Evidence-based answer
Complications from cosmetic Botox injections depend on the location of treatment; they include headache (2%–17%), brow ptosis (3.1%), blepharoptosis (2.5%), muscle imbalance (6.9%), muscle bulge (5.9%), and bruising (9.2%–25%). Complications may last from several hours (headache) up to 1 month (some nerve paralysis) (SOR: A, systematic review of RCTs and subsequent RCT). Treatment applied to the upper and periorcular regions of the face have the highest complication rates (4%–8%) (SOR: C, case series).

Evidence summary
A systematic review of 35 randomized, double-blind and open-label trials (N=8,787) evaluated the safety of botulinum toxin from 2000 to 2012. Studies with safety as a primary or secondary endpoint and studies indicating treatment for aesthetic conditions were included. Formulations used included onabotulinumtoxinA, abobotulinumtoxinA, and incobotulinumtoxinA at doses of 2 to 199 units. Only some results were pooled; the rest were reported separately, and statistical testing between treatment and placebo was not reported.

The most common adverse events were documented based on treatment location. All events were temporary and resolved spontaneously, within hours (headache) or up to a month later (eyelid ptosis).

For treatment of glabellar lines, 2% to 16.8% experienced a transient mild to moderate headache, which was reportedly not significantly different from 0% to 20% for placebo (15 studies, n=6,183). For any site, blepharoptosis occurred in 2.5% compared with 0% in the placebo group (11 studies, n=5,689) while eyebrow ptosis occurred in 3.1% (9 studies, n=1,661). For treatment of crow’s feet, mild bruising was the most common adverse event in 9.2% to 25%, similar to the 12.5% rate in the control group (2 studies, n=212). Lower face treatments induced perioral and labial muscular imbalance in 6.9% (4 studies, n=203). In 1 study (n=82), treatment of hypertrophied masseter muscles produced a muscle bulge in 5.9% compared with 0% of controls.
A 2013 double-blind, multicenter RCT evaluated 276 patients to assess the safety and efficacy of botulinum injections for glabellar frown lines.\(^2\) Patients received 5 injections of 0.1 mL per site that contained either 4 units of incobotulinumtoxinA or placebo at each site and were observed for 120 days. Patients in the study were adults with moderate to severe glabellar frown lines based on a facial wrinkle scale (FWS) score of 2 or 3. The FWS measures the severity of glabellar frown lines as 0-none, 1-mild, 2-moderate, or 3-severe.

Adverse reactions occurred in 7.1% of the treated patients versus 2.2% in those receiving placebo injections (no statistical analysis reported). The most common reaction was headache, reported in 3.8% of the botulinum group versus 2.2% in the placebo group (no statistical analysis reported). Two cases of transient ptosis resulted from botulinum injection of the bilateral eyebrows and brow.\(^2\)

A retrospective case series of 1,819 patients (5,310 treatments) at a single medical center analyzed the adverse events of botulinum toxin treatments performed by several specialists (neurology, ophthalmology, dermatology, plastic surgery) for conditions including cervical dystonia, hemifacial spasm, jaw dyskinesia, masseter hyperplasia, and wrinkle correction.\(^3\) Patients who received at least 1 injection of botulinum toxin A (onabotulinumtoxinA or abobotulinumtoxinA) were included. Most patients were women (73%), with an average age of 54 years.

Side effects were categorized as either muscle-related adverse events (ptosis, drooling, neck discomfort, dysphagia, and facial paralysis) or muscle-unrelated adverse events (edema, bruising, injection pain, and dissatisfaction). Ptosis (0.8%) and dissatisfaction (0.4%) were the most common. Treatment of blepharospasm, upper facial wrinkles, and hemifacial spasm incurred the highest adverse event rates of 8%, 4%, and 4%, respectively, with muscle-related and muscle-unrelated events occurring at similar rates. A periocular injection site carried the highest adverse event rates. All adverse events (not described) were considered tolerable and lasted less than 4 weeks.\(^3\)

Matthew Lomeli, MD  
Michael Mercado, MD  
Naval Hospital Camp Pendleton  
Camp Pendleton, CA

The opinions and assertions contained herein are those of the authors and are not to be construed as official or as reflecting the views of the US Navy Medical Department, the US Navy at large, or the US Department of Defense.

REFERENCES
Small bites

Recently, I visited a tapas restaurant with several friends. We each ordered a different tapas plate to share as a group. Soon, pretty little dishes started to appear. The herb-crusted strip sirloin on edible meadow flowers was masterful. The skewed eggplant and oyster mushroom micro-kebab with wasabi aioli was nearly divine. Unfortunately, even though every item was a work of art, all these tiny masterpieces collected together did not seem to make a proper meal for any of us.

This got me thinking about small research studies, which have a number of unfortunate features that we need to be aware of in evidence-based medicine—features well-described in a recent paper in Neuroscience. The authors discussed 3 problems that are always present in small studies, because they arise directly or indirectly from the rules of probability:

- The chance of discovering a true effect is low (the definition of a low-powered study)
- There is a greater chance that something will pass the P<.05 confidence threshold that is not actually true
- When a true effect is discovered, follow-up research is likely to reduce the magnitude of the measured effect (a.k.a., the “winner’s curse”)

In addition, the authors noted that smaller research studies are also at increased risk of other types of problems, such as:

- Uncomfortably wide estimates of effect size across studies (due to differences in protocol)
- Publication bias, whereby small studies with null results may not be written up for publication
- Poor design quality, because smaller samples likely correlate with lower design budgets

The authors then pulled 49 meta-analyses in the field of neuroscience and compared them with the 730 primary studies that these meta-analyses contained. The authors found that the average statistical power of the original research to find the outcomes later established in the meta-analyses was a paltry 8% to 31%. The neuroscientists concluded that many associations reported in the neuroscience literature were probably overstated. They also concluded that neuroscience studies generally need to be bigger.

So, waiter! Take away these research tapas. Bring me your largest, juiciest RCT. And make it well done!

---

REFERENCE

Diving for PURPs

Nasal balloon autoinflation for otitis media with effusion does not pass the sniff test


This RCT compared the effectiveness of nasal balloon autoinflation (inflating a balloon with each nostril through a specialized nozzle) 3 times daily for 3 months to usual care in 320 children (4–11 years old) with otitis media with effusion. Patients were included if they had hearing loss or other ear-related problems for more than 3 months and tympanometric confirmation of otitis media with effusion.

The primary outcome was tympanometric normalization of middle ear pressure, and secondary outcomes included ear-related quality of life and a weekly symptom diary completed by parents.

No difference was noted between the groups in tympanometric resolution of middle ear effusion at 1 month (adjusted relative risk [ARR] 1.36; 95% CI, 0.99–1.9), but there was higher likelihood of tympanometric resolution in the autoinflation group at 3 months compared with the usual care group (ARR 1.37; 95% CI, 1.03–1.83).

Ear-related quality of life at 3 months (measured with the 14-item quality of life in children’s ear problems scale [OMQ-14]) improved slightly more from baseline in the autoinflation group than in the usual care group (mean difference –0.33; 95% CI, −0.59 to –0.07).

No difference was noted between groups in days with symptoms at 1 month (8 vs 9 days; OR 0.66; 95% CI, 0.41–1.05), but there were fewer symptom days over 3 months in the autoinflation group (14 vs 22 days; OR 0.58; 95% CI, 0.37–0.90).

<table>
<thead>
<tr>
<th>Relevant</th>
<th>Yes</th>
<th>Medical care setting</th>
<th>Yes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Valid</td>
<td>Yes</td>
<td>Implementable</td>
<td>Yes</td>
</tr>
<tr>
<td>Change in practice</td>
<td>Yes</td>
<td>Clinically meaningful</td>
<td>No</td>
</tr>
</tbody>
</table>

**Bottom line:** Nasal autoinflation is a simple and safe procedure, but having children perform it 3 times daily for 3 months may be a difficult task and not worth the effort for the small gain in symptoms.

