serum BNP level shows moderate correlation with PVR.
BNP compared with RHC-RAP	12 (645)	0.47 (0.40 to 0.54)	SOE = Moderate Serum BNP level shows moderate correlation with RAP.
BNP compared with RHC-CI	10 (550)	-0.42 (-0.54 to -0.28)	SOE = Low Serum BNP level shows negative moderate correlation with cardiac index.
BNP compared with RHC-PCWP	5 (319)	0.16 (0.01 to 0.31)	SOE = Low Serum BNP level shows poor correlation with PCWP.
BNP compared with 6MWD (absolute)	9 (484)	-0.46 (-0.55 to -0.35)	SOE = Moderate Serum BNP level shows negative moderate correlation with 6MWD.
Echocardiography-derived sPAP compared with RHC-sPAP	9 (362)	0.76 (0.53 to 0.89)	SOE = Low sPAP estimated by echocardiography shows good correlation with sPAP from RHC.

6MWD = 6-minute walk distance; BNP = B-type natriuretic peptide; CI = confidence interval; FAC = fractional area change; PCWP = pulmonary capillary wedge pressure; PVR = pulmonary vascular resistance; RA = right atrium; RAP = right atrial pressure; RHC = right heart catheterization; SOE = strength of evidence; sPAP=systolic pulmonary artery pressure

Table D. Summary of strength of evidence and effect estimates for the use of echocardiography or biomarkers in the prediction of mortality (KQ 2)^a

Marker	Number of Studies (Patients)	Summary Hazard Ratio Estimate (95% CI)	SOE and Findings
BNP	6 (407)	2.42 (1.72 to 3.41)	SOE = Moderate Increase in log-transformed BNP level is a good predictor of mortality.
Pericardial effusion	8 (2,590)	2.43 (1.57 to 3.77)	SOE = Moderate Presence of pericardial effusion is a strong predictor of mortality, although there was wide variability in results for this measure.
RA size	4 (242)	1.06 (1.01 to 1.10)	SOE = Moderate RA size is a predictor of mortality.
FAC	4 (242)	0.98 (0.96 to 1.01)	SOE = Moderate FAC is a poor predictor of mortality.
Uric acid	4 (246)	1.01 (1.00 to 1.01)	SOE = Low Small increase in mortality but imprecision of estimates limit these data.
TAPSE	4 (251)	0.94 (0.82 to 1.08)	SOE = Insufficient Inconsistent results between studies lead to uncertainty.

BNP = B-type natriuretic peptide; CI = confidence interval; FAC = fractional area change; RA = right atrium; RAP = right atrial pressure; SOE = strength of evidence; TAPSE = tricuspid annular plane systolic excursion

^aDarker background indicates insufficient strength of evidence.

- We did not identify any eligible studies that evaluated the comparative effectiveness of calcium channel blockers on intermediate-term and long-term patient outcomes, or that randomized treatment-naïve patients to monotherapy versus combination therapy, or that directly compared two drug classes.
- Although we did not intend to exclude studies of children, the inclusion criterion requiring reporting intermediate-term and long-term patient outcomes had the effect of eliminating randomized clinical trials of children with PAH.
- Prostanoids were associated with lower mortality when compared with standard therapy or placebo (low strength of evidence). Current evidence is inconclusive regarding a reduction in mortality associated with treatment with endothelin antagonists or phosphodiesterase inhibitors (insufficient strength of evidence).
- Endothelin antagonists, phosphodiesterase inhibitors, and prostanoids were all associated with improved 6MWD after 8 to 16 weeks of therapy, with a magnitude of effect that is approximately equal to the estimated minimal important difference (MID) for PAH of 6MWD of 33 meters (moderate strength of evidence).
- Endothelin antagonists and phosphodiesterase inhibitors were associated with lower incidence of hospitalization when compared with standard therapy or placebo (moderate strength of evidence). Current evidence is inconclusive regarding a reduction in hospitalization associated with treatment with prostanoids (insufficient strength of evidence).
- Endothelin antagonists, phosphodiesterase inhibitors, and prostanoids were associated with statistically significant improvements in most or all hemodynamic measures such as PVR, mPAP, and cardiac index (low strength of evidence), compared with placebo or standard therapy. The clinical significance of the magnitude of the observed changes in these intermediate outcomes is unclear.
- Among commonly reported adverse events, there was a higher incidence of jaw pain associated with

aerosolized prostanoid treatment compared with placebo (high strength of evidence) and cough associated with aerosolized prostanoids versus placebo (high strength of evidence). In addition, headache was associated with phosphodiesterase inhibitors compared with placebo or standard therapy (moderate strength of evidence), and flushing was associated with phosphodiesterase inhibitors (moderate strength of evidence) and aerosolized prostanoids (moderate strength of evidence), compared with placebo or standard therapy.

