Controlling blood pressure in diabetes is more important than what agents we use

We often are not aggressive enough in controlling blood pressure for those with diabetes, despite evidence that it impacts outcomes more than glycemic control. Though there does not appear to be direct evidence that other blood pressure agents prevent renal failure in those with diabetes, it is reassuring that BP control, even when we are unable to use ACE inhibitors or ARBs, is a worthy goal.

Allen Daugird, MD
University of North Carolina, Chapel Hill

Evidence summary

Diabetic nephropathy is the leading cause of end-stage renal disease, and it occurs in 20% to 40% of patients with diabetes. Optimal glycemic (glycosylated hemoglobin [HbA1c] level <7%) and hypertension control (<130/80 mm Hg) can prevent or slow the progression of diabetic nephropathy. An average of 3 antihypertensive medications are needed to achieve currently recommended blood pressure.
goals in those with diabetes. In hypertensive and normotensive patients with type 2 diabetes and microalbuminuria, ACE inhibitors have been well studied and found to reduce the risk of mortality, major cardiovascular events, and slow the progression to overt nephropathy, in patients with diabetes and at least 1 other risk factor. In patients with type 2 diabetes and hypertension, macroalbuminuria, and serum creatinine >1.5 mg/dL, ARBs are effective in slowing the progression of diabetic nephropathy.

Some patients, however, are intolerant to ACE inhibitors and ARBs. When patients are intolerant to these medications, diuretics, NDCAs, or beta-blockers are recommended agents for the treatment of hypertension.

According to a systematic review, NDCAs cause a greater reduction in proteinuria compared with DCAs (dihydropyridine calcium antagonists, such as nifedipine and amlodipine), although there was no significant differences in lowering blood pressure. Mean change in proteinuria was +2% for DCAs and –30% for NDCAs (95% confidence interval [CI], 10%–54%; P=.01). In another RCT, amiodipine was no more effective than placebo in reducing proteinuria, while irbesartan effectively reduced end-stage renal disease (number needed to treat [NNT]=25 over 2.6 years).

In the UKPDS-Hypertension in Diabetes study (a multicenter randomized study in patients with type 2 diabetes that evaluated the effects of different levels of blood pressure control on diabetic complications), researchers found that patients assigned to the tight-control group (blood pressure goal <150/85 mm Hg) had 37% risk reduction in microvascular endpoints (nephropathy and advanced retinopathy). There was no difference in study endpoints between the ACE inhibitor captopril and the beta-blocker atenolol. Selective beta-blockers like carvedilol appear to have fewer adverse metabolic effects, although the clinical significance of this difference is unclear. In insulin-dependent patients and patients with hypoglycemic episodes, peripheral vascular disease, and bronchospastic disease, beta-blockers should be used with caution.

The NESTOR study—a multinational, multicenter, double-blind, randomized controlled, 2-parallel-groups study over 1 year—found that indapamide SR (a thiazide-type diuretic) treatment is as efficacious as enalapril in reducing proteinuria and lowering blood pressure.

A meta-analysis of RCTs in patients with non-diabetic renal disease and RCTs or time-controlled studies with nonrandomized crossover design in patients with diabetic nephropathy revealed that dietary protein restriction effectively slows the progression of both diabetic and non-diabetic renal disease.

**Recommendations from others**

From the American Diabetes Association’s “Standards of Medical Care in Diabetes” (position statement): to reduce the risk or slow the progression of nephropathy, optimize glucose and blood pressure control.

- Patients with diabetes should be treated to a blood pressure <130/80 mm Hg.
- For patients with diabetes and albuminuria or nephropathy who are intolerant to ACE inhibitors or ARBs, NDCAs, diuretics, or beta blockers are recommended for treating hypertension. NDCA use may reduce albuminuria in patients with diabetes, including during pregnancy.

**References**


**Diuretics are effective for treating hypertension in patients with diabetes at high risk for cardiovascular disease**


---

**CASE STUDIES IN MENOPAUSE**

**Applying the data from clinical trials to your practice**

**CASE 1**

**Estrogen deficiency during menopause: Its role in the metabolic syndrome**

Steven R. Goldstein, MD
Professor, Obstetrics and Gynecology
New York University School of Medicine

**CASE 2**

**Assessing fracture risk in women with osteopenia**

Michael McClung, MD
Director, Oregon Osteoporosis Center

Look for the CME supplement at [www.obgmanagement.com](http://www.obgmanagement.com)
What is the role of tacrolimus and pimecrolimus in atopic dermatitis?