Review and Summary Authors: Corey Lyon, DO, and Shannon Langner, MD, University of Colorado FMR, Denver, CO

Beers, beers, good for the heart?


A trial examined the association of alcohol consumption and medical outcomes such as cardiovascular disease, cancer, and mortality in countries with different average incomes.

Levels of alcohol consumption were divided up into never drinkers, former drinkers, low intake (1–7 drinks per week), moderate intake (7–14 drinks per week for women, 7–21 drinks for men), and high intake (>14 drinks per week for women, >21 drinks for men).

Compared with never drinkers, current drinkers were at increased risk of cancer (HR 1.51; 95% CI, 1.22–1.89) and injury (HR 1.29; 95% CI, 1.04–1.61), and reduced risk of myocardial infarction (MI) (HR 0.76; 95% CI, 0.63–0.93) and hospital admission (HR 0.86; 95% CI, 0.78–0.94). However, there was no association with mortality (HR 1.00; 95% CI, 0.87–1.14) except in the high intake group vs never drinkers (HR 1.31; 95% CI, 1.04–1.66). No difference was noted for current drinkers in risk of CVD (HR 0.97; 95% CI, 0.87–1.09) or stroke (HR 1.01; 95% CI, 0.82–1.24) vs never drinkers. Low or moderate intake was associated with a reduction in risk of MI vs never drinkers (low: HR 0.77; 95% CI, 0.63–0.94; moderate: HR 0.65; 95% CI, 0.44–0.97), but no difference was noted in risk of MI among those with high intake vs never drinkers.

Current drinking was associated with reduced risk of MI in high- and upper-income countries (HR 0.53; 95% CI, 0.40–0.73), but not in low-income countries (HR 0.97; 95% CI, 0.76–1.25) compared with never drinkers in the same region.

<table>
<thead>
<tr>
<th>Relevant</th>
<th>Yes</th>
<th>Medical care setting</th>
<th>Yes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Valid</td>
<td>No</td>
<td>Implementable</td>
<td>Yes</td>
</tr>
<tr>
<td>Change in practice</td>
<td>No</td>
<td>Clinically meaningful</td>
<td>Yes</td>
</tr>
</tbody>
</table>

**Bottom line:** These outcomes are difficult to apply due to report bias (patients may misreport alcohol intake) and unidentified confounders. Varying levels of alcohol intake in countries with different incomes had opposite clinical outcomes, which may also suggest an unmeasured confounder.

Review and Summary Author: Corey Lyon, DO, University of Colorado FMR, Denver, CO
Does continuous labor support decrease rates of cesarean and assisted vaginal delivery?

**Bottom line**

Yes. Continuous labor support reduces cesarean and instrumented vaginal birth rates. Its effectiveness appears to be strongest when the support provider is neither a hospital staff member nor a part of the woman’s social network. It may be less effective in hospitals that allow for the presence of other support people, and in hospitals where epidural analgesia and external fetal monitoring are routine.

**Case**

A 26-year-old primipara presents with her partner to your office for a routine OB visit at 36 weeks, 6 days gestation. She has heard of a local doula program and wonders if having a doula in the labor room would help her avoid a cesarean delivery.

**Evidence summary**

A 2013 Cochrane meta-analysis of 22 RCTs involving 15,288 women compared the effect of continuous, one-to-one intrapartum support versus usual care on rates of spontaneous vaginal delivery (SVD), assisted vaginal birth, and cesarean birth. Hospital setting and support provider characteristics varied significantly among studies.

Overall, women who received continuous labor support were more likely to have an SVD (RR 1.08; 95% CI, 1.04–1.12) and less likely to have a cesarean birth (RR 0.78; 95% CI, 0.67–0.91) or instrumented vaginal birth (RR 0.90; 95% CI, 0.85–0.96).

Subgroup analyses showed that continuous labor support was more effective for increasing the likelihood of SVD and decreasing the likelihood of cesarean birth when the support provider was a third party (5 trials evaluating SVD, n=1,935; RR 1.12; 95% CI, 1.07–1.17; 7 trials evaluating cesarean birth, n=2,330; RR 0.72; 95% CI, 0.60–0.86) rather than a hospital staff member (9 trials of SVD, n=10,813; RR 1.03; 95% CI, 1.01–1.06; 9 trials of cesarean birth, n=10,786; RR 0.95; 95% CI, 0.85–1.05) or part of the women’s social network (5 trials of SVD, n=1,470; RR 1.07; 95% CI, 0.99–1.15; 6 trials evaluating cesarean birth, n=2,059; RR 0.83; 95% CI, 0.69–1.01; P=.007 for subgroup comparison of SVD, and P=.03 for subgroup comparison of cesarean section).

But subgroup analyses also found that the effect of continuous labor support on both outcomes was not statistically significant in hospitals that permit companions to be present (9 trials of SVD, n=10,889; RR 1.03; 95% CI, 1.00–1.05; 11 trials of cesarean birth, n=11,326; RR 0.94; 95% CI, 0.85–1.03), and the effect on cesarean birth was not statistically significant in hospitals where epidural analgesia was available (14 trials, n=13,064; RR 0.93; 95% CI, 0.86–1.02) or external fetal monitoring was routine (9 trials, n=10,123; RR 0.92; 95% CI, 0.83–1.01).

The largest RCT included in the review randomized 6,915 women in the United States and Canada to continuous labor support from a specially trained nurse or usual care, and found no significant difference in rates of cesarean (12.5% vs 12.6%, respectively; P=.44) or operative vaginal delivery (15.7% vs 16.2%; P=.54). The authors concluded that the benefits of continuous labor support could be overpowered by interventionist birth environments, especially when support providers are hospital employees.

Relevant to this concern, another RCT from the review randomized 420 women in the Cleveland area to receive support from a doula or usual care during labor. All but 3 women had additional support people present. All doulas were trained professionals not employed by the hospital. Women who labored with doula support were significantly less likely to have a cesarean birth (13.4% vs 25.0%; P=.002); instrumented vaginal delivery was not an outcome of this study.

**Case Wrap-Up**

Based on your review of the evidence, your patient employs a doula for her delivery. She goes into labor at 39 weeks, 4 days. Nine hours after arrival to the hospital, she delivers a healthy 3,650-g infant with APGAR scores of 9 and 9. The new mother was happy with the doula and thanks you for the advice.

James Conniff, MD
Lee Dresang, MD
University of Wisconsin–Madison
Madison, WI

**REFERENCES**

Dementia tests and gold standards

<table>
<thead>
<tr>
<th>Brief single domain tests</th>
<th>Brief multi-domain tests</th>
<th>Gold standards</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clock Drawing Test</td>
<td>Mini-Mental Status Examination (MMSE)</td>
<td>Diagnostic and Statistical Manual (DSM)</td>
</tr>
<tr>
<td>Three Word Recall</td>
<td>General Practitioner Assessment of Cognition (GPCOG)</td>
<td>National Institute of Neurological and Communicative Disorders and Stroke—Alzheimer’s Disease and Related Disorders Association (NINCDS-ADRADA)</td>
</tr>
<tr>
<td>John Brown Test</td>
<td>Abbreviated Mental Status Test Score (AMTS)</td>
<td></td>
</tr>
<tr>
<td>Verbal Fluency Animals</td>
<td>Prueba Cognitive de Legans (PCL)</td>
<td></td>
</tr>
<tr>
<td>Orientation (from Mini Mental Status Examination)</td>
<td>Short Portable Mental Status Questionnaire (SPMSQ)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Blessed Information Memory Concentration Test (BIMCT)</td>
<td></td>
</tr>
</tbody>
</table>

Does bracing in adolescents with idiopathic scoliosis reduce scoliosis progression or surgical intervention?