Twenty-eight RCTs involving 3,613 patients evaluated the comparative effectiveness and safety of monotherapy or combination therapy for PAH. Of these RCTs, 18 (64%) were rated good quality, 9 (32%) fair quality, and 1 (4%) poor quality. Nineteen studies (68%) were funded by industry, one by private foundation, one by government and private funding, one by private and industry funding, one by industry and “other” funding, and five did not report funding sources.

The mean patient ages ranged from 28 to 50 years old. Twenty studies enrolled patients with PAH, four studies enrolled patients with PAH associated with systemic sclerosis (formerly scleroderma), and two studies enrolled patients with Eisenmenger syndrome. Two studies enrolled a minority of patients with PH other than PAH: one included patients with chronic thromboembolic PH (28%), and another included patients with PH owing either to lung disease or to chronic thromboembolic PH (37%).

Twenty-one studies compared a single drug (monotherapy) with placebo or standard therapy and included the following drugs: bosentan (6 studies), sildenafil (2), iloprost (2), epoprostenol (3), tadalafil (3), ambrisentan (2), treprostinil (3), and vardenafil (1). For the purposes of this analysis, the standard therapy arms were grouped with the placebo arms. Standard therapies included supportive therapy (diuretics, oxygen, digoxin, oral anticoagulants) with or without calcium channel blockers, but not including newer specific vasodilator medications. One study was a head-to-head comparison of bosentan and sildenafil. The remaining five studies compared combination therapy with monotherapy: (1) intravenous (IV) epoprostenol plus bosentan versus IV epoprostenol plus placebo, (2) sildenafil plus IV epoprostenol versus IV epoprostenol plus placebo, (3) bosentan plus aerosolized iloprost versus bosentan, (4) bosentan plus aerosolized iloprost versus bosentan plus placebo, and (5) aerosolized treprostinil plus bosentan or sildenafil versus bosentan or sildenafil plus placebo. We did not identify any eligible studies published after 1990 that evaluated the safety or

efficacy of calcium channel blockers on intermediate-term and long-term patient outcomes.

Most studies (85%) were multicenter trials; three were single-center trials, and four did not report the number of centers. The studies reported the following outcomes: 6MWD (27 studies), mortality (21), dyspnea (17), right heart catheterization indices (18), functional class (13), hospitalization for worsening PAH (10), quality of life (11), lung transplantation (5), right heart failure or right ventricular dysfunction (4), and brain natriuretic peptide (4). Twenty-one studies reported harms or adverse events. Table E summarizes the strength of evidence ratings for the key outcomes of mortality, 6MWD, and hospitalization. Details about the specific components of these ratings (risk of bias, consistency, directness, and precision) and information on other outcomes are available in the full report.

Discussion

Key Findings and Strength of Evidence

A single study compared the combination of biomarker tests and echocardiography with echocardiography alone to screen for PAH (KKQ 1). This good-quality prospective cohort study of 372 patients suggested that biomarker testing with NT-proBNP may be useful in ruling out PAH among those suspected of PH who also have elevated sPAP by echocardiography;²⁴ however, this finding is limited by the lack of replication, small sample size (wide confidence limits) and confounding with RV strain on ECG. No data are available regarding combined echocardiography and biomarker screening in asymptomatic patients at high risk for PAH. In the absence of other direct comparative trials, we attempted to address this question by evaluating the efficacy of biomarker and echocardiography independently for screening and diagnosis of PAH. We reviewed 60 studies involving 7,096 patients that evaluated biomarker tests, echocardiography, or both, to screen for PAH. The associations between natriuretic peptide testing and PAH diagnosis is insufficiently strong to support its use alone as a screening test in either asymptomatic or symptomatic patients suspected of PAH. Data on biomarker testing were essentially limited to a single test—NT-proBNP—which showed moderate correlation with RHC hemodynamic measures and a great deal of variability between studies in its diagnostic accuracy and discrimination.