Sara Trammell, MD, Amer Shakil, MD
Department of Family Medicine and Community Medicine, University of Texas Southwestern, Dallas
Laura Wilder, MLS
University of Texas Southwestern Medical Center Library, Dallas

**Evidence summary**

A recent meta-analysis included 25 randomized controlled trials involving tacrolimus and pimecrolimus. This review included trials of tacrolimus and pimecrolimus in comparison with placebo, topical corticosteroids of varying strengths, and each other. They reported on both safety and efficacy. Fifteen vehicle-controlled trials of pimecrolimus and tacrolimus were reviewed. Both medications proved to be significantly more effective than the vehicle alone. A total of 3 trials (732 patients) compared tacrolimus 0.1% with potent topical corticosteroids (hydrocortisone butyrate 0.1%,
beta-methasone valerate 0.1%) and found it to be as effective as the topical steroids after 3 weeks of application (number needed to treat [NNT]=6).2,3

At both the 0.03% and 0.1% strengths, tacrolimus was found to be more effective than mild topical corticosteroids (hydrocortisone acetate 1%) in 2 studies enrolling a total of 1183 children with moderate to severe atopic dermatitis4,5 (NNT=5 for the tacrolimus 0.03%, and NNT=3 for tacrolimus 0.1%).6 A randomized, double-blinded, multicenter trial compared the use of pimecrolimus 1% cream with 0.1% triamcinolone acetonide cream and 1% hydrocortisone acetate cream for 658 adults with moderate-to-severe atopic dermatitis.7 The majority of patients used either form of treatment for 1 year.

Although long-term safety and tolerability were similar, topical corticosteroids were more efficacious (NNT=13). Another study compared pimecrolimus 1% with betamethasone valerate 0.1% (a potent corticosteroid) in a study of 87 patients.8 At the end of 3 weeks, the pimecrolimus 1% cream was significantly less effective than betamethasone valerate 0.1% (NNT=4).

In a meta-analysis of 3 randomized studies of head-to-head comparison of pimecrolimus 1% and tacrolimus 0.03% or 0.1% among children and adults, tacrolimus ointment was more effective than pimecrolimus cream at the end of the study for adults (P<.0001), for children with moderate-to-severe disease (P=.04), in the combined analysis (P<.0001), and at week 1 for children with mild disease (P=.04). No significant difference was seen in the incidence of adverse effects, although more pimecrolimus-treated patients withdrew from the studies because of a lack of efficacy (P=.03) or adverse events (P=.002; pediatric mild).9

The authors of the first meta-analysis concluded that pimecrolimus 1% was more effective compared with placebo, less effective than potent topical corticosteroids, and had yet to be studied in comparison with low-potency topical corticosteroids. Tacrolimus 0.1% was more effective than placebo, more effective than mild corticosteroids, and as effective as potent topical corticosteroids. It was noted that both these agents caused more burning of the skin than topical corticosteroids—pimecrolimus 1% compared with betamethasone valerate 0.1% (number needed to harm [NNH]=50); tacrolimus 0.1% compared with betamethasone valerate 0.1% and hydrocortisone butyrate 0.1% (NNH=3); and tacrolimus 0.03% compared with the mild corticosteroid hydrocortisone acetate 1% (NNH=10). However, there was no significant difference in the rate of skin infections.

**Recommendations from others**

In 2003, a work group of dermatologists appointed by the president of the American Academy of Dermatology published a technical report on the guidelines of care for atopic dermatitis.10 This group evaluated the effectiveness of several topical treatments for the treatment of atopic dermatitis. They noted that coal tar and its derivatives may reduce the severity of atopic dermatitis symptoms, but there are significant barriers to compliance. The severity of pruritus associated with atopic dermatitis may be reduced with short-term use of topical doxepin.

Evidence supports the use of emollients in combination with other topical corticosteroid treatments to reduce the severity of atopic dermatitis. However, emollients need frequent application, which may be associated with poor compliance. The work group also concluded that both tacrolimus and pimecrolimus are effective and safe in reducing the severity of atopic dermatitis symptoms for both children and adults up to 1 year of treatment.

In March 2005, the FDA posted a Public Health Advisory and Alerts for Healthcare Professionals regarding the potential cancer risk from these products when used to treat atopic dermatitis.
dermatitis. These creams will carry a “black box” warning regarding this potential risk. They recommended use only as a second-line therapy, at minimal amounts necessary, and for short periods of time, not continuously. They also recommended against their use for children aged <2 years and for people with diminished immune systems.

REFERENCES


What physical exam techniques are useful to detect malingering?