Evidence-Based Answer
Bracing patients with adolescent idiopathic scoliosis (AIS) reduces curve progression (SOR: B, systemic review of RCTs and nonrandomized trials with some disease-oriented outcome measures). Patients with AIS who wear their braces >10 hours a day are less likely to experience curve progression requiring surgical care than patients wearing braces <2 hours a day (SOR: B, nonrandomized prospective study).

A 2015 systematic review included 7 studies (4 RCTs and 3 prospective controlled trials) to evaluate the effectiveness of bracing on curve progression in patients with AIS. Patients were 10 years old or older at the time of AIS diagnosis and followed until bone maturity. Patients with secondary scoliosis (congenital, neurological, metabolic, posttraumatic, etc.) were excluded. Duration of the trials ranged between 1 and 5 years.

Of the 662 patients, 483 were treated with a brace, 133 underwent observation, and 46 were treated with electrical stimulation. Studies were heterogeneous in terms of population characteristics. The mean age was approximately 12.5 years. At baseline, Cobb angles were between 15° and 40°. Success was defined as curves not evolving to >50°.

Rigid bracing compared with observation significantly increased the success rate at 2 years of follow-up in 1 RCT (n=116; RR 1.79; 95% CI, 1.29–2.50), 1 nonrandomized prospective controlled trial (n=146; RR 1.50; 95% CI, 1.19–1.89), and 1 cohort study (n=242; RR 1.50; 95% CI, 1.19–1.89). At 3 years of follow-up, an additional cohort study also showed an increased success rate with rigid bracing compared with observation (n=240; RR 1.75; 95% CI, 1.42–2.16). Elastic bracing increased the success rate compared with observation at 3 years of follow-up in 1 RCT (n=47; RR 1.88; 95% CI, 1.11–3.20).

A 2014 nonrandomized prospective controlled study (N=100) investigated if time spent wearing a brace affected the rate of surgical intervention. Previously untreated patients older than 10 years with AIS and Cobb angles between 25° and 45° were sequentially assigned to wear a heat-sensor-fitted Boston brace either 16 or 23 hours per day. The threshold for surgery was set at curve progression to >50°. Patients were analyzed based on time compliant with wearing the brace and progression to the surgical threshold with data collected every 4 months until skeletal maturity.

Overall, 28% of patients progressed to the surgical threshold. Of the patients who wore the brace for <2 hours a day (n=27), 44% progressed to the surgical threshold, but of the 31 patients who wore the brace for >10 hours a day, only 7% progressed to the surgical threshold (risk difference 38%; 95% CI, 16–57; NNT=3). Wearing the brace >14 hours a day provided no greater benefit than >10 hours a day.

Won Woo (Chris) Lee, MD
Deborah E. Miller, MD, MACM, FAAFP
University of Chicago (NorthShore) FMR
Glenview, IL


Are statins effective for primary prevention of CVD?

Evidence-Based Answer
Statin use is effective in lowering rates of cardiovascular disease (CVD) events (stroke and myocardial infarction [MI]) and all-cause mortality in adult men and women without prior history of CVD events, including in those with low (<10% 5-year) cardiovascular risk (SOR: A, systematic reviews of RCTs).

A recent Cochrane systematic review (18 RCTs, N=56,943) examined the effectiveness of statins for the primary prevention of CVD in adults (mean age 57 years). Trials compared statin versus placebo or usual care; 10% or fewer participants had a history of CVD and all patients were followed for 1 to 5 years.

All-cause mortality was significantly reduced by statins (13 trials, n=48,060; OR 0.86; 95% CI, 0.79–0.94, NNT=125). The composite outcome of fatal and nonfatal CVD events was also reduced (9 trials, n=23,805; RR 0.75; 95% CI, 0.70–0.81, NNT=34), as were combined fatal and nonfatal strokes (10 trials, n=40,295; RR 0.78; 95% CI, 0.68–0.89, NNT=20). No evidence was found for increased risk of myalgia or muscle pain (9 trials, n=37,938; RR 1.0; 95% CI, 0.97–1.1), risk of cancer (11 trials, n=38,739; RR 1.0;
Evidence-Based Practice / May 2016

0.93–1.1), or risk of type 2 diabetes (2 trials, n=24,407; RR 1.2; 95% CI, 1.0–1.4). All but 1 of the trials had some form of pharmaceutical industry sponsorship.¹

A meta-analysis examined 22 RCTs of statin versus placebo (N=134,537, minimum follow-up 4.0 years) for primary and secondary prevention of CVD among patients in 5 CVD risk categories, based on 5-year major vascular event (MVE) risk calculations.² The stratification by MVE risk calculation (with MVE defined as first nonfatal MI, coronary death, stroke, or coronary revascularization procedure) was established by the researchers.

In 23,798 patients with a 5-year predicted risk of <5% without a known history of vascular disease, a statistically significant reduction in annual risk of major vascular events was noted for every 1.0 mmol/L reduction in low-density lipoprotein (LDL) (22 trials; RR 0.61; 95% CI, 0.45–0.81). For 24,674 patients with 5-year predicted risk of ≥5% to <10% without a known history of vascular disease, there was also a similar statistically significant reduction in annual MVEs per 1.0 mmol/L reduction in LDL (22 trials; RR 0.66; 95% CI, 0.57–0.77) as well as all-cause mortality (RR 0.83; 95% CI, 0.69–0.99). All risk groups of patients without a history of vascular disease had a significant reduction in major coronary events (22 trials, n=69,959, RR 0.71; 95% CI, 0.65–0.77) and all-cause mortality (RR 0.91; 0.85–0.97).²

The American College of Cardiology and the American Heart Association (ACC/AHA) 2013 evidence-based guidelines³ recommend use of the new Pooled Cohort Equations calculator to estimate 10-year CVD risk in adults, and identified 4 groups most likely to benefit from statins to reduce CVD events in both primary and secondary prevention:

- Patients with clinical atherosclerotic CVD (Grade A, strong recommendation with high certainty based on the evidence that the benefits are substantial)
- Patients with primary elevations of LDL cholesterol of ≥190 mg/dL (Grade B, with moderate certainty based on evidence that the net benefit is moderate to substantial)
- Adults with diabetes aged 40 through 75 years with LDL levels between 70 and 189 mg/dL (Grade A)
- Adults aged 40 through 75 who have LDL levels 70 to 189 mg/dL and estimated 10-year CVD risk of ≥7.5% (Grade A)

The guidelines state that moderate-intensity statin therapy can be considered in individuals with a 5% to 7.5% 10-year atherosclerotic CVD risk (Grade C, with at least moderate certainty based on evidence that there is a small net benefit).³

Makristy Caratao, MD
John Van Buskirk, DO
Tacoma FMR
Tacoma, WA


Does nasal saline irrigation decrease the symptoms of seasonal allergic rhinitis when used as adjunctive therapy?

Evidence-Based Answer
It appears likely, although the magnitude of any additional effect is unclear. Saline nasal irrigation decreases symptoms of allergic rhinitis, but does not improve quality of life in children and adults compared with a heterogeneous collection of control treatments (SOR: B, extrapolated from a meta-analysis with some nonactive comparators). In children, saline nasal irrigation as adjunctive therapy to nasal steroids decreases symptoms more than either therapy individually (SOR: B, single RCT).

A 2012 meta-analysis of 10 RCTs and prospective cohort studies (N=406) assessed the efficacy of 1.5 to 500 mL nasal saline irrigation used every 4 to 12 hours in 275 adults, 86 children, and 45 pregnant women with allergic rhinitis.³ Control groups consisted of patients using no intervention, oil drops, nasal steroid spray, or cetirizine. The primary outcome was improvement of allergy symptoms, using various severity scales ranging from 0 to 5 points per symptom. Scores were translated to percentages of improvement after nasal irrigation. A secondary outcome was quality of life, using the validated Rhinitis Quality of Life Questionnaire (28 questions, 0–6 scale) and validated Rhinasthma Questionnaire (30 questions, 0–5 scale).

Over 1 to 12 weeks, saline irrigation improved nasal symptom scores by 30.1% (95% CI, 15.7–44.4) compared with control, while quality-of-life scores did not change significantly.¹
A 2014 open-label RCT compared the efficacy of using intranasal corticosteroid spray and saline spray together and separately for allergic rhinitis in children treated by a referral clinic (N=61). The study enrolled children (age range 2–15 years) with moderate to severe allergic rhinitis and a positive skin prick test for common allergens. Patients were excluded if they had marked septum deviation, prior nasal surgery, nasal polyposis, or active infection. Intranasal fluticasone was started at 200 mcg/d and decreased by half every 4 weeks to 50 mcg/d. For nasal saline irrigation, 4–6 sprays of nasal saline spray were applied twice daily using a positive pressure applicator. Patients visited the clinic at 4, 8, and 12 weeks and were assessed for nasal symptoms, based on a 12-point scale ranging from 0 to 3 for itching, rhinorrhea, nasal congestion, and sneezing. The average initial symptom score ranged from 7 to 7.2 (numerical values estimated from graphs).

After 12 weeks, the average symptom score was 6.6 for the irrigation group (7% decrease in symptoms, no \( P \) value reported), 4.2 for the steroid group (41% decrease in symptoms, no \( P \) value reported), and 3.3 for the steroid and irrigation group (64% decrease in symptoms, \( P < .05 \)).

Liya Milgram, DO
Margaret Wiedmann, MD
Advocate Illinois Masonic FMR
Chicago, IL


Do macrolides have long-term clinical benefits in patients with severe asthma?

Evidence-Based Answer

Macrolide antibiotics given for at least 6 weeks improve asthma symptom scores, quality of life (QOL), peak expiratory flow, and airway hyperreactivity (SOR: A, meta-analysis of RCTs). In patients with severe asthma, macrolide therapy does not decrease the number of asthma exacerbations (SOR: B, single RCT).

A 2013 meta-analysis of 12 RCTs (N=831) examined the effectiveness of a prolonged course of macrolide antibiotics (clarithromycin, erythromycin, roxithromycin, or troleandomycin) versus placebo or standard treatment for the long-term management of asthma. Ten studies included adults and 2 included children. Five studies included patients with mild to moderate asthma, 3 studies included patients with severe asthma, and 4 did not specify.

Azithromycin and clarithromycin were the most common antibiotics, and duration of treatment varied from 6 to 26 weeks (mean 8 weeks). Outcomes evaluated included forced expiratory volume in 1 second (FEV1), peak expiratory flow (PEF), airway hyperactivity (measured by dose or concentration of methacholine needed to reduce the FEV1 by 20%), symptom scores, and QOL measures. Details of symptom scores were not given and QOL was measured using the asthma QOL questionnaire, which included asthma symptoms and activity restrictions, with 0.5 indicating a clinically meaningful change.

Macrolides improved final asthma symptom scores (5 trials, n=199; weighted mean difference [WMD] −0.56; 95% CI, −0.73 to −0.39) and QOL (5 trials, n=346; WMD 0.18; 95% CI, 0.001–0.37). Macrolides also improved PEF (4 trials, n=419; WMD 6.70 L/min; 95% CI, 1.35–12.06), but on subgroup analysis PEF was improved in adults only and not children. Macrolides decreased airway hyperactivity (2 trials, n=131; standardized mean difference [SMD] 1.99; 95% CI, 0.46–3.52), although there was significant heterogeneity between the studies. (A SMD of 0.2 is considered small, 0.6 moderate, and 1.2 large.) No improvement was noted in FEV1 (8 trials, n=381; SMD 0.05; 95% CI, −0.14 to 0.25).

Patients on macrolides had more nausea (3 trials, n=403; RR 2.47; 95% CI, 1.22–5.0). Study quality on the 0 to 5 Jadad scale ranged from 3 to 5 (5 being the highest quality score).

A 2013 RCT examined the effect of azithromycin 250 mg 3 times weekly versus placebo over 26 weeks in 109 adult patients, 18 to 75 years old, with severe asthma. Severe asthma was defined as patients on inhaled steroids and long-acting beta-agonists for at least 6 months, and 2 exacerbations requiring corticosteroids in the past 12 months. Current smokers and ex-smokers with a >10 pack-year smoking history were excluded. The primary outcome was the rate of exacerbations over the duration of macrolide therapy.

The rate of exacerbations over 6 months in the azithromycin group was 0.75 compared with 0.81 in the placebo group (\( P = .682 \)). The strict exclusion criteria of smokers and any noncompliance with inhaled steroid
therapy may limit the ability to apply this study to a general patient population.

The National Heart, Lung, and Blood Institute updated guidelines on the management of asthma in 2007. Macrolide antibiotics are not mentioned in the asthma treatment algorithm, citing insufficient evidence for macrolides in patients with chronic asthma.

Russell Lemmon, DO
University of Wisconsin School of Medicine and Public Health
Madison, WI


What is an appropriate age to stop screening for colorectal cancer?

Evidence-Based Answer
Screening patients 75 to 85 years of age with colonoscopy has a small benefit in life-years gained but at significant cost, leading the US Preventive Services Task Force (USPSTF) to recommend stopping routine screening at age 75 (SOR: B, decision analysis). However, the increasing proportion of colorectal cancers (CRC) diagnosed in patients older than 75 in recent years may shift the balance (SOR: C, case-series).

A 2008 decision analysis evaluated the benefits and costs of CRC screening strategies to determine optimal screening intervals and ages to start and stop screening using 2 independent microsimulation models. The models used data from 1975 to 1979 on adenoma prevalence and CRC incidence in the United States, a time period unaffected by widespread screening. All-cause mortality and CRC survival data were obtained from US public health records and the sensitivity of colonoscopy was estimated from recent studies.

The 2 analyses showed that raising the colonoscopy screening stop age from 75 to 85 years would have a small increase in life-years gained (5 and 2 life-years per 1,000 individuals), but result in a large increase in the number of colonoscopies required (398 and 358 colonoscopies per 1,000 individuals). Based on this finding, the USPSTF recommended that clinicians screen for CRC in adults up to age 75 and recommended against routine screening for CRC in adults 76 to 85 years of age, but with cases evaluated on an individual basis.

A 2012 case-series evaluated trends in the age at diagnosis of 124,314 patients with CRC in Taiwan’s national cancer registry from 1988 to 2007. The 20-year study was divided into four 5-year increments and patients were organized into 4 age groups: <50, 50 to 75, 75 to 85, and >85.

The main finding of this study was that the proportion of total CRC cases occurring in the <50 and 50-to-75 age groups decreased during the 20-year period, but the proportion in the 75-to-85 age group increased from 12% to 25% for men and from 12% to 22% for women. Based on the increasing proportion of CRC cases in patients older than 75 years, the authors concluded that screening should be extended beyond age 75.

Muhammad Mustafa Jawaid, MD
St Mary Mercy Hospital
Livonia, MI


What is the best imaging test for the diagnosis of acute appendicitis in children?

Evidence-Based Answer
A staged approach using ultrasound (US) followed by CT scan (if the US is equivocal) is effective in diagnosing acute appendicitis (sensitivity 99% and specificity 91%) (SOR: A, systematic review of diagnostic cohort studies and consistent cohort study). Triaging cases with a pediatric appendicitis score (PAS) prior to US and follow-up with CT if the US is equivocal lowers sensitivity to 92% and increases specificity to 95% (SOR: B, prospective cohort trial). Follow-up MRI after an equivocal US has a sensitivity of 100% and a specificity of 96% (SOR: B, retrospective cohort trial).

A meta-analysis of 57 cohort studies involving 9,356 children (age range 7–12 years) evaluated the accuracy of CT and US for acute appendicitis. The gold standard was pathology after surgery.
Results showed CT had a sensitivity of 94% and a specificity of 95% (positive likelihood ratio [LR+] 19; negative likelihood ratio [LR–] 0.06) and US had a sensitivity of 88% and a specificity of 94% (LR+ 15; LR– 0.13). The authors of the study supported the use of a staged approach (US followed by CT scan) to lower the long-term risk of cancer.

A retrospective study involving 631 pediatric patients (2 months–18 years old) presenting to the emergency department for suspected appendicitis used a staged protocol to rule out appendicitis. The protocol involved conducting an US, followed by a CT scan if the US results were equivocal.

Compared with pathology after surgery, the protocol had a sensitivity of 99% and a specificity of 91% (LR+ 11; LR– 0.01). With the staged evaluation, 8.1% of surgeries removed a normal appendix, while the missed appendicitis rate was 0.5%. CT was avoided in 333 patients (53%).

A prospective, observational cohort trial of 196 pediatric patients (average age 11 years) evaluated the use of US with a PAS (range 0–10; measuring RLQ tenderness to light palpation and percussion, anorexia, fever ≥38°C, nausea/vomiting, WBC >10,000, >75% neutrophils, and migration of pain to RLQ). Low-risk patients (PAS score <3) were discharged, intermediate-risk patients (PAS score 4–7) underwent US evaluation, and high-risk patients (PAS 8–10) underwent surgery. CT scan was used as the gold standard in cases of uncertainty. This pathway resulted in a sensitivity of 92% and specificity of 95% (LR+ 18; LR– 0.08).

A retrospective cohort study involving 331 patients (age range 0–18 years) evaluated a simple algorithm to reduce the use of CT in evaluating appendicitis. Patients with high concern for acute appendicitis based on clinical presentation, physical examination, and CBC were immediately seen by surgery prior to imaging. Patients in the equivocal group underwent US and CBC testing. Patients with a nondiagnostic US were evaluated by surgery prior to a CT scan.

The authors found that CT utilization decreased from 39% to 18% after implementation of the algorithm (P<.001). There was no increase in the negative appendectomy rate (9%–11%; P=.59).

A retrospective review of 60 pediatric patients (age range 7–17 years) evaluated the accuracy of MRI for acute appendicitis after inconclusive US. Pathology after surgery was used for the gold standard. MRI had a sensitivity of 100% and a specificity of 96% (LR+ 25; LR– 0).
CI, 0.58–190), although the confidence interval was wide. No significant difference was noted between multiple groups of estrogen-based products for hyperplasia or increased endometrial stripe compared with placebo. Nevertheless, the authors recommended the use of progesterone when total daily estrogen was high enough to result in significant systemic absorption.1

A 2005 open-label RCT, published after the search date of the Cochrane review discussed above, enrolled 185 postmenopausal women with urogenital atrophy based on symptoms or examination and randomized them to ESTring® (n=126) or Vagifem® (n=59) for 12 months.2 Primary outcomes were changes in the transvaginal ultrasound and progestogen challenge test. ESTring (a vaginal ring containing 2 mg micronized 17β-estradiol) was placed vaginally every 3 months. Vagifem (a vaginal tablet with 25 mcg 17β-estradiol) was inserted daily for the first 2 weeks, then twice a week.

At 48 weeks, no differences were noted between the 2 treatments in endometrial thickness. Three women using ESTring and 2 using Vagifem had endometrial strips >5 mm. Biopsies were taken of strips >7 mm (n=4) and all were negative for endometrial proliferation.2

The North American Menopause Society published a 2013 evidence-based position statement formed after a literature review stating that progesterone is not needed for a woman with a uterus when using low-dose vaginal estrogen.3 The authors noted that endometrial safety had not been studied in clinical trials past 1 year. Progesterone was not recommended for women without a uterus.

For a women at high risk for endometrial cancer, the position paper recommended considering surveillance (transvaginal ultrasound or progesterone challenge), but noted that data were insufficient to recommend routine screening for women using long-term vaginal estrogen.3

Gwen Fitz-Gerald, MD
Corey Lyon, DO
University of Colorado Denver FMR
Denver, CO

Does annual monofilament testing in patients with diabetes improve outcomes?

Evidence-Based Answer
The answer is unclear. A patient’s inability to perceive monofilament testing is a positive predictor of increased likelihood of future foot ulceration, but it is not known if a patient will have an improved outcome if monofilament testing is performed (SOR: B, consistent cohort trials).

A 2011 systematic review of 9 prospective cohort studies of 11,007 patients with type 1 or type 2 diabetes evaluated the use of monofilament testing in predicting the occurrence of diabetic foot ulcers.1 The study compared outcomes of positive monofilament testing with negative monofilament testing at 1 to 4 years. A positive monofilament test was defined variably as decreased sensation in 1 of 3 to 1 of 9 sites on both feet. A positive test result was shown to be a significant predictor of future ulceration versus a negative test result (1-year RR 4.4; 95% CI, 2–10; 4-year RR 4.9; 95% CI, 2.5–9.6).1

In 2006, a prospective cohort study analyzed the effectiveness of monofilament testing for predicting the development of a diabetic foot ulcer in 1,285 adult VA patients with diabetes mellitus type 2 (98% white men).2 Baseline monofilament testing was completed and a positive test result was recorded if there was insensitivity to a Semmes-Weinstein 5.07 monofilament at 1 or more sites. The follow-up interval was 12 to 18 months; follow-up consisted of physical examination, questionnaire, or a phone call, checking for ulceration development, spanning an average of 3.4 years (with 75% of patients followed up for 5 years). During follow-up, 216 patients developed foot ulcers, 210 died before foot ulcer development, and 277 patients were lost to follow-up. Monofilament insensitivity was associated with a higher incidence of future foot ulceration compared with a negative examination result (HR 2.03; 95% CI, 1.50–2.76; P<.001).2

A 2013 prospective observational study evaluated risk factors for the development of new ulceration in 563 undifferentiated diabetic patients over 18 months.3 Screening clinics were held in 12 Irish general practice offices. HbA1C results, eGFR results, neuropathy symptoms scores, cutaneous pressure perception using a 10-g monofilament, vascular symptoms and

signs, and an assessment of foot deformity were evaluated.

Patients with an abnormal monofilament test result (n=16) developed ulcers at a higher rate than patients with a normal monofilament test result (n=370) (87.5% vs 22%; P<.001).³

Steve Schrock, MD
Bethany Perrin, DO
Louis Cavadini, DO
Benjamin Garman, MD
Saint Joseph Regional Medical Center FMR
Mishawaka, IN

Is there a benefit from antiplatelet therapy for preeclampsia prophylaxis in patients at risk for developing preeclampsia?

Evidence-Based Answer

The use of low-dose aspirin in the late first or early second trimester in women at high risk for preeclampsia reduces the incidence of preeclampsia, preterm birth, and intrauterine growth restriction (SOR: B, systematic reviews of RCTs at risk of bias).

A 2014 systematic review and meta-analysis of 21 RCTs (N=24,666) investigated aspirin prophylaxis during pregnancy for women at high risk of preeclampsia.¹ Each woman’s risk of preeclampsia was defined as high (previous preeclampsia, chronic or gestational diabetes, chronic hypertension, renal disease, autoimmune diseases, or multiple gestation) or moderate (primigravid, advanced maternal age [≥35 years], pregnancy interval ≥10 years, pre-pregnancy BMI ≥35 kg/m², or preeclampsia in a first-degree relative). Most of the trials used an aspirin dose of 60 or 100 mg, but doses ranged from 0.5 mg/kg to 150 mg daily. Most trials started aspirin at 12 to 16 weeks’ gestational age and discontinued aspirin at delivery, but some stopped aspirin at 35 weeks’ gestational age or once preeclampsia developed.

Low-dose aspirin reduced the incidence of preeclampsia, preterm birth, and intrauterine growth restriction (IUGR) compared with placebo (see TABLE). No statistically significant differences were noted in perinatal mortality, placental abruption, postpartum hemorrhage, or neonatal intracranial hemorrhage. The optimal dose of aspirin was not identified because secondary comparisons between different doses were underpowered. There was evidence of small-study effects, based on funnel plot asymmetry, in the preterm birth and IUGR outcomes, which means that future results could reduce the estimated benefit.¹

The 2014 US Preventive Services Task Force statement on prevention of preeclampsia recommended the use of low-dose aspirin (81 mg/d) after 12 weeks of gestation in women who are at high risk for preeclampsia (Grade B, moderate certainty of substantial net benefit).²

The 2013 task force recommendation from the American College of Obstetricians and Gynecologists recommended initiating daily low-dose (60–80 mg) aspirin in the late first trimester for women with a medical history of early-onset preeclampsia and preterm delivery or preeclampsia in more than 1 pregnancy (Quality of Evidence: Moderate Strength of Recommendation: Qualified).³

John Paik-Tesch, MD
Theresa Schill, MD
Aimi Rughani, MD
Mercy Medical Center, UC Davis Family Medicine Network
Merced, CA


Maternal and neonatal outcomes with aspirin prophylaxis versus placebo in women at risk of preeclampsia¹

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Trials</th>
<th>N</th>
<th>RR</th>
<th>95% CI</th>
<th>RD</th>
<th>NNT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preeclampsia</td>
<td>13</td>
<td>12,184</td>
<td>0.76</td>
<td>0.62–0.95</td>
<td>−1.9%</td>
<td>53</td>
</tr>
<tr>
<td>Preterm birth</td>
<td>10</td>
<td>11,779</td>
<td>0.86</td>
<td>0.76–0.98</td>
<td>−2.7%</td>
<td>38</td>
</tr>
<tr>
<td>IUGR</td>
<td>13</td>
<td>12,504</td>
<td>0.80</td>
<td>0.65–0.99</td>
<td>−0.9%</td>
<td>111</td>
</tr>
</tbody>
</table>

IUGR=intrauterine growth restriction; NNT=number needed to treat; RD=risk difference; RR=relative risk.

How effective is circumcision in reducing recurrent UTIs in boys with prior UTI?

Evidence-Based Answer
For boys with a history of urinary tract infection (UTI) or high-grade (grade 3 or higher) vesicoureteral reflux, circumcision will prevent 87 to 261 UTIs per 1,000 boys. Circumcision will also likely reduce the risk of recurrent UTIs in premature infants (SOR: B, meta-analysis of RCT, cohort and case-control trials, and individual retrospective study).

In 2005, a meta-analysis of 12 studies (1 RCT, 4 cohort, 7 case-control) examined the effect of circumcision on the risk of UTI in boys.1 These 12 studies provided data on 402,908 children and 1,953 separate episodes of UTI. Most of the studies were of infants, 1 included adults, and 4 others included boys beyond the first year of life.

Compared with uncircumcised children, circumcision was associated with a reduced risk of UTI (OR 0.13; 95% CI, 0.08–0.20). The single RCT (n=70) did not show a reduced risk (OR 0.13; 95% CI, 0.01–2.6); however, the 4 cohort studies (n=400,700) showed a reduced risk (OR 0.13; 95% CI, 0.07–0.23), as did the 7 case-control studies (n=2,138; OR 0.13; 95% CI, 0.07–0.23). The TABLE shows the benefit versus harm for circumcision in preventing UTI in boys at different levels of risks for UTI. This table was constructed based on older literature showing 1% of boys experience a UTI within their first 10 years of life, a UTI rate of 0.5% in the uncircumcised group, a 2% surgical complication rate, and a UTI recurrence rate of 10% in normal children and 30% in children with grade 3 or higher vesicoureteral reflux.1

A 2000 retrospective chart review of male preterm infants admitted to the neonatal intensive care unit examined the frequency and potential prevention of recurrent UTI with circumcision in these hospitalized infants.2 Among the 744 infants admitted, 38 infants had 53 episodes of UTI. Three infants >37 weeks old were excluded from the analysis. The remaining 35 infants were <37 weeks old and divided into 2 groups based on number of UTIs: A1 (1 UTI, n=24) and A2 (>1 UTI, n=11).

The frequency of UTI in all (A1 and A2) infants was 12.1%. Of these 35 infants, 11 with a first UTI developed a second UTI, and 4 developed a third UTI. All recurrent UTIs occurred in uncircumcised premature infants >60 days of age. No UTIs recurred after circumcision. Premature uncircumcised males had an increased risk for UTI compared with circumcised infants (OR 11; 95% CI, 3.3–29).2

Lauren Alderson, MD
Chih-Wen Shi, MD, PhD
Naval Hospital Camp Pendleton FMR
Camp Pendleton, CA

The opinions and assertions contained herein are those of the authors and are not to be construed as official or as reflecting the views of the US Navy Medical Department, the US Navy at large, or the US Department of Defense.


<table>
<thead>
<tr>
<th>Patient group</th>
<th>Risk of UTI</th>
<th>Number of UTIs</th>
<th>Number of complications</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>If uncircumcised</td>
<td>If circumcised</td>
<td>Prevented by circumcision</td>
</tr>
<tr>
<td>Normal</td>
<td>1%</td>
<td>10</td>
<td>1</td>
</tr>
<tr>
<td>Past UTI</td>
<td>10%</td>
<td>100</td>
<td>13</td>
</tr>
<tr>
<td>High-grade VUR</td>
<td>30%</td>
<td>300</td>
<td>39</td>
</tr>
</tbody>
</table>

UTI=urinary tract infection; VUR=vesicoureteral reflux.

INTERESTED IN SUBMITTING A LETTER TO THE EDITOR?
Visit www.fpin.org/letters or email ebp@fpin.org for more information.
Is cytisine effective for smoking cessation?

**Bottom line**
Cytisine is up to 3 times more effective than placebo for smoking abstinence (SOR: A, meta-analysis of RCTs) and noninferior to nicotine replacement therapy (SOR: B, single noninferiority trial). Sustained quit rates decline from 40% at 1 month to 8.4% at 12 months.

**Evidence summary**
A 2013 meta-analysis of 7 RCTs (N=4,020) compared cytisine versus placebo for smoking cessation. Standard cytisine dosing is a tapered regimen of 1.5-mg tablets taken over 25 days during which patients stop smoking by day 3 to 5 (see **TABLE**).

Cytisine, either alone or combined with varying levels of behavioral support, was more effective than placebo with or without behavioral support, for smoking abstinence at follow-up times ranging from 26 days to 2 years (RR=1.6; 95% CI, 1.4–1.8). A subgroup analysis including the 2 highest quality RCTs (n=911) with intention-to-treat analysis and verification of smoking status biochemically showed cytisine was more effective than placebo for smoking abstinence at 6 months (RR 3.3; 95% CI, 1.8–5.9).

Treatment with cytisine caused more gastrointestinal adverse events (AEs) than placebo, but there was no significant difference in overall AEs. Trials ranged from low to high quality. Lower quality studies had no validation of abstinence, unclear blinding, no placebo, and lack of concealed allocation.

The longest study with validation of abstinence in the above meta-analysis was a double-blinded RCT of 740 adults who smoked 10 or more cigarettes per day and were interested in quitting. Patients were randomized to cytisine or placebo for 25 days. Complete data were available for 77% of patients.

At 12 months, biochemically verified smoking abstinence was 8.4% in the cytisine group versus 2.4% (P≤.001) in the placebo group (NNT=17).

A 2014 noninferiority RCT included 1,310 adult daily smokers recruited through a national quit line randomized to standard cytisine dosing for 25 days or nicotine replacement therapy (NRT) with patches, gum, or lozenges for 8 weeks. The type and strength of NRT was decided by participant preference and quit line advisors in accordance with national smoking-cessation guidelines. Both groups were offered low-intensity telephone behavioral support.

One month after the set quit date, continuous self-reported abstinence rates in the cytisine group were 40% compared with 31% in the NRT group (risk difference [RD] 9.3%; 95% CI, 4.2–14.5; NNT=11). Secondary outcomes in this study included continuous abstinence rates at 2 months (31% for cytisine vs 22% for NRT; RD 9%; 95% CI, 4.3–13.8; NNT=11) and 6 months (22% for cytisine vs 15% for NRT; RD 6.6%; 95% CI, 2.4–10.8; NNT=14).

**TABLE**

<table>
<thead>
<tr>
<th>Day</th>
<th>Dosing frequency</th>
<th>Total no. of daily tablets</th>
</tr>
</thead>
<tbody>
<tr>
<td>1–3</td>
<td>1 tab every 2 hours</td>
<td>6</td>
</tr>
<tr>
<td>4–12</td>
<td>1 tab every 2.5 hours</td>
<td>5</td>
</tr>
<tr>
<td>13–16</td>
<td>1 tab every 3 hours</td>
<td>4</td>
</tr>
<tr>
<td>17–20</td>
<td>1 tab every 4 hours</td>
<td>3</td>
</tr>
<tr>
<td>21–25</td>
<td>1 tab every 6 hours</td>
<td>2</td>
</tr>
</tbody>
</table>

Sarah Kirchhoff, MD
Laura Morris MD, MSPH
University of Missouri Family Medicine
Columbia, MO

**REFERENCES**

**EVIDENCE-BASED PRACTICE**

**Feature Articles** allow residency and fellowship programs throughout the FPIN Network to showcase their unique area of excellence.

Does your program have a specialty that you would like to publish in EBP?

Can you commit to a quarterly publication schedule?

We want to hear from you!

Apply now for the 2015-2016 academic year by contacting Senior Editorial Assistant Adelina Colbert at adelina@fpin.org.
FPIN is recruiting for NEW PURL Jam Sites!

PURL Jam Sites are responsible for reviewing potentially practice-changing literature 6 to 12 times per year. This information is disseminated through PURL articles (published in The Journal of Family Practice and PURLs Journal Club toolkits).

**QUALIFICATIONS**

Department with proven ability to create family medicine scholarship, meet tight deadlines, and productively engage with the PURLs editorial team.

5-7 committed faculty, 1 local editor, and 1 EBM champion.

**Take your faculty to the next level of EBM**

- Work with a committed team of PURL authors, reviewers, nominators, & editors
- Encourage elite faculty development in EBM
- Publications in *JFP* and *EBP*
- Positively influence family medicine on a national level

Contact PURLs@fpin.org for more information and to schedule a meeting with the Editor-In-Chief.
Is single-dose fosfomycin an effective treatment for uncomplicated UTIs?

Evidence-Based Answer
Yes, single-dose fosfomycin is an effective treatment for uncomplicated urinary tract infections (UTIs) compared with more commonly prescribed antibiotics (SOR: B, RCT, case-control study, and case series).

An RCT compared the efficacy of a single 3-g dose of fosfomycin with a 7-day course of nitrofurantoin (100 mg daily) for 749 female patients with UTIs. E. coli was the pathogen in 84% of the fosfomycin group and 79% of the nitrofurantoin group. Of the E. coli isolates, 99% were susceptible to fosfomycin and 90% were susceptible to nitrofurantoin.

At 5 to 11 days, fosfomycin cured 216 (82.1%) patients and nitrofurantoin cured 206 (84.1%) patients (P=.3). Bacteriologic cure rates, measured 6 weeks after therapy, were 60% for fosfomycin and 59% for nitrofurantoin (P value not reported). Relapse or reinfection at 6 weeks after the last day of therapy occurred in 27 of 246 (10.9%) patients in the fosfomycin group and 40 of 219 (18.2%) patients in the nitrofurantoin group (P=.08).

A case-control study evaluated the efficacies of a single 3-g dose of fosfomycin and amoxicillin/clavulenate 500 mg/125 mg every 8 hours for 5 to 7 days in 73 patients (63% >60 years old, 77% female) with complicated and uncomplicated cystitis due to an extended-spectrum beta lactamase (ESBL)-producing E. coli infection. E. coli susceptibility was comparable to other first-line agents for treatment of UTI such as nitrofurantoin (100%), TMP-SMX (87%), and fluoroquinolones (97%; no P value given for any comparison).

Single-dose fosfomycin was listed in recent guidelines of the Infectious Diseases Society of America as an appropriate treatment for uncomplicated UTIs. Noted benefits included one-time dosing, pregnancy category B, and no need for renal adjustments. (Editor’s note: My quick Internet search found a single dose of fosfomycin costs $70–$90.)

Danielle N. Rhyne, PharmD, BCPS
Joseph J. Saseen, PharmD, BCPS
University of Colorado Skaggs School of Pharmacy and Pharmaceutical Sciences, Aurora, CO

No difference was noted in median pain reduction between groups: in the morning, the acupuncture group was +1/100 point higher (P=.288) and in the evening –5/100 points lower (P=.483) than the sham group.²

An unblinded RCT assigned 386 pregnant women (at 12–31 weeks’ gestation) with pelvic-girdle pain to 3 groups: standard treatment/usual care (ST), ST with acupuncture at 10 body points for 30-minute treatments twice weekly for 6 weeks (ST+Ac), or ST with stabilizing exercises (ST+Ex).³

Only nighttime self-rated pain (100-point rating scale) showed a significant difference between groups in change from baseline. Both exercise and acupuncture reduced pain significantly more than ST alone: ST+Ac versus ST (median difference [Md] between groups –27; 95% CI, –13.3 to –29.5), ST+Ex versus ST (Md –13; 95% CI, –2.7 to –17.5). ST+Ac reduced pain more than ST+Ex (Md –14; 95% CI, –18.1 to –3.3). The lack of a sham treatment group may have biased results.³

An unblinded RCT randomized 72 pregnant women (at 24–37 weeks’ gestation) with pelvic and back pain to acupuncture or usual care.⁴ Acupuncture at 3 to 10 body points, 1 to 2 times weekly for 3 weeks, resulted in greater reductions from baseline in self-reported pain (60% vs 40% improvement on 3-point rating scale; OR 8.8; 95% CI, 2.48–33.25). The lack of a sham treatment group and a significant difference between groups at baseline may have biased results.⁴

An unblinded RCT assigned 60 pregnant women (<32 weeks’ gestation) with low back and pelvic pain to acupuncture or physiotherapy.⁵ Acupuncture at 2 to 10 ear and body sites for 30 minutes, 10 times in 1 month resulted in greater reduction of self-reported pain from baseline than 10 physiotherapy treatments for morning pain (MD –1.4 on an 11-point rating scale; P<.02) and evening pain (MD –2.8; P<.01). The lack of a sham treatment group and a 40% attrition rate from the physiotherapy group may have biased results.⁵

Evidence-Based Answer
Pooled together, probiotics improve irritable bowel syndrome (IBS) symptoms compared to placebo with an NNT of 7 (SOR: A, meta-analyses of RCTs). Lactobacillus plantarum and species in the genus Escherichia reduce the risk of persistent symptoms. Streptococcus faecium may also reduce the risk of persistent symptoms. Combination products containing Bifidobacterium infantis, Lactobacillus casei, and Lactobacillus plantarum or Bifidobacterium longum, Lactobacillus acidophilus, Lactobacillus bulgaricus, and Streptococcus salivarius ssp thermophilus decrease distension and flatulence. Products containing Bifidobacterium breve, B longum, or L acidophilus improve abdominal pain, distention, and flatulence (SOR: B, meta-analyses of small RCTs).

A 2014 systematic review and meta-analysis of 35 RCTs (N=3,452) studied the efficacy of probiotics compared with placebo for management of IBS.¹ Patients were 16 years old or older and were predominantly female with a diagnosis of IBS, made by either clinician opinion or diagnostic criteria such as the Rome criteria.

Probiotic preparations were either single species or combinations of 3 to 4 species and data were pooled by probiotic genera with some pooling of data for individual species. Global IBS symptom scores and abdominal pain scores from the studies were converted to a dichotomous outcome of either persistent or improved IBS symptoms. Clinical outcome scores were also combined and reported as a standardized mean difference (SMD).¹

Taken as a whole, probiotics improved global IBS symptoms (24 trials, n=2,001; SMD –0.25; 95% CI –0.36 to –0.14) and reduced persistent symptoms (23 trials, n=2,575; RR 0.79; 95% CI 0.70–0.89; NNT=7) compared with placebo. Results for specific probiotic regimens shown to be effective (see TABLE 1) and regimens not shown to be effective are also listed (see TABLE 2). Risk of bias in many trials was unclear because of incomplete reporting of randomization and allocation concealment. Many

Susan C. Weller, PhD
Kyu Jana, MD
Laura Porterfield, MD
Monica Prado, MD
University of Texas Medical Branch
Galveston, TX

Evidence-based practice analyses were limited by small study size so power to detect a difference was limited.

A 2013 systematic review and meta-analysis of 24 RCTs (N=2,009) evaluating probiotics for IBS included many of the same studies as the meta-analysis above, but had more stringent study inclusion criteria for IBS diagnosis, including only trials using the Rome criteria. Patients were 18 years old or older and predominantly women.

The included studies evaluated 20 different species of probiotics either as single agents or in combination with other species. For each individual species, the meta-analyses pooled data from every study containing that species regardless of whether it was used alone.

### TABLE 1

<table>
<thead>
<tr>
<th>Probiotic preparation</th>
<th>Treatment protocol</th>
<th>No. of trials</th>
<th>No. of patients</th>
<th>Outcome</th>
<th>Results</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lactobacillus plantarum DSM 9843¹</td>
<td></td>
<td>3</td>
<td>314</td>
<td>Persistence of symptoms</td>
<td>RR 0.67</td>
<td>0.51–0.87</td>
</tr>
<tr>
<td></td>
<td>a) 1 x 10¹⁰ cfu daily for 4 weeks</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>b) 1 x 10¹⁰ cfu BID for 4 weeks</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>c) 2 x 10¹⁰ cfu daily for 4 weeks</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Streptococcus faecium¹</td>
<td>Paraghurt (S. faecium) for 4 weeks</td>
<td>1</td>
<td>54</td>
<td>Persistence of symptoms</td>
<td>RR 0.72</td>
<td>0.53–0.99</td>
</tr>
<tr>
<td>Escherichia genus¹</td>
<td></td>
<td>2</td>
<td>418</td>
<td>Persistence of symptoms</td>
<td>RR 0.86</td>
<td>0.79–0.93</td>
</tr>
<tr>
<td></td>
<td>a) 1.1–3.4 x 10⁷ cfu TID for 1 week, then 2.2–6.8 x 10⁷ cfu TID for weeks 2–8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>b) 2.5–25 x 10⁷ cfu daily for 4 days then BID for 12 weeks</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bifidobacterium breve²</td>
<td>B breve containing probiotic preparation used for 1–6 months</td>
<td>3</td>
<td>154</td>
<td>Abdominal pain</td>
<td>SMD –0.34</td>
<td>–0.66 to –0.02</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Abdominal distention</td>
<td></td>
<td></td>
<td></td>
<td>SMD –0.45</td>
<td>–0.77 to –0.13</td>
</tr>
<tr>
<td></td>
<td>Flatulence</td>
<td></td>
<td></td>
<td></td>
<td>SMD –0.42</td>
<td>–0.75 to –0.10</td>
</tr>
<tr>
<td>B. longum²</td>
<td>B longum containing probiotic for 4–8 weeks</td>
<td>4</td>
<td>202</td>
<td>Abdominal pain</td>
<td>SMD –0.48</td>
<td>–0.91 to –0.06</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L. acidophilus²</td>
<td>L acidophilus containing probiotic for 4–8 weeks</td>
<td>6</td>
<td>328</td>
<td>Abdominal pain</td>
<td>SMD –0.31</td>
<td>–0.61 to –0.01</td>
</tr>
<tr>
<td>B. infantis, L. casei, and L. plantarum²</td>
<td>Combination used for 8 weeks</td>
<td>2</td>
<td>73</td>
<td>Abdominal distention</td>
<td>SMD –0.53</td>
<td>–1.0 to –0.06</td>
</tr>
<tr>
<td>B. infantis, L. casei, and L. plantarum²</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>SMD –0.60</td>
<td>–1.1 to –0.13</td>
</tr>
<tr>
<td>B. longum, L. acidophilus, L. bulgaricus, and S. salivarius ssp thermophilus²</td>
<td>Combination used for 4–8 weeks</td>
<td>3</td>
<td>102</td>
<td>Flatulence</td>
<td>SMD –0.61</td>
<td>–1.0 to –0.21</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

cfu=colony forming units; RR=relative risk; SMD=standard mean difference.

Note: An SMD of 0.2 is considered small, 0.6 moderate, 1.2 large, and 2.0 very large.

### TABLE 2

<table>
<thead>
<tr>
<th>Probiotic preparation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lactobacillus paracasei ssp. paracasei F1⁹¹</td>
</tr>
<tr>
<td>L. acidophilus La5¹</td>
</tr>
<tr>
<td>Bifidobacterium lactis Bb12¹</td>
</tr>
<tr>
<td>Streptococcus salivarius ssp. thermophilus³</td>
</tr>
<tr>
<td>S. boulardii²</td>
</tr>
<tr>
<td>VSL #3 (containing 1 strain of S. salivarius ssp. thermophilus, 3 strains of Bifidobacterium, and 4 strains of Lactobacillus)⁴</td>
</tr>
<tr>
<td>Combination of L. paracasei ssp. paracasei F19/L. acidophilus La5/B. lactis Bb12³</td>
</tr>
<tr>
<td>Combination of B. lactis DN-173 010/S. salivarius ssp. thermophilus/L. bulgaricus⁴</td>
</tr>
<tr>
<td>Combination of B. animalis/B. longum/L. acidophilus/S. boulardii/L. bulgaricus/B. animalis/B. infantis/L. casei/L. plantarum/L. bulgaricus²</td>
</tr>
<tr>
<td>Combination of L. bulgaricus and S. salivarius ssp. thermophilus²</td>
</tr>
</tbody>
</table>
or as part of a combination. Different symptom scales were used in the included studies so pooled results were reported as SMDs.²

Results for specific probiotic regimens shown to be effective compared with placebo were reported (see TABLE 1) and regimens not shown to be effective were also listed (see TABLE 2). Most trials evaluated different combination products, but meta-analysis was performed on individual species from these combination products leading to heterogeneity. Quality assessment using the Jadad scale revealed most studies were of good quality.

Olusola Adegoke, MBBS, MPH
Thomas Satre, MD
University of Minnesota/Saint Cloud Hospital FMR
Saint Cloud, MN