March 2022

EVIDENCE-BASED PRACTICE

A Peer-Reviewed Journal of the Family Physicians Inquiries Network

March 2022
Volume 25 | Number 3

EVIDENCE-BASED PRACTICE
Volume 25 | Number 3

FPIN
Family Physicians Inquiries Network

FPIN envisions a primary care workforce that thinks critically, communicates expertly, and utilizes the best current evidence to improve the health of patients.

TABLE OF CONTENTS

VOL 25 | NO. 3 | March 2022

EVIDENCE-BASED PRACTICE

A PEER-REVIEWED JOURNAL OF THE FAMILY PHYSICIANS INQUIRIES NETWORK

IN DEPTH

DIVING FOR PURLs

GEMS

EBM ON THE WARDS

HELPDESK ANSWERS

Evidence-Based Practice, (ISSN: 2473-3717 [online]), is published monthly online on behalf of the Family Physicians Inquiries Network, Inc., by Wolters Kluwer Health, Inc., at 1800 Dual Highway, Suite 201, Hagerstown, MD 21740-6636. Business and production offices are located at Two Commerce Square, 2001 Market St., Philadelphia, PA 19103.

Postmaster: Send address changes to Evidence-Based Practice, PO Box 1610, Hagerstown, MD, 21741, USA.

TABLE OF CONTENTS *Continued*

EVIDENCE-BASED PRACTICE

A PEER-REVIEWED JOURNAL OF THE FAMILY PHYSICIANS INQUIRIES NETWORK

Publication Staff

Paige Smith, BS
FPIN
Columbia, MO

STATEMENT OF PURPOSE

Evidence-Based Practice (EBP) addresses important patient care questions asked by practicing family physicians, using the best sources of evidence in a brief, clinically useful format. Our goal is to instruct our authors on how to write peer-reviewed scholarly research for the medical and scientific community.

EDITORIAL POLICY

Statements and opinions expressed in articles and communications in this journal are those of the author(s) and not necessarily those of the editor, publisher, or any organizations endorsing this journal. The Publisher and editors of EBP do not endorse any methods, products, or ideas mentioned in the journal, and disclaim any liability which may arise from any material herein. Unless noted, authors have reported no competing interests and have nothing to disclose.

DISCLOSURE

The PURLs Surveillance System is supported in part by Grant Number UL1RR024999 from the National Center for Research Resources, a Clinical Translational Science Award to the University of Chicago. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Center for Research Resources or the National Institutes of Health.

Does early physical therapy result in cost savings for patients with acute low back pain?

EVIDENCE-BASED ANSWER

No clear benefit of early physical therapy (PT) on overall health care costs was noted for patients with acute low back pain (LBP) when compared with no PT (SOR: **A**, consistent evidence from randomized controlled trial and retrospective cohort). However, for patients who do receive PT for acute LBP, therapy started within 14 days is associated with lower overall costs (SOR: **A**, systematic review of large cohorts and single retrospective cohort).

DOI 10.1097/EBP.0000000000001422

A 2017 randomized controlled trial (RCT; n=220) investigated the total health care costs of early physical therapy (PT) versus usual care in adult patients who presented to a primary care provider with a current acute nonradicular low back pain (LBP) episode of less than 16 days.[1] Those with LBP of 16 days or more, radicular symptoms within 72 hours of presentation, and clinical signs of nerve root compression were excluded. Patients were randomly assigned to receive early PT (n=107) or usual care (n=113). The usual care group received only patient education while the treatment group received education and four PT sessions over a three-week span. Education involved providing both a book and a handout that emphasized favorable prognosis of LBP, importance of maintaining activity levels, and the limited utility of imaging such as x-ray or magnetic resonance imaging. Education was provided to all study participants before randomization. The first PT session was initiated within three days of initial evaluation. Health care utilization related to LBP information was collected monthly via online self-reporting of provider visits, procedures or tests, or medication specific to LBP. After one year of follow-up, patients treated in the early PT group had significantly higher overall costs related to their LBP compared with the usual care–only group (mean difference [MD] +$980; 95% CI, $175–$984).