**Evidence-Based Answer**

No examination technique objectively proves malingering (strength of recommendation [SOR]: C, expert opinion). Waddell’s signs are associated with poor treatment outcomes but cannot discriminate organic from nonorganic causes (SOR: B, systematic review of low-quality studies). Hoover’s and the Abductor sign indicate nonorganic paralysis (SOR: C, small, lower-quality case-control studies) (TABLE 1).

**Clinical Commentary**

Meticulous examination and documentation will save time and trouble down the road

Warning flags for malingering include persistent noncompliance during prescribed evaluation or treatment, striking inconsistency between physical findings and stated symptoms, and an attorney or insurance company referring the patient to you. If monetary compensation is involved, malingering can potentially be prosecuted as fraud.

Meticulous examination and documentation will save you time and trouble down the road. If you find evidence of malingering, confronting the patient directly will likely result in animosity towards you from the patient and may result in litigation. The confrontation may escalate into violent behavior. Further complicating matters, specialist referral often reinforces the malingering behavior. A common option at approaching the potentially malingering patient is to allow him or her the opportunity to save face: “Well, Mr. Q, I am not finding the usual signs that go along with the complaints you are having....”

If you are in doubt of a diagnosis of malingering, it is generally safest to assume a person is not malingering until you specifically witness a contradictory event.

Tim Huber, MD
US Navy, Camp Pendleton, Calif

**Evidence Summary**

The 4th edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-IV) defines malingering as “the intentional production of false or grossly exaggerated physical or psychological symptoms motivated by external incentives such as avoiding military duty, avoiding work, obtaining financial compensation, evading criminal prosecution, or obtaining drugs.”

Malingering is not considered a mental disorder because symptoms are intentionally produced for external incentives.

No physical exam maneuver can determine a patient’s external incentives. Traditionally, a physician uses certain exam techniques to determine if symptoms are of functional, or nonorganic, origin. Both terms denote the absence of a structural or physiological source for the phenomena, and include malingering and mental disorders such as factitious disorder, conversion disorder, and somatoform disorders. Our literature search only found studies concerning the detection of nonorganic causes of back pain, paralysis, and sensory loss.

Several exam tests are commonly thought to detect nonorganic causes of low back pain. Gordon Waddell described 8 signs in 5 categories (TABLE 2) used to “identify [back pain] patients who require more detailed psychological assessment.”

A systematic review critiqued 60 studies of

---

**TABLE 1**

<table>
<thead>
<tr>
<th>Examination Technique</th>
<th>SOR</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hoover's sign</td>
<td>C</td>
<td>Small, lower-quality case-control studies</td>
</tr>
<tr>
<td>Abductor sign</td>
<td>C</td>
<td>Small, lower-quality case-control studies</td>
</tr>
</tbody>
</table>

**TABLE 2**

<table>
<thead>
<tr>
<th>Waddell's Signs</th>
<th>Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pain on coughing</td>
<td>Physical</td>
</tr>
<tr>
<td>Pain on sneezing</td>
<td>Physical</td>
</tr>
<tr>
<td>Pain on coughing</td>
<td>Psychological</td>
</tr>
<tr>
<td>Pain on sneezing</td>
<td>Psychological</td>
</tr>
</tbody>
</table>

---

Steven Greer, MD, Lee Chambliss, MD, MSPH
Moses Cone Memorial Hospital Family Medicine Residency, Greensboro, NC

Leslie Mackler, MSLS
Moses Cone Health System Library, Greensboro, NC

Tim Huber, MD
US Navy, Camp Pendleton, Calif

---

---

---
Waddell’s signs published between 1980 and 2000. The authors performed a thorough database search, including hand searches of key pain journals, meeting abstracts, and textbooks. The majority of the reviewed studies were small and of lower quality. The review found little evidence on test-retest or interrater reliability. There was consistent evidence that Waddell’s signs are associated with poorer treatment outcomes and generally consistent evidence that they are not associated with secondary gain and cannot discriminate organic from nonorganic problems.

A small, diagnostic case-control study of Mankopf’s test, which is based on the theory that pain increases heart rate, investigated 20 chronic low back pain patients considered nonorganic vs 20 pain-free controls using mechanical pain stimulus applied to subjects’ fingers. There was no significant difference in heart rate response between groups, and no significant effect of pain on heart rate in either group. The authors did not define their criteria for determining patients’ back pain as non-organic, nor did they include patients with low back pain