We found that echocardiography estimates of pulmonary artery pressures (sPAP, TG, and TRV) and pulmonary vascular resistance (TRV/VTI^{RVOT}) demonstrated good

Table E. Summary of strength of evidence and effect estimates for monotherapy versus combination therapy for PAH (KQ 3)^a

Intervention	Mortality	6MWD (m)	Hospitalization
Endothelin antagonist vs. placebo	SOE = Insufficient (6 studies, 838 patients) Inconclusive benefit (few studies, few deaths lead to wide CI) OR 0.60 (95% CI, 0.23 to 1.59)	SOE = Moderate (6 studies, 663 patients) Improved 6MWD with endothelin antagonists compared with placebo Mean difference 39.9 (95% CI, 21.4 to 58.4)	SOE = Moderate (3 studies, 606 patients) Reduced risk of hospitalization OR 0.34 (95% CI, 0.17 to 0.69)
Phosphodiesterase inhibitors vs. placebo	SOE = Insufficient (4 studies, 1,011 patients) Inconclusive benefit (few studies, few deaths lead to wide CI) OR 0.30 (95% CI, 0.08 to 1.11)	SOE = Moderate (4 studies, 991 patients) Improved 6MWD with PDE5 therapy compared with placebo or standard therapy Mean difference 38.9 (95% CI, 22.0 to 55.9)	SOE = Moderate (4 studies, 1,011 patients) Reduced risk of hospitalization OR 0.48 (95% CI, 0.25 to 0.91)
Prostanoids vs. placebo or standard therapy	SOE = Low (8 studies, 1,229 patients) Lower mortality with prostanoids, but inconsistent results and wide confidence intervals OR 0.52 (95% CI, 0.29 to 0.95)	SOE = Moderate (7 studies, 933 patients) Improved 6MWD with prostanoid therapy compared with placebo Mean difference 27.9 (95% CI, 10.3 to 45.4)	SOE = Insufficient (2 studies, 301 patients) Inconclusive benefit (few studies, wide CI) OR 0.42 (95% CI, 0.06 to 3.08)
Combination vs. monotherapy	SOE = Insufficient (3 studies, 566 patients) Inconclusive benefit (few studies, few deaths lead to wide CI) OR 0.37 (95% CI, 0.04 to 3.32)	SOE = Low (3 studies, 363 patients) Improved 6MWD with combination therapy compared with monotherapy Mean difference 23.9 (95% CI, 8.0 to 39.9)	SOE = Insufficient (3 studies, 566 patients) Inconclusive benefit (few studies, wide CI) OR 0.64 (95% CI, 0.31 to 1.36)

6MWD = 6-minute walk distance; CI = confidence interval; NS = not statistically significant; OR = odds ratio; SOE = strength of evidence

^aDarker background indicates insufficient strength of evidence.

accuracy in screening for PAH. In low-prevalence populations (<10%), negative predictive value of a normal sPAP is high, suggesting that echocardiography with a low threshold may be an appropriate test in asymptomatic high-risk populations or in patients with symptoms suggesting PAH. (This is shown in studies of liver transplant patients with complete verification).

Our findings suggest that echocardiographic estimation of sPAP is sufficiently accurate to justify its role in screening for PAH in symptomatic patients suspected of having PH. However, this conclusion has several important caveats. First, echocardiography in a small but significant number of patients may not produce an estimate of sPAP because

of poor-quality Doppler visualization of the tricuspid regurgitant jet. Second, echocardiographic estimates of sPAP often over- or under-estimate pulmonary artery pressure enough to result in misclassification according to PAH diagnostic threshold—hence the selection of a test threshold is critical for the aim of screening. A single test threshold is insufficient to perform with simultaneously high sensitivity and specificity (or simultaneously high positive and negative predictive values), especially in populations with higher risk or higher prevalence (more symptomatic), where echocardiography cannot be relied upon to exclude pulmonary hypertension if pretest probability is high. In asymptomatic patients at high risk

for PH, echocardiography seems to perform with similar sensitivity and specificity; however, these studies suffer from verification bias, which likely inflates both the sensitivity and specificity estimates. For example, consider two prospective studies that show that approximately 10 percent of asymptomatic patients with systemic sclerosis and normal sPAP develop PH when serially retested with echocardiography. These findings are consistent with either misclassification at baseline echocardiographic screening or prospective development of PH. This ambiguity suggests that if echocardiographic screening of asymptomatic patients with a high-risk diagnosis were to be undertaken, then serial testing would be necessary.