A 2016 systematic review (N=564,921) of two RCTs and 10 nonrandomized and cohort studies assessed cost and medical utilization of delayed versus early PT in patients with acute and subacute LBP.[2] Patients were at least 15 years old with a nonspecific diagnosis of LBP. Significant variability was noted in the definition of early PT in the studies overall. A subanalysis of two large

TABLE. Adjusted annual probability of service use and average costs for 6,668 patients with acute LBP by timing of physical therapy initiation[3,a]

Cost	Immediate (3 d)	Early (4–14 d)	Difference of early vs immediate	Delayed (15–28 d)	Difference of delayed vs immediate	Late (29–90 d)	Difference of late vs immediate
Health care costs incurred over 1-y follow-up period, $, mean							
Pain medication	115	143	28[c]	186	70[b]	223	107[b]
Advanced imaging	550	618	68[d]	601	51	634	84[c]
LBP-related medical	2,746	2,973	226	4,068	1,321[b]	6,067	3,320[b]
Total LBP-related	2,901	3,135	233	4,283	1,615[b]	6,366	3,464[b]
Non–LBP-related	10,066	10,566	500	12,709	2,643[b]	10,891	825

[a] Covariates in the multivariable generalized linear models included age group and sex. [b] $P<.001$. [c] $P<.01$. [d] $P<.05$. LBP=low back pain.

retrospective cohorts (N=124,800) examining military and nonmilitary patients with LBP of six months or less who underwent PT within 90 days of initial visit was identified. Early PT was defined as initiation within 14 days and delayed PT between 15 and 90 days. Patients enrolled in early PT saved significantly more cost per pain episode than delayed PT in both the non-military (MD $2,736; 95% CI, $1,811–$3,662) and the military population (MD $1,202; 95% CI, $1,143–$1,262). The authors concluded that early PT resulted in cost savings without significant differences in additional utilization, function, or quality of life.

A 2018 large retrospective cohort study of patients seen for acute LBP (n=46,916) investigated the impact of receiving PT and the timing of PT on health care utilization and costs.[3] Patients were 17 years old or older (54% female) who were all given a diagnosis of LBP by a physician. Most patients did not receive PT (n=40,246). Differing costs measured were pain medication, imaging, LBP-related costs, total LBP-related costs, and non–LBP-related costs at one-year follow-up. Patients not receiving PT had significantly lower average imaging costs ($82 vs $245, $P<.0001$) and LBP-related costs ($772 vs $3,469, $P<.0001$). However, pain medication costs were higher for the non-PT group compared with the PT group ($140 vs $102, $P<.0001$). A subanalysis of patients who did receive PT were analyzed according to timing of PT as immediate (3 days or less), early (4–14 days), delayed (15–28 days) and late (29–90 days). In summary, among the patients who did receive PT, those who started immediately had consistently lower health care utilization and cost measures (see **TABLE**). EBP

Andrew Cunningham, MD
Mustafa Hussein, MD
Simin Emadzadeh, MD
St. Joseph Mercy Livingston Family Medicine Residency, Brighton, MI

The authors declare no conflicts of interest.

References

1. Fritz J, Kim A, Magel J, Asche C. Cost-effectiveness of primary care management with or without early physical therapy for acute low back pain. *Spine (Phila Pa 1976)*. 2017; 42(5):285–290. [STEP 2]
2. Ojha H, Wyrsta N, Davenport T, Egan W, Gellhorn A. Timing of physical therapy initiation for nonsurgical management of musculoskeletal disorders and effects on patient outcomes. *J Orthopaedic Sports Phys Ther*. 2016;46(2):56–70. [STEP 1]
3. Liu X, Hanney W, Masaracchio M, et al. Immediate physical therapy initiation in patients with acute low back pain is associated with a reduction in downstream health care utilization and costs. *Phys Ther*. 2018;98(5):336–347. [STEP 2]

Efficacy and safety of gabapentin in the treatment of vasomotor symptoms

Shan D, Zou L, Liu X, Shen Y, Cai Y, Zhang J. Efficacy and safety of gabapentin and pregabalin in patients with vasomotor symptoms; a systematic review and meta-analysis. *Am J Obstet Gynecol*. 2020;222(6):564-579.e12. doi:10.1016/j.acog.2019.12.011.