---

**TABLE 1**

Summary of tests for the detection of malingering

<table>
<thead>
<tr>
<th>TEST</th>
<th>SYMPTOMS</th>
<th>DESCRIPTION</th>
<th>EVIDENCE/OUTCOMES</th>
<th>SOR</th>
</tr>
</thead>
<tbody>
<tr>
<td>McBride’s</td>
<td>Back pain with radicular symptoms</td>
<td>Stand on one leg. Flex symptomatic leg and raise to chest. Refusal or pain = nonorganic</td>
<td>No published studies</td>
<td>C (expert opinion)</td>
</tr>
<tr>
<td>Mankopf’s</td>
<td>Back pain</td>
<td>1700 g pressure applied to the middle phalanx of the second finger of the nondominant hand. True pain should increase heart rate.</td>
<td>Did not correlate with organic pain</td>
<td>C (small inconclusive diagnostic case-control study)</td>
</tr>
<tr>
<td>Waddell’s</td>
<td>Back pain</td>
<td>Positive signs from 3 or more categories (TABLE 2)</td>
<td>Cannot discriminate organic from nonorganic</td>
<td>C (from SR)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Associated with poorer treatment outcomes</td>
<td></td>
<td>C (from SR)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Not associated with secondary gain</td>
<td></td>
<td>B (from SR)</td>
</tr>
<tr>
<td>Hoover’s</td>
<td>Leg paresis</td>
<td>Cup heels and have patient press down with paretic limb. Then have patient raise opposite limb. True paresis if no difference in downward pressure at heels</td>
<td>Indicates nonorganic paresis</td>
<td>C (extrapolated from small diagnostic case-control study using strain gauge)</td>
</tr>
<tr>
<td>Abductor</td>
<td>Leg paresis</td>
<td>Ask patient to abduct paretic leg to resistance. In true paresis, opposite leg should abduct.</td>
<td>Indicates nonorganic causes</td>
<td>C (small, lower-quality case-control study)</td>
</tr>
<tr>
<td>Arm Drop</td>
<td>Arm paresis</td>
<td>Hold paretic hand above face and drop it. If hand misses face, paresis is nonorganic</td>
<td>No published studies</td>
<td>C (expert opinion)</td>
</tr>
<tr>
<td>Midline Split</td>
<td>Sensory loss</td>
<td>Test facial sensation to pinprick. Nonorganic loss of sensation is delineated by the midline.</td>
<td>Very weakly indicates nonorganic cause</td>
<td>C (small diagnostic case-control study)</td>
</tr>
</tbody>
</table>

SOR, strength of recommendation (see page 722); SR, systematic review.
caused by an identifiable pathology. There was no mention of blinding. This literature search found no published studies of McBride’s test, where the patient’s refusal to stand on the unaffected leg and flex the affected leg to the chest determines a feigned radiculopathy.

A few tests attempt to detect nonorganic causes of paralysis. In Hoover’s test, a patient is asked to alternately press down with the paralyzed leg and raise the unaffected leg to resistance, while the hand of the examiner cups the heel of the affected leg. A small, diagnostic case-control study using a computer-assisted strain gauge to measure movement effort during Hoover’s test involved 7 women with true paresis, 9 with nonorganic paresis, and 10 controls. The investigators diagnosed nonorganic paresis by history, neurological exam, and lack of positive neuroradiologic findings. The authors calculated a maximal involuntary to voluntary ratio for each patient’s extremities. The calculation discriminated between all 9 nonorganic patients and both the normal controls and patients with true paresis. The authors did not mention blinding in the study. No attempt was made to compare the strain gauge measurements with a clinician-performed Hoover’s test.

The Abductor sign, based on a similar theory that thigh abductors work in concert, was developed and studied by one individual. In this diagnostic case-control study, the single author tested 33 patients from his practice, 17 with organic paresis, and 16 with nonorganic paresis. The author differentiated organic from nonorganic paresis by history, physical exam, and various imaging studies with no independent assessment. He reported his test as 100% accurate. We did not find any published studies of the Arm Drop test, where feigned paralysis of an upper extremity is tested by holding the arm over the face of the supine patient and letting go.

The Midline Split test attempts to detect nonorganic causes of sensory loss. The fact that cutaneous nerves cross the midline is the basis for the idea that a sharp midline split denotes nonorganic sensory loss. In 1 diagnostic cohort study of 100 people presenting to a neurology department with complaints of decreased sensation on one side of the face, 80 patients were determined to have organic deficits such as multiple sclerosis or stroke. The author did not describe how these diseases were diagnosed. Of those with organic deficits, 7.5% showed midline splitting of sensory loss, falsely suggesting a nonorganic process. Only 20% of the patients with nonorganic sensory loss showed the expected midline split.