We reviewed 99 studies, involving 8,655 patients, that evaluated biomarker tests or echocardiography to diagnose and follow progression of disease as well as response to therapy for PAH (KQ 2). Our review found that BNP showed only moderate correlation with most RHC measures (mPAP, PVR, cardiac index, RAP) and clinical measures of disease severity (6MWD) and showed weak correlation with PCWP. Most effect estimates were precise (mPAP, PVR, cardiac index, RAP, 6MWD), but estimates for PCWP were imprecise, making it difficult to interpret the clinical importance of the findings for this measure. For the other measures, correlation with BNP was moderate, indicating that BNP levels alone could not serve as an accurate surrogate marker for disease severity. Alternatively, echocardiography-derived sPAP showed strong correlation with RHC-sPAP with a precise effect estimate, and may be useful as an alternative to RHC to assess disease severity. However, there was a great deal of heterogeneity among these studies.

BNP level and the presence of pericardial effusion were predictors of mortality and may be useful clinically, though results were not highly precise. RA size and uric acid were also associated with mortality, but studies were less consistent than for BNP. FAC showed no significant ability to predict mortality; data on TAPSE were too inconsistent to be conclusive.

Our findings do not support any recommendations for replacing existing measurement tools to assess disease severity, prognosis, or response to therapy. Echocardiography-derived sPAP shows promise as a possible surrogate marker for RHC-sPAP, but it is unclear whether or not this measure alone is adequate to assess disease severity, prognosis, or response to therapy.

We reviewed 37 studies involving 4,192 patients that assess the effectiveness of drug treatment for PAH in

adults. Our review found inconclusive evidence regarding mortality reduction for 11 of the 12 drug treatment comparisons: (1) ambrisentan versus placebo (OR 0.40; 95% CI, 0.10 to 1.51), (2) bosentan versus placebo (OR 0.72; CI, 0.14 to 3.60), (3) epoprostenol versus placebo or standard therapy (OR 0.33; CI, 0.07 to 1.50), (4) iloprost versus placebo (OR 0.43; CI, 0.08 to 2.47), (5) sildenafil versus placebo (OR 1.01; CI, 0.10 to 9.92), (6) tadalafil versus placebo (OR 0.50; CI, 0.05 to 5.63), (7) treprostinil versus placebo (OR 0.50; CI, 0.12 to 2.12), (8) vardenafil versus placebo (OR 0.08; CI, 0.00 to 1.82), (9) endothelin antagonists versus placebo (OR 0.60; CI, 0.23 to 1.59), (10) phosphodiesterase inhibitors versus placebo (OR 0.30; CI, 0.08 to 1.11), and (11) combination therapy versus monotherapy (OR 0.37; CI, 0.04 to 3.32).

Few deaths were observed in these limited-duration studies, leading to wide confidence intervals and lack of statistical power to detect a difference in mortality; however, a consistent direction of effect and demonstrated improvements in other outcomes, including functional and hemodynamic measures, support that a mortality reduction might exist.

Increases in 6MWD ranging from 27.9 meters (95% CI, 10.3 to 45.4) to 39.9 meters (CI, 21.4 to 58.4) were observed in trials of all drug classes when compared with placebo or standard therapy; however, comparisons between agents are inconclusive. The magnitude of these statistically significant improvements in 6MWD associated with treatment are very close to a recently published estimate of 33 meters for the minimal important difference for the 6MWD in patients with PAH.²⁵ Combination therapy in patients already on monotherapy also showed improved 6MWD compared with continuation of monotherapy (OR 23.9; CI, 8.0 to 39.9), but the diversity of treatment regimens and the small number of combination therapy trials again make comparisons between specific regimens inconclusive. In studies evaluating hospitalization, endothelin receptor antagonists and phosphodiesterase-5 inhibitor treatment was associated with lower odds of hospitalization compared with placebo (OR 0.34 and 0.48, respectively). The magnitude of the odds ratio associated with prostanoids was similar (OR 0.42), but the 95% confidence interval included 1.0, thereby making this finding not statistically significant. Combination therapy compared with monotherapy also showed a similar nonsignificant effect on hospitalizations (OR 0.64). Endothelin antagonists, phosphodiesterase inhibitors, and prostanoids each had favorable effects on most hemodynamic outcomes including cardiac index, mPAP, and PVR.

In studies reporting adverse effects, we found that phosphodiesterase-5 inhibitors were more likely than endothelin receptor antagonists to cause headache, and endothelin antagonists still were more likely than placebo to cause headache. Drugs did not significantly differ in their odds of causing dizziness or diarrhea. Aerosolized prostanoids were much more likely to cause jaw pain and cough compared with placebo. Phosphodiesterase-5 inhibitors and prostanoids were associated with flushing, while data on endothelin receptor antagonists were inconclusive. Phosphodiesterase-5 inhibitors had a significant association with peripheral edema, while data on prostanoids and endothelin receptor antagonists were inconclusive.

The findings from our meta-analyses of the few studies that compared combination therapy with monotherapy suggest, but do not prove, that combination therapy confers more benefit than does monotherapy in the treatment of PAH. These findings are generally consistent with the European Society of Cardiology (ESC)/European Respiratory Society (ERS) guideline recommendation for monotherapy as initial treatment, with combination treatment reserved for patients who have an inadequate clinical response to monotherapy.

Applicability

The principal limitations to applicability of data on the diagnosis of PAH all relate to the patient populations studied. First, the studies may not be applicable to the screening of asymptomatic patients. None of the study populations consisted entirely of asymptomatic patients, and although many studies included some patients without symptoms, they were not reported separately in terms of outcomes. Some studies of populations in whom PAH was suspected failed to adequately describe the basis for a clinical suspicion of PAH, whether symptoms of dyspnea, clinical signs, or other test results, such as diffusion capacity of the lung for carbon monoxide (DLCO), thus the applicability of these studies for screening symptomatic patients was also limited.

A second kind of limitation resulted from the fact that the spectrum of disease among study populations was often skewed, particularly in case-control studies, by selection criteria that selected from patients with known PAH (cases) and patients known not to have PAH (controls). Such studies usually excluded participants with other conditions that might be confused with PAH such as PH due to left-sided heart failure, thrombotic disease, or chronic obstructive pulmonary disease.

A third limitation was that participants in many studies had a wide range of disease severity, particularly those cases in case-control design studies, making these studies a poor match for the question at hand. Other applicability issues identified in the KQ 1 studies were less frequent and judged to be less severe.

Our findings in KQ 2 assessing the prognostic or predictive value of biomarkers and echocardiography may not be applicable to all PAH populations. The greatest concern is that studies reviewed in KQ 2 included participants at widely differing points in the natural history of disease, who had widely differing degrees of disease severity and different underlying etiologies of PAH. There was also concern that the population was not adequately described to assess applicability, included patients with conditions other than PAH, or in general did not match the review question. Applicability may also be limited by the use of surrogate markers that may not be clinically relevant; also by insufficient followup time. In a few studies, it was also felt that the intervention arm or cointerventions did not adequately reflect current clinical practice or that the study setting was widely divergent from the current typical U.S. setting. Finally, there is concern that some studies did not provide adequate information about adverse events.

Applicability considerations were somewhat different for KQ 3 than for the KQs about screening and management of PAH. Most of the studies included in this review for KQ 3 were RCTs with generally good internal validity. Patient populations, however, differed between studies; variation in eligibility criteria resulted in differences between study populations in severity of illness, underlying etiology of PAH, comorbid conditions, and prior and concurrent treatment. Many different countries were represented, thereby introducing potential differences in clinical practice and care delivery settings relative to current practice in typical settings in the United States. There was also concern that the population was not always adequately described to assess applicability, with few studies exploring potential differences in response to treatment among different patient subgroups. Finally, the studies that compared combination therapy with monotherapy were all of similar design, randomizing patients who had previously received monotherapy to either continued monotherapy with that drug or continued therapy with that drug plus the addition of a second drug. While we considered these studies to represent a comparison of combination therapy with monotherapy, we do so with the understanding that this study design does not address the question of whether initiating two drugs

is superior to initiating a single drug to treatment-naïve patients.

Research Gaps

The available evidence leaves numerous gaps and areas for potential future research. We used the framework recommended by Robinson et al.²⁶ to identify gaps in evidence and describe why these gaps exist. Results are as follows:

KQ 1: Screening for PAH

- Patients at elevated risk for PAH, other than those with systemic sclerosis, have seldom been studied in screening test studies.
 - Consider cohort studies of testing for PH among high-risk populations other than those with systemic sclerosis; including patients with HIV, sickle cell anemia or trait portal hypertension, family history of PAH, or catecholaminergic drug use.
 - Different populations may have different risks of PAH and different benefits from screening; in studies where heterogeneous populations are included, the effectiveness of screening should be examined according to risk factor.
- Relatively few data exist on screening of asymptomatic patients with a combination of echocardiography and biomarker testing.
 - Consider cohort studies that apply echocardiography and biomarker screening in a coordinated or algorithmic way, and studies that verify diagnosis in at least a sample of test-negative patients by RHC or lengthy followup.
 - Future tests of the added value of biomarkers should use well validated echocardiography parameters as a screening test, including estimates of pulmonary artery pressures (sPAP, TG, and TRV) and pulmonary vascular resistance (TRV/VTI_{RVOT}).
- Studies of echocardiography for diagnosis of PH have focused on the association of single measures or parameters at a time rather than an integrated diagnostic assessment based on an entire examination and multiple echocardiographic measures or parameters.
 - Consider studies that evaluate a global echocardiographic assessment based not only on sPAP but also on right heart chamber size wall thickness and function, estimated PVR, and left heart measures.

- Consider further development of data on the use of echocardiography to measure exercise response to sPAP.
- Consider further development of echocardiographic estimation of mPAP, which would better align with the diagnostic criteria for PAH.
- Consider studies of additional promising measures such as end diastolic pulmonary regurgitation gradient, mean tricuspid regurgitation gradient, and Doppler tissue imaging of the tricuspid annulus.

KQ 2: Management of PAH

- Echocardiographically guided and BNP-guided treatment strategies have not been explicitly tested.
 - Consider cohort studies evaluating prognosis, as well as treatment trials examining association of baseline echocardiographic parameters and BNP levels with response to treatment.
- Other imaging modalities, such as magnetic resonance imaging, have been little studied as alternative noninvasive tests to assess RV function.
- Cardiopulmonary exercise testing and exercise echocardiography have yielded relatively few data, and their clinical utility and relationship to PH diagnostic criteria are uncertain.
 - Consider validation studies to demonstrate prognostic value, particularly for patients with normal resting echocardiography but abnormal exercise echocardiography.

KQ 3: Pharmacotherapy for PAH

- Relatively few data exist on the efficacy of treating PAH early in the disease course (WHO functional class I-II).
 - Improved data on efficacy of early PAH treatment would strengthen linkage to data on efficacy of screening testing.
 - Consider treatment trials in early-stage PAH, particularly among patients identified by case finding or screening interventions.
- Relatively few data exist on children with persistent PH or congenital heart disease.
 - Consider controlled trials in children.
- Few treatment trials address direct comparison of alternative drug treatments, particularly for PAH patients early in the disease course.

- Consider trials designed to compare clinical alternative treatments to permit more evidence-based treatment selection, such as head-to-head treatment comparisons rather than placebo-control, or combination versus monotherapy trials.
- The majority of RCTs have been too short and small to generate definitive data on major patient-centered outcomes. Although surrogate markers have limitations, more complete collection, analysis, and correlation of these markers with patient-centered outcomes may not only help to validate surrogate outcomes but also provide more practical outcome measures.
 - Consider including biomarker and imaging techniques with conventional clinical outcomes to improve data on validity and responsiveness of surrogate outcomes.
- Few data are available from trials about differences in response to treatment based on patient characteristics.
 - Consider subgroup analysis of treatment efficacy by WHO functional class, underlying etiology, and other patient-level factors.
- Data on the efficacy of combination treatments are limited.
 - Consider more combination treatment trials, in particular trials with clear criteria for starting combination therapy, and trials in patients who have not failed monotherapy.
- The duration of controlled trial efficacy data are limited.
 - Consider, particularly for clinically relevant comparisons (e.g., head-to-head treatment or combo versus monotherapy trials), longer term followup studies that retain randomized group comparisons while assessing long-term efficacy.

Conclusions

Further research is needed to confirm the single good-quality study suggesting that echocardiography and the biomarker NT-proBNP in combination may be sufficiently accurate to rule out PAH when testing symptomatic patients. In asymptomatic populations, more research is needed to draw conclusions regarding the effectiveness for screening. BNP, RA size, the presence of pericardial effusion and uric acid had prognostic value in patients with PAH, but other echocardiographic parameters and biomarkers either were not predictive or had insufficient

data. Although no treatments demonstrate a strong and consistent mortality reduction, many are associated with improved 6MWD and reduced hospitalization rates. Comparisons of different drug combinations are inconclusive regarding mortality reduction but suggest an improvement in 6MWD compared with continuation of monotherapy.

References

1. Galie N, Torbicki A, Barst R, et al. Guidelines on diagnosis and treatment of pulmonary arterial hypertension. The Task Force on Diagnosis and Treatment of Pulmonary Arterial Hypertension of the European Society of Cardiology. *Eur Heart J*. 2004;25(24):2243-78. PMID: 15589643.
2. Badesch DB, Abman SH, Simonneau G, et al. Medical therapy for pulmonary arterial hypertension: updated ACCP evidence-based clinical practice guidelines. *Chest*. 2007;131(6):1917-28. PMID: 17565025.
3. Simonneau G, Robbins IM, Beghetti M, et al. Updated clinical classification of pulmonary hypertension. *J Am Coll Cardiol*. 2009;54(1 Suppl):S43-54. PMID: 19555858.
4. D'Alonzo GE, Barst RJ, Ayres SM, et al. Survival in patients with primary pulmonary hypertension. Results from a national prospective registry. *Ann Intern Med*. 1991;115(5):343-9. PMID: 1863023.
5. Hoeper MM, Lee SH, Voswinckel R, et al. Complications of right heart catheterization procedures in patients with pulmonary hypertension in experienced centers. *J Am Coll Cardiol*. 2006;48(12):2546-52. PMID: 17174196.
6. Benza RL, Miller DP, Gomberg-Maitland M, et al. Predicting survival in pulmonary arterial hypertension: insights from the Registry to Evaluate Early and Long-Term Pulmonary Arterial Hypertension Disease Management (REVEAL). *Circulation*. 2010;122(2):164-72. PMID: 20585012.
7. Cacciapuoti F. Echocardiographic evaluation of right heart function and pulmonary vascular bed. *Int J Cardiovasc Imaging*. 2009;25(7):689-97. PMID: 19634000.
8. Barst RJ, McGoon M, Torbicki A, et al. Diagnosis and differential assessment of pulmonary arterial hypertension. *J Am Coll Cardiol*. 2004;43(12 Suppl S):40S-47S. PMID: 15194177.
9. Farber HW, Foreman AJ, Miller DP, et al. REVEAL Registry: correlation of right heart catheterization and echocardiography in patients with pulmonary arterial hypertension. *Congest Heart Fail*. 2011;17(2):56-64. PMID: 21449993.
10. Kiatchoosakun S, Wongvipaporn C, Nanagara R, et al. Right ventricular systolic pressure assessed by echocardiography: a predictive factor of mortality in patients with scleroderma. *Clin Cardiol*. 2011;34(8):488-93. PMID: 21717471.
11. Sachdev A, Villarraga HR, Frantz RP, et al. Right ventricular strain for prediction of survival in patients with pulmonary arterial hypertension. *Chest*. 2011;139(6):1299-309. PMID: 21148241.

12. Galie N, Hoepfer MM, Humbert M, et al. Guidelines for the diagnosis and treatment of pulmonary hypertension: the Task Force for the Diagnosis and Treatment of Pulmonary Hypertension of the European Society of Cardiology (ESC) and the European Respiratory Society (ERS), endorsed by the International Society of Heart and Lung Transplantation (ISHLT). *Eur Heart J*. 2009;30(20):2493-537. PMID: 19713419.
13. Bernal V, Pascual I, Esquivias P, et al. N-terminal brain natriuretic peptide as a diagnostic test in cirrhotic patients with pulmonary arterial hypertension. *Transplant Proc*. 2009;41(3):987-8. PMID: 19376405.
14. Cavagna L, Caporali R, Klersy C, et al. Comparison of brain natriuretic peptide (BNP) and NT-proBNP in screening for pulmonary arterial hypertension in patients with systemic sclerosis. *J Rheumatol*. 2010;37(10):2064-70. PMID: 20634241.
15. Aubert JD. Biochemical markers in the management of pulmonary hypertension. *Swiss Med Wkly*. 2005;135(3-4):43-9. PMID: 15729606.
16. Soon E, Holmes AM, Treacy CM, et al. Elevated levels of inflammatory cytokines predict survival in idiopathic and familial pulmonary arterial hypertension. *Circulation*. 2010;122(9):920-7. PMID: 20713898.
17. Macchia A, Marchioli R, Tognoni G, et al. Systematic review of trials using vasodilators in pulmonary arterial hypertension: why a new approach is needed. *Am Heart J*. 2010;159(2):245-57. PMID: 20152223.
18. Sitbon O, Humbert M, Jais X, et al. Long-term response to calcium channel blockers in idiopathic pulmonary arterial hypertension. *Circulation*. 2005;111(23):3105-11. PMID: 15939821.
19. McLaughlin VV, Archer SL, Badesch DB, et al. ACCF/AHA 2009 expert consensus document on pulmonary hypertension a report of the American College of Cardiology Foundation Task Force on Expert Consensus Documents and the American Heart Association developed in collaboration with the American College of Chest Physicians; American Thoracic Society, Inc.; and the Pulmonary Hypertension Association. *J Am Coll Cardiol*. 2009;53(17):1573-619. PMID: 19389575.
20. Agency for Healthcare Research and Quality. *Methods Guide for Effectiveness and Comparative Effectiveness Reviews*. Rockville, MD: Agency for Healthcare Research and Quality. www.effectivehealthcare.ahrq.gov/index.cfm/search-for-guides-reviews-and-reports/?pageaction=displayproduct&productid=318. Accessed March 19, 2012.
21. Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2: A Revised Tool for the Quality Assessment of Diagnostic Accuracy Studies. *Ann Intern Med*. 2011;155(8):529-36. PMID: 22007046.
22. Owens DK, Lohr KN, Atkins D, et al. AHRQ series paper 5: Grading the strength of a body of evidence when comparing medical interventions--Agency for Healthcare Research and Quality and the Effective Health-Care Program. *J Clin Epidemiol*. 2010;63(5):513-23. PMID: 19595577.
23. Atkins D, Chang SM, Gartlehner G, et al. Assessing applicability when comparing medical interventions: AHRQ and the Effective Health Care Program. *J Clin Epidemiol*. 2011;64(11):1198-207. PMID: 21463926.
24. Bonderman D, Wexberg P, Martischinig AM, et al. A noninvasive algorithm to exclude pre-capillary pulmonary hypertension. *Eur Respir J*. 2011;37(5):1096-103. PMID: 20693249.
25. Mathai SC, Puhan MA, Lam D, et al. The minimal important difference in the 6-minute walk test for patients with pulmonary arterial hypertension. *Am J Respir Crit Care Med*. 2012;186(5):428-33. PMID: 22723290.
26. Robinson KA, Saldanha IJ, Mckoy NA. *Frameworks for Determining Research Gaps During Systematic Reviews. Methods Future Research Needs Report No. 2*. (Prepared by the Johns Hopkins University Evidence-based Practice Center under Contract No. 290-2007-10061-I.) AHRQ Publication No. 11-EHC043-EF. Rockville, MD: Agency for Healthcare Research and Quality. June 2011. www.effectivehealthcare.ahrq.gov/reports/final.cfm. Accessed May 22, 2012.

Glossary

6MWD	6-minute walk distance
AHRQ	Agency for Healthcare Research and Quality
BID	two times per day
BNP	B-type natriuretic peptide
CI	confidence interval
CHF	congestive heart failure
COPD	chronic obstructive pulmonary disease
CTEPH	chronic thromboembolic pulmonary hypertension
CVD	collagen vascular disease
FAC	fractional area change
FC	functional class
HR	hazard ratio
HRQOL	health-related quality of life
IQR	interquartile range
KQ	Key Question
MI	myocardial infarction
mo	month/months
mPAP	mean pulmonary artery pressure
MPI	myocardial performance index
NA	not applicable
NR	not reported
NT-proBNP	N-terminal pro-B-type natriuretic peptide
OR	odds ratio
PAH	pulmonary arterial hypertension
PADP	pulmonary artery diastolic pressure
PASP	pulmonary artery systolic pressure
PCWP	pulmonary capillary wedge pressure
PH	pulmonary hypertension

PPH	primary pulmonary hypertension
PVR	pulmonary vascular resistance
QOL	quality of life
RA	right atrium
RAP	right atrial pressure
RHC	right heart catheterization
RIMP	right index of myocardial performance
RR	risk ratio
RV	right ventricle
RVEF	right ventricle ejection fraction
SD	standard deviation
SEM	standard error of the mean
SOE	strength of evidence
sPAP	systolic pulmonary artery pressure
SSc	systemic sclerosis
TAPSE	tricuspid annular plane systolic excursion
TEP	Technical Expert Panel
TRV	tricuspid regurgitant jet velocity
VSD	ventricular septal defect
VTIRVOT	velocity-time integral of right ventricular outflow tract
yr	year/years

Full Report

This executive summary is part of the following document: McCrory DC, Coeytaux RR, Schmit KM, Kraft B, Kosinski AS, Mingo AM, Vann LM, Gilstrap DL, Hargett CW, Lugogo NL, Heidenfelder BL, Posey R, Irvine RJ, Wing L, Pendergast K, Dolor RJ. Pulmonary Arterial Hypertension: Screening, Management, and Treatment. Comparative Effectiveness Review No. 117. (Prepared by the Duke Evidence-based Practice Center under Contract No. 290-2007-10066-I.) AHRQ Publication No. 13-EHC087-EF. Rockville, MD: Agency for Healthcare Research and Quality. April 2013. www.effectivehealthcare.ahrq.gov/reports/final.cfm.