DOI 10.1097/EBP.0000000000001357

This meta-analysis examines the safety and efficacy of gabanoids for the treatment of vasomotor symptoms (aka "hot flashes") in postmenopausal female patients. Nineteen prospective randomized controlled trials (RCTs) and two randomized crossover studies were included (N=3,519). The effectiveness and safety of oral gabapentin (300–2,400 mg daily) or oral pregabalin (150–300 mg daily) was compared with placebo as well as numerous other pharmaceuticals, supplements, behavioral therapies, and acupuncture. Primary outcomes included reduction of vasomotor symptom frequency, severity, and duration from baseline. Secondary outcomes included adverse events related to gabapentin and pregabalin. Intervention periods ranged from four weeks to 24 weeks. Compared with placebo or vitamin E, gabapentin reduced hot flash frequency from baseline (mean difference [MD] –1.62 per week; 95% CI, –1.98 to –1.26 after four weeks, MD –2.77 per week; 95% CI, –4.29 to –1.24 after 12 weeks, and MD –1.43; 95% CI, –2.56 to –0.29 after 24 weeks). No significant difference was noted at eight weeks. Overall level of evidence was assessed as low to moderate quality by the GRADE criteria. Evidence for gabapentin reducing symptom severity and duration when compared with placebo or vitamin E was insignificant. Two RCTs (N=139) compared estrogen with gabapentin. Estrogen was more effective than gabapentin in reducing hot flash frequency (MD 1.11 per week; 95% CI, 0.69–1.52). Estrogen was more effective than gabapentin in reducing hot flash frequency (MD –1.11 per week; 95% CI, –1.52 to –0.69) and hot flash severity score (standardized mean difference –0.50; 95% CI, –0.85 to –0.14). No difference was noted between gabapentin and antidepressants in reducing hot flash severity score. The effectiveness of pregabalin was limited and unable to be meta-analyzed. Adverse events reported with gabapentin included dizziness, somnolence, headache, fatigue, insomnia, weight gain, gastrointestinal disturbance, visual change, tremor, and nausea. Compared with placebo, gabapentin participants suffered more from dizziness (risk ratio [RR] 4.45; 95% CI, 2.50–7.94; number needed to harm [NNH]=10) and somnolence (RR 3.29; 95% CI, 1.97–5.48; NNH=18) and the medication had elevated dropout rates (RR 1.99; 95% CI, 1.50–2.62; NNH=17). Side effects seemed to be dose dependent. No difference in adverse events was found in the comparison of gabapentin with estrogen.

Methods

This article was identified as a potential PURL through the standard systematic methodology that has been described here (https://journals.lww.com/ebp/Documents/PURLs%20Methods%20AC.pdf). An additional literature search was conducted by searching American Academy of Family Physicians, National Institute of Health and North American Menopause Society with the terms "nonhormonal treatment of vasomotor symptoms" to find additional literature to place this research into the context of current clinical practice.

Does this meet PURL criteria?			
Relevant	Yes	Medical care setting	Yes
Valid	Yes	Implementable	Yes
Change in practice	No	Clinically meaningful	Yes

Bottom line: In postmenopausal female patients, very low to moderate evidence exists that gabapentin reduces frequency of hot flashes; however, this does not represent a novel approach. Compared with estrogen replacement therapy, gabapentin is less effective in reducing frequency and severity of symptoms although it may serve as an important alternative in women with contraindications to estrogen therapy. Unfortunately, the number needed to treat data for benefit cannot be calculated from any available study data. Yet, NNHs for side effects can be noted, with the most common being dizziness (NNH of 10) and somnolence (NNH of 18).

Kimberly W. Roman, MD
Michael Brisson, DO
South Baldwin Regional Medical Center FMR, Foley, AL

The authors declare no conflicts of interest.

A validated screening tool for penicillin allergy testing

Trubiano JA, Vogrin S, Chua KYL, et al. Development and Validation of a Penicillin Allergy Clinical Decision Rule. *JAMA Intern Med*. 2020;180(5):745-752.

DOI 10.1097/EBP.0000000000001356

This was a two-phase diagnostic study. Part one was a multicenter, prospective cohort study from two tertiary care sites in Australia that developed and internally validated a penicillin allergy decision rule (N=622). The tool was derived using reverse stepwise logistic regression, with a bootstrap regression model for internal validation. The authors used a standard definition of antibiotic allergy testing (AAT) positivity in all cohorts. Primary outcome is any positive test result with penicillin AAT. Part two was a retrospective cohort study at two sites in Australia and one in the United States to externally validate the tool (N=945).

The derivation cohort included only those patients with a negative skin test result followed by an oral challenge. Thirty patients who reported an allergy to an intravenous penicillin without an oral form were excluded. Of the remaining 622 patients, 9.3% had a positive AAT. Of those with a negative AAT (n=590), 95.6% underwent oral penicillin challenge.

Based on multivariable analysis from all 622 patients, a mnemonic/scoring tool was developed: penicillin allergy, five or fewer years ago (2 points); anaphylaxis/angioedema or severe cutaneous adverse reaction (2 points); treatment required for allergy episode (1 point) (PEN-FAST). Tool interpretation (risk of having a positive penicillin test result): very low (0 points), overall allergy risk 0.6%; low (1 or 2 points), with a 5% risk; moderate (three points), with a 20% risk; and high (4 or 5 points), with a 50% risk probability. A cutoff of less than three points was chosen, and that classified 460 of 622 patients (74.0%) as low risk of true allergy. Of the remaining 162 patients, classified as higher risk, 41 (25.3%) had a positive allergy test result. Using this cutoff (<3 points), sensitivity to identify penicillin allergy was 70.7%; specificity 78.5%; positive predictive value 25.3%; and negative predictive value (NPV) 96.3% (95% CI, 94.1%–97.8%).

External validation trial sites confirmed the high NPV for the PEN-FAST tool with an NPV=95.0% (95% CI, 89.4–98.1; n=334), NPV=84.9% (95% CI, 72.4–93.3; n=80), and NPV=98.4% (95% CI, 96.3–99.5; n=531).

Methods

This article was identified as a potential PURL through the standard systematic methodology that has been described here (https://journals.lww.com/ebp/Documents/PURLs%20Methods%20AC.pdf).

Does this meet PURL criteria?			
Relevant	Yes	Medical care setting	Yes
Valid	Yes	Implementable	Yes
Change in practice	No	Clinically meaningful	Yes

Bottom line: PEN-FAST is a validated tool with a high negative predictive value that could be used at the point of care to allow safe prescribing practices of beta-lactams in patients with reported penicillin allergies who require these antibiotic classes for therapy.

Tyler S. Rogers, MD, FAAFP
Derrick Thiel, MD
Samuel Tiglao, DO, FAAFP
Madigan Army Medical Center, Joint Base Lewis-McCord, WA

The authors declare no conflicts of interest.
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 Army Medical Department, the Army at large, or the Department of Defense.

Getting a head start on concussion recovery

Leddy JJ, Haider MN, Ellis MJ, et al. Early subthreshold aerobic exercise for sport-related concussion: a randomized clinical trial. *JAMA Pediatr*. 2019;173(4):319-325.

DOI 10.1097/EBP.0000000000001377

This multicenter prospective randomized clinical trial compared subsymptom threshold aerobic exercise to placebo-like stretching in 103 adolescent male and female athletes (13–18 years old) presenting within 10

days of a sports-related concussion (SRC). Researchers determined patient exercise tolerance by a treadmill test then randomly assigned participants to aerobic exercise (n=52) or stretching (n=51) for 20 minutes per day. The aerobic exercise group was on a stationary bike or treadmill at home or in a gym under supervision with a monitored heart rate at 80% of the heart rate with symptom exacerbation. The stretching group was given a stretching program and provided resting instructions. Patients reported daily symptoms and completed weekly physician visits. The aerobic exercise group recovered in a median of 13 days (interquartile range [IQR] 10–18.5 days) compared with a median of 17 days (IQR 13–23 days) for the stretching participants (P=.009). The incidence of delayed recovery (>30 days) was higher in the stretching group (n=7) compared with the aerobic exercise group (n=2), but this was not statistically significant. Daily symptom reporting occurred at a high rate in both groups with patients stating they performed their prescribed exercise 89% of the time. No adverse events were noted.

Methods

This article was identified as a potential PURL through the standard systematic methodology that has been described here.

Does this meet PURL criteria?			
Relevant	Yes	Medical care setting	Yes
Valid	Yes	Implementable	Yes
Change in practice	Yes	Clinically meaningful	Yes

Bottom line

Subsymptom threshold exercise not only seems safe but may also help speed recovery compared with stretching in adolescents with SRC.

Leah Stem, MD
UPMC, St. Margaret FMRP, Pittsburgh, PA

The author declares no conflict of interest.

Safety and efficacy of DOACs in morbidly obese patients with atrial fibrillation

Kido K, Shimizu M, Shiga T, Hashiguchi M. Meta-analysis comparing direct oral anticoagulants versus warfarin in morbidly obese patients with atrial fibrillation. *Am J Cardiol*. 2020;126:23-28.

DOI 10.1097/EBP.0000000000001378

Current guidelines from the International Society of Thrombosis and Haemostasis in 2016 recommend warfarin over direct oral anticoagulants (DOACs) as the anticoagulant of choice for patients with body mass index >40 kg/m^2 or weight >120 kg and nonvalvular atrial fibrillation because of lack of efficacy and safety data in this population.[1] This meta-analysis of four retrospective cohort studies and one post-hoc study of a randomized controlled trial compared DOAC and warfarin use in 8,732 adult patients with atrial fibrillation and morbid obesity as described above. Patients were excluded if they had mechanical heart valves, were pregnant, or required dialysis. The primary outcome of stroke or systemic embolism incidence showed no statistically significant difference between the DOAC and warfarin groups (odds ratio [OR] 0.85; 95% CI, 0.60–1.19; P=.35). DOAC use had a significantly lower major bleeding event rate compared with the warfarin group (OR 0.63; 95% CI, 0.43–0.94; P=.02). The majority of the data compared apixaban and rivaroxaban with warfarin, so conclusions on dabigatran and edoxaban cannot be made. This meta-analysis concluded that in morbidly obese patients with atrial fibrillation, DOACs could be considered although a randomized controlled trial is needed to confirm these findings. If using apixaban or rivaroxaban, current guidelines recommend measuring anti-factor Xa levels.[1]

Methods

This article was identified as a potential PURL through the standard systematic methodology that has been

described here https://journals.lww.com/ebp/Documents/PURLs%20Methods%20AC.pdf.

Does this meet PURL criteria?			
Relevant	Yes	Medical care setting	Yes
Valid	Yes	Implementable	Yes
Change in practice	No	Clinically meaningful	Yes

Bottom line

In morbidly obese patients with atrial fibrillation, apixaban and rivaroxaban use can be considered for anticoagulation because some evidence of equal efficacy is present for preventing stroke and systemic embolic events, and fewer bleeding event is noted when compared with warfarin.

Krista Cowan, MD
Steven Fox, MD
Michael Shepherd, MD
University of Tennessee College of Medicine Chattanooga
Chattanooga, TN

The authors declare no conflicts of interest.

Reference

1. Martin K, Beyer-Westendorf J, Davidson BL, Huisman MV, Sandset PM, Moll S. Use of the direct oral anticoagulants in obese patients: guidance from the SSC of the ISTH. *J Thromb Haemost*. 2016;14:1308-1313.

Should you keep injecting knees?

The effect of multiple intra-articular steroid injections on knee cartilage volume loss and pain relief in patients with knee osteoarthritis

McAlindon TE, LaValley MP, Harvey WF, et al. Effect of intra-articular triamcinolone vs saline on knee cartilage volume and pain in patients with knee osteoarthritis: a randomized clinical trial. *JAMA.* 2017; 317(19):1967–1975. doi:10.1001/jama.2017.5283
DOI 10.1097/EBP.0000000000001260

KEY TAKEAWAY: For treatment of symptomatic knee osteoarthritis, corticosteroid injections, compared to saline injections, cause increased cartilage loss. Both have minimal effect on pain relief.

STUDY DESIGN: Randomized, placebo-controlled, double blind trial.

LEVEL OF EVIDENCE: Step 2.

BRIEF BACKGROUND INFO: Pain and synovitis from knee osteoarthritis is commonly treated with corticosteroid injections. For decades, the efficacy and negative consequences of these injections have been questioned, with studies showing conflicting results on the risks and benefits.

PATIENTS: Patients with symptomatic knee osteoarthritis and signs of synovitis on ultrasound.

INTERVENTION: Intra-articular triamcinolone every three months for 24 months.

CONTROL: Intra-articular saline every three months for 24 months.

O: Cartilage volume and structural damage. Osteoarthritis (WOMAC) index for pain.

METHODS BRIEF DESCRIPTION:

- 140 patients, over age 45 (mean age 58), with mild-moderate knee osteoarthritis (by a validated scale classifying joint space narrowing, sclerosis, osteophyte formation, bony end deformity) as diagnosed by x-ray and synovitis on ultrasound.
- Randomized to receive saline or triamcinolone injections every three months for two years with annual MRI to evaluate cartilage depth. Both subjects and investigators were blinded to treatment group.
- Primary outcome was volume of cartilage loss.
- No analgesia use was permitted for 48 hours pre-procedure. Symptoms assessed using WOMAC scale for pain, function, stiffness.
- Initial pain score of eight on 20-point scale (7.5 in triamcinolone group, 8.2 in saline group).

INTERVENTION (# IN THE GROUP): 70 initially, 59 completed.

COMPARISON (# IN THE GROUP): 70 initially, 60 completed.

FOLLOW UP PERIOD: 2 years.

RESULTS: The triamcinolone group had greater cartilage loss, both in the index compartment and overall.

- Total mean cartilage thickness decreased 0.29 mm (triamcinolone) vs 0.13 mm (saline), $P=.04$. Clinical significance of this finding is uncertain.
- Triamcinolone group did not see more cartilage damage (–177.63 vs –82.01 mm^3, $P=.6$)
  No significant differences noted in knee pain (between group difference –0.64 points; 95% CI, –1.6 to 0.29) or in other WOMAC scale components or functional tests.

LIMITATIONS:

1. This study only measured pain every three months. Some studies show steroid injections are most efficacious in first four weeks after injection.
2. Patients continued their usual medications during the trial. This might have attenuated difference between treatment groups.
3. Additionally, this only addressed knees that had US evidence of inflammation, and perhaps that paradoxically does not respond to triamcinolone (as a previous study has found)
4. Did not include patients with severe pain or severe osteoarthritis, nor did it include patients with very mild osteoarthritis. Either of these groups may have had a different response to treatment. EBP

Charles Nielsen, DO
Jessica Bloom-Foster, MD
Eastern Maine Center Medical Program
Bangor, ME

More than meets the eye: The role of social determinants of health in maternal morbidity and mortality in the United States

Social determinants of pregnancy-related mortality and morbidity in the United States: a systematic review

Wang E, Glazer KB, Howell EA, Janevic TM. Social determinants of pregnancy-related mortality and morbidity in the United States: A systematic review. *Obstet Gynecol.* 2020;135:896-915. DOI 10.1097/EBP.0000000000001410

KEY TAKEAWAY: Race/ethnicity, insurance coverage, and maternal education all impact maternal mortality rates. Non-Hispanic Black women experience the most severe maternal morbidity. This systematic review identifies large gaps in the literature addressing health disparities in pregnancy.

STUDY DESIGN: Systematic review of 83 studies (majority retrospective cross-sectional and cohort studies).

LEVEL OF EVIDENCE: STEP 3 (downgraded because of observational study design).

BRIEF BACKGROUND INFO: The rate of maternal mortality has increased over the last 30 years. Few studies have examined the role of social determinants of health (social/environmental conditions where people live and work) on maternal health and pregnancy outcomes.

PATIENTS: Pregnant and postpartum women.

INTERVENTION: Social determinants of health and socioeconomic position indicators.
CONTROL: Patients without the same socioeconomic position indicators.
OUTCOME: Pregnancy-related morbidity and mortality.

METHODS BRIEF DESCRIPTION:
- Examined associations between social determinants and adverse maternal outcomes.
  - Social determinants: Most common socioeconomic position indicators (race, insurance, and education).
  - Maternal outcomes: Pregnancy-related death, severe maternal morbidity, and emergency hospitalization/readmissions.
- Included women all gave birth in the United States.
- Three authors screened the abstracts and two screened full articles for inclusion.
- All outcomes were measured using non-Hispanic White women as a reference.

INTERVENTION # IN THE GROUP: Not available.
COMPARISON # IN THE GROUP: Not available.

FOLLOW-UP PERIOD: Not applicable.

RESULTS:
National- and state-level studies consistently found higher pregnancy-related mortality and severe maternal morbidity risks among Black women when controlling for factors such as insurance coverage, marital status, and medical conditions. The data were not pooled.
- 58 of 67 studies found minority groups experienced higher levels of severe maternal morbidity when compared with non-Hispanic White women:
  - Non-Hispanic Black women: risk ratios (RR) range from 1.2 to 5.
  - Hispanic women: RR range from 1.3 to 3.4.
  - Asian-Pacific Islander women: RR range from 1.2 to 1.6.
  - Native American women: RR range from 1.3 to 1.8.
- 21 of 30 studies found that those with no insurance or Medicaid had worse maternal outcomes than those with private insurance.
- Eight of 12 studies found those with less education had worse maternal outcomes than those with more education.
  - However, mortality was higher for Black women at every education level than White women at any education level.

LIMITATIONS:
Most studies were retrospective cross-sectional or cohort studies using medical chart review, so a lack of evidence is present as to what could be driving the associations. EBP

Te-Ana Harris, MD
LewisGale Family Medicine Residency Program
Roanoke, VA

The authors declare no conflicts of interest.

Psychotherapy delivered via telehealth is comparable to in-person psychotherapy

DOI 10.1097/EBP.0000000000001487

With a global pandemic rapidly changing the way clinical behavioral health services are delivered, a look at the evidence on whether differences are present in treatment outcomes based on treatment delivery modality is indicated.

A comprehensive review of 755 articles between 2003 and 2013 yielded 70 studies on telemental health, of which 15 were retained because they primarily addressed effectiveness. Effectiveness was determined on the basis of "clinical parameters, beneficial effects of a program under optimal conditions of delivery, and other data under more real-world conditions," with a key component being feasibility and adaptability in other settings.[1] The review found that telemental health was effective for diagnosis and assessment across adults, children, geriatrics, and ethnicity, and was comparable with in-person care.[1] They pointed out that telemental health is actually better for some populations because it increases access to care, including consultation in primary care.[1]

One systematic review and one meta-analysis evaluated the treatment of depressive symptoms via telehealth.[2,3] The systematic review of 33 articles included nine studies that evaluated depression as a primary outcome; six examined depression and other symptomology (anxiety, post-traumatic stress disorder); and 18 measured depression as a secondary outcome.[2] The majority of studies (n=23) received a strong or moderate rating on research design (14 randomized controlled studies, 4 quasi-experimental studies, and 15 uncontrolled studies). Twenty-two of these studies had statistically significant reductions in depressive symptoms after telehealth psychotherapy. The rest of the studies indicated no statistical difference between telehealth and in-person delivery or reported inconclusive results.[2] Little demographic information was included in this systematic review other than gender. The second meta-analysis of 14 studies compared telehealth for treatment of depression to standard nontelehealth modalities. The included populations were adult veterans, adult patients from outpatient mental health clinics, and adult patients who sought mental health services in health-related settings. Similarly, no statistically significant difference was found between the modalities (Hedges' g=0.14; 95% CI, −0.03 to 0.30; P=.098; I^2=49.74%).[3]

A systematic review of articles between 2004 and 2014 included measures of treatment satisfaction and therapeutic alliance in video teleconferencing or telephone-based psychotherapy compared with in-person delivery.[4] Fourteen studies (9 were randomized controlled trials and 5 were nonexperimental design) out of 552 met inclusion criteria. The researchers found comparable treatment satisfaction ratings for both modalities. The results in the article were descriptive with no specific statistical analysis