### TABLE 2

<table>
<thead>
<tr>
<th>CATEGORY</th>
<th>SIGNS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tenderness</td>
<td><strong>Superficial:</strong> light pinching causing pain = positive&lt;br&gt;<strong>Nonanatomic:</strong> deep tenderness over a wide area = positive</td>
</tr>
<tr>
<td>Simulation</td>
<td><strong>Axial loading:</strong> downward pressure on the head causing low back pain = positive&lt;br&gt;<strong>Rotation:</strong> Examiner holds shoulders and hips in same plane and rotates patient. Pain = positive</td>
</tr>
<tr>
<td>Distraction</td>
<td>Straight leg raise causes pain when formally tested, but straightening the leg with hip flexed ninety degrees to check Babinski does not</td>
</tr>
<tr>
<td>Regional</td>
<td><strong>Weakness:</strong> multiple muscles not enervated by the same root&lt;br&gt;<strong>Sensation:</strong> glove and stocking loss of sensation.</td>
</tr>
<tr>
<td>Overreaction</td>
<td>Excessive show of emotion</td>
</tr>
</tbody>
</table>

If you are in doubt about a diagnosis of malingering, it is safest to assume a person is not—unless you witness a contradictory event.
Evidence-based medicine ratings

**The Journal of Family Practice** uses a simplified rating system called the Strength of Recommendation Taxonomy (SORT). More detailed information can be found in the February 2003 issue, “Simplifying the language of patient care,” pages 111–120.

**Strength of Recommendation (SOR) ratings** are given for key recommendations for readers. SORs should be based on the highest-quality evidence available.

- A Recommendation based on consistent and good-quality patient-oriented evidence.
- B Recommendation based on inconsistent or limited-quality patient-oriented evidence.
- C Recommendation based on consensus, usual practice, opinion, disease-oriented evidence, or case series for studies of diagnosis, treatment, prevention, or screening.

**Levels of evidence** determine whether a study measuring patient-oriented outcomes is of good or limited quality, and whether the results are consistent or inconsistent between studies.

**STUDY QUALITY**
1. Good-quality, patient-oriented evidence (eg, validated clinical decision rules, systematic reviews and meta-analyses of randomized controlled trials [RCTs] with consistent results, high-quality RCTs, or diagnostic cohort studies)
2. Lower-quality patient-oriented evidence (eg, unvalidated clinical decision rules, lower-quality clinical trials, retrospective cohort studies, case control studies, case series)
3. Other evidence (eg, consensus guidelines, usual practice, opinion, case series for studies of diagnosis, treatment, prevention, or screening)

**Consistency across studies**
Consistent—Most studies found similar or at least coherent conclusions (coherence means that differences are explainable); or If high-quality and up-to-date systematic reviews or meta-analyses exist, they support the recommendation.
Inconsistent—Considerable variation among study findings and lack of coherence; or If high-quality and up-to-date systematic reviews or meta-analyses exist, they do not find consistent evidence in favor of the recommendation.

The author apparently performed the sensory exam without blinding or independent confirmation.

**Recommendations from others**
The *DSM-IV* recommends suspicion of malingering for patients who present with 2 or more of the following: medicolegal issues, disagreement between objective and subjective stress or disability, noncompliance with evaluation or treatment, or antisocial personality disorder.

The American Medical Association published the Guides to the Evaluation of Permanent Impairment, which states, “Confirmation of malingering is extremely difficult and generally depends on intentional or inadvertent surveillance.”

**REFERENCES**
Do TZDs increase the risk of heart failure for patients with diabetes?

Serena Baskin, MD, Jon O. Neher, MD
Valley Family Medicine Residency Program, Renton, Wash

Sherry Dodson, MLS
University of Washington Health Sciences Library, Seattle

Evidence summary
A retrospective cohort study of health insurance claims compared the incidence of CHF among 5441 patients with diabetes who had taken TZDs (rosiglitazone, troglitazone, or pioglitazone) vs 28,103 who had not. Patients were allowed other oral agents and insulin, and they were followed for up to 6 years. The TZD group had more patients on insulin and with pre-existing comorbidities. Based on Kaplan-Meier estimates, which control for censored information, the incidence of new heart failure at 40 months was 8.2% in the TZD group and 5.3% in the non-TZD group (number needed to harm [NNH]=34.5). Using a multivariate analysis that controlled for the coadministration of insulin, the hazard ratio for TZD use was 1.76 (95% confidence interval [CI], 1.43–2.17). The incidence of CHF was 3.24% in the troglitazone group (n=1665), 2.39% in the rosiglitazone group (n=1882), and 1.63% in the pioglitazone group (n=1347). The difference in these rates is not statistically significant. Of the 28,103 patients not on a TZD, 1.41% developed heart failure. Individual agents were not compared with placebo.

A manufacturer-sponsored study that combined data from 4 separate unpublished randomized controlled trials compared the incidence of CHF at 1 year for patients treated with pioglitazone (as monotherapy and in combination with other oral agents) with those treated only with other oral agents.

CLINICAL COMMENTARY
Consider stopping TZDs for patients developing edema or CHF
Improved glycemic control decreases the risk of end organ damage and heart failure in patients with diabetes. Thiazolidinediones are very useful drugs, particularly for patients with marked insulin resistance and hyperlipidemia. However, they do precipitate edema and heart failure. The edema can be severe enough to lead to discontinuation of the drug, and the risk of heart failure limits the population in which they can be used. They can be used safely in some cardiac patients but, as noted in the article, they should be avoided or used with caution in patients with CHF. Patients taking a TZD who subsequently develop edema should be carefully evaluated for CHF.

Richard Hoffman, MD
Chesterfield Family Practice, Richmond, Va
Cardiac failure was noted in 12 of 1857 in the pioglitazone group vs 10 of 1856 subjects in the non-pioglitazone groups (not statistically significant). The paper did not comment on how the patients were recruited, how outcomes were measured, or why the 4 original studies were not published.2

Another manufacturer-sponsored retrospective cohort study of pioglitazone analyzed insurance claims data to compare the incidence of CHF among 1668 adult patients taking pioglitazone (and possibly other medications, but not insulin) vs 1668 adult patients taking insulin (and possibly other medications, but not a TZD). The 2 groups were matched in terms of comorbid conditions, but statistical analysis did not take disease severity into account. The incidence of CHF was 2% of pioglitazone users compared with 4% of patients using insulin (NNH for insulin=50). In addition, CHF-related hospitalizations were 0.7% for CHF in the pioglitazone group vs 2.5% in the insulin group (NNH for insulin=55). Both of these findings are statistically significant.3

Recommendations from others
The American Diabetes Association/American Heart Association recommends that patients be evaluated for heart disease or heart failure before starting TZD therapy and monitored for symptoms thereafter. Patients who are at risk for developing CHF, who already have New York Heart Association class I or II CHF, or who take insulin should begin TZD therapy with low doses that are titrated up gradually. The US Food and Drug Administration has not approved TZDs for patients with class III or IV CHF, as there are no studies in these populations.4

REFERENCES
Is therapy based on endoscopy results better than empiric therapy for dyspepsia?

Kevin Rich, MD
University of Washington Family Practice Residency of Idaho, Boise

Sarah Safranek, MLIS
University of Washington Health Sciences Libraries, Seattle

Evidence-based answer

In the initial management of dyspepsia for patients without “alarm” symptoms (weight loss, recurrent vomiting, dysphagia, anemia, evidence of bleeding, onset of dyspepsia after age 45 years), therapy based on the results of early endoscopy was not better than empiric acid suppression (anti-secretory therapy) or a Helicobacter pylori “test and treat” strategy in reducing symptoms or improving quality of life (strength of recommendation [SOR]: A, based on a systematic review). Results from studies of patient satisfaction comparing early endoscopy with empiric medication therapy are conflicting (SOR: A, based on 2 randomized controlled trials [RCTs]).

Though formal cost analyses are not available, a strategy using “test and treat,” as opposed to early endoscopy, results in significantly fewer endoscopies, which when formally evaluated, may translate into a more cost-effective strategy of care (SOR: A, based on a systematic review). Long-term follow-up suggests that patients receiving “test and treat” therapy may require fewer antisecretory medication prescriptions compared with patients receiving early endoscopy (SOR: B, based on a single RCT).

Clinical commentary

Test-and-treat for H pylori a reasonable first option

Guidelines for treating dyspepsia have to consider several factors: clinical outcomes, risk vs benefit to the patient, direct and indirect medical costs, and patient preference and satisfaction. This well-constructed review clearly demonstrates there is no significant difference in symptom control between early endoscopy and empiric acid suppression or testing and treating for H pylori. The evidence regarding 2 other outcomes—patient satisfaction and cost (especially if the indirect cost of sick days is considered)—is less clear.

In my experience, testing and treating for H pylori is a reasonable first option, which often avoids long courses of antisecretory therapy or costly endoscopy. I treat patients who are negative for H pylori with 8 weeks of acid suppression therapy, and refer those with persistent symptoms for endoscopy. I follow patients carefully and try to distinguish between symptoms of dyspepsia and reflux, which requires longer courses of acid suppression. For patients with alarm symptoms, I recommend early endoscopy.

Wail Malaty, MD
Mountain Area Health Education Center, Rural Track Family Practice Residency, Hendersonville, NC

Evidence summary

Though individual studies have suggested that therapy based on endoscopy performed before any other study (early endoscopy) may be superior to empiric antisecretory therapy and as efficacious as a “test and treat” strategy in symptom relief, a Cochrane systematic review of 20 RCTs (11 in primary care settings) provides the best evidence on the role of early endoscopy.1 A subgroup analysis of 5 RCTs, which compared early endoscopy with empiric antisecretory therapy (typically for 4 weeks), revealed that early endoscopy demonstrated a trend towards improvement in self-reported symptoms and in dyspepsia symptom relief scores, but the difference was not statistically significant (relative risk [RR]=0.89; 95% confidence interval [CI], 0.77–1.1). Because each study
used different symptom scores, the relative risk as calculated may under-represent the true benefit of early endoscopy when compared with empiric antisecretory therapy.

When patient satisfaction was evaluated, results were dependent on the location of care. In a primary care setting, patients undergoing early endoscopy were as satisfied as those receiving empiric antisecretory therapy. In a trial of 414 patients randomized after referral to specialty care, patients in the early endoscopy group were more satisfied with their medical care than those receiving empiric antisecretory therapy (RR=0.13; 95% CI, 0.06–0.29). Results from studies comparing the benefits of *H pylori* “test and treat” strategies to early endoscopy are conflicting. A subgroup analysis reported on 3 RCTs from both primary and secondary settings with 931 patients comparing *H pylori* “test and treat” to initial endoscopy. It found no significant difference in symptom reduction (RR=1.06; 95% CI, 0.98–1.26). A recent follow-up study of 1 of the trials included in the Cochrane systematic review reported on outcomes of a “test and treat” vs early endoscopy strategy at 6 years. There was no difference in days without symptoms demonstrated between the 2 groups (mean difference=0.05; 95% CI, –0.03 to 0.14 days). Self-reported symptom tracking and a poor response rate (62%) to patient questionnaires reduces the strength of this study’s conclusions.

Formal cost-effective analyses comparing the “test and treat” with early endoscopy strategy have not been done. A subgroup analysis of 4 trials from the Cochrane review (1 from primary care) demonstrated a significant reduction of the number of endoscopies among patients receiving “test and treat” care vs those receiving early endoscopy (RR=0.23; 95% CI, 0.12–0.44). In the long-term follow-up study, fewer antisecretory medication prescriptions were needed by those patients in the “test and treat” group (*P*=.047). These figures are more robust; they were obtained from national registry data rather than personal recall and questionnaire submission.

**Recommendations from others**

Guidelines from the American Gastroenterological Association for the initial approach to young patients with dyspepsia without alarm symptoms is to first “test and treat” for those testing positive for *H pylori*, prescribe empiric antisecretory therapy for those testing negative, and proceed with endoscopy for recurrent or persistent dyspepsia at 4 to 8 weeks. The American Society for Gastrointestinal Endoscopy does not recommend any of initial endoscopy, empiric antisecretory therapy, or “test and treat” over another for the reduction of symptoms. The British Society of Gastroenterology recommends that initial management of dyspepsia consist of empiric acid suppression and *H pylori* testing. Persons testing positive for *H pylori* should undergo endoscopy. The Institute for Clinical Systems Improvement recommends nonurgent upper endoscopy for those aged 50 years and older with symptoms of uncomplicated dyspepsia. They recommend initial *H pylori* testing and treating those with positive results, and empiric proton pump inhibitor treatment for 4 weeks for those who are *H pylori*–negative.

**References**

permanent hearing loss before 10 months of age (at cost of 200 extra referrals for false-positives). Sensitivity and specificity of the hypothetical model’s 2-stage screening was 85% and 97%, respectively. The estimated positive predictive value was 6.7%.1,3

Individually, OAE and ABR accurately diagnose neonatal hearing loss. One multicenter cohort of 2995 infants measured test performance of OAE and ABR against the gold standard (visual reinforcement audiometry performed at 8–12 months).4 The authors used a receiver operating characteristics (ROC) curve to plot speech awareness thresholds for both tests. When middle-ear pathology and progressive hearing loss were excluded, the area under the ROC curves for ABR and OAE were 0.91 and 0.94, respectively, indication that both tests had excellent test accuracy (a perfect test would have an area under the curve of 1.0).

Evidence summary

In the United States, approximately 5000 infants with moderate-to-profound hearing loss are born annually.1 Affected children graduate high school averaging 4th-grade academic performance skills.2 Efforts to reduce the impact on these children have focused on early diagnosis and treatment.

A systematic review gathered studies comparing universal hearing screening with selective screening.1 Most included studies used a 2-stage universal screening protocol. Infants who failed initial testing were retested within 12 weeks. Testing methods included otoacoustic emissions (OAE) and auditory brainstem response (ABR). Infants who failed the second test were referred for audiological evaluation. Using these data, a hypothetical model was created, which found that 1441 newborns would need to be screened to diagnose 1 additional case of moderate-to-profound hearing loss before 10 months of age (at cost of 200 extra referrals for false-positives). Sensitivity and specificity of the hypothetical model’s 2-stage screening was 85% and 97%, respectively. The estimated positive predictive value was 6.7%.1,3

Should we recommend universal neonatal hearing screening?

Evidence-based answer

Universal neonatal hearing screening leads to both earlier detection and earlier treatment of infants with hearing loss (strength of recommendation [SOR]: A, based on a systematic review). Available evidence suggests early identification and intervention may improve language outcomes (SOR: C, based on retrospective cohort studies).

Clinical commentary

Despite lack of evidence, early intervention could aid future language skills

Despite the lack of hard outcomes data to support neonatal hearing screening, it seems reasonable that early intervention will aid future language skills. Hopefully, future evidence will support the notion that early treatment leads to tangible school performance improvement. For most, however, the decision to universally screen neonates will be guided by state law rather than clinical evidence alone; 38 states currently have mandated screening programs with legislation pending in others.

Evidence summary

In the United States, approximately 5000 infants with moderate-to-profound hearing loss are born annually.1 Affected children graduate high school averaging 4th-grade academic performance skills.2 Efforts to reduce the impact on these children have focused on early diagnosis and treatment.

A systematic review gathered studies comparing universal hearing screening with selective screening.1 Most included studies used a 2-stage universal screening protocol. Infants who failed initial testing were retested within 12 weeks. Testing methods included otoacoustic emissions (OAE) and auditory brainstem response (ABR). Infants who failed the second test were referred for audiological evaluation. Using these data, a hypothetical model was created, which found that 1441 newborns would need to be screened to diagnose 1 additional case of moderate-to-profound hearing loss before 10 months of age (at cost of 200 extra referrals for false-positives). Sensitivity and specificity of the hypothetical model’s 2-stage screening was 85% and 97%, respectively. The estimated positive predictive value was 6.7%.1,3

Individually, OAE and ABR accurately diagnose neonatal hearing loss. One multicenter cohort of 2995 infants measured test performance of OAE and ABR against the gold standard (visual reinforcement audiometry performed at 8–12 months).4 The authors used a receiver operating characteristics (ROC) curve to plot speech awareness thresholds for both tests. When middle-ear pathology and progressive hearing loss were excluded, the area under the ROC curves for ABR and OAE were 0.91 and 0.94, respectively, indication that both tests had excellent test accuracy (a perfect test would have an area under the curve of 1.0).
Strategies based on selective screening of high-risk infants fails to identify permanent hearing loss in many affected infants. In a cohort study of more than 10,000 infants, only 43% of infants with permanent hearing loss were identified with selective versus universal screening. Most affected infants would have been missed using risk-based criteria.\(^5\)

Limited evidence suggests that early identification of infants with permanent hearing loss improves language skills. In a retrospective cohort study of 150 infants examining language outcomes, participants were grouped according to age at identification of hearing loss.\(^6\) All participants received comprehensive in-home language intervention services plus amplification devices.

Of the 85 children with normal cognitive ability, the mean receptive and expressive language quotients at 13 to 36 months were higher in the early-identified group vs the late-identified group (receptive language quotients, 79.6 vs 64.6, \(P<.001\); expressive language quotients, 78.3 vs 63.1, \(P<.001\)). Total language quotient was also higher in the early group (language quotients, 79 vs 64; \(P<.001\)).

The conclusions were limited by multiple factors: retrospective study design, cohort selection drawn from different hospitals during different time periods, unblinded participant selection, and unblended outcome assessments. Other published studies have inconclusive outcome data. The Cochrane Collaboration published a systematic review in which no studies were found that fulfilled the inclusion criteria to evaluate the effectiveness of universal hearing screening.\(^7\)

**Recommendations from others**

The Joint Committee on Infant Hearing recommended universal neonatal hearing screening during hospital birth admission in their Year 2000 Position Statement.\(^8\) For infants whose hearing is impaired on re-screening, the committee recommends audiology referral and medical evaluation to rule out associated conditions before age 3 months. They further recommend interventional services begin before age 6 months for infants with confirmed hearing loss.

The US Preventive Services Task Force does not recommend for or against universal hearing screening, citing insufficient outcomes data.\(^9\)

**REFERENCES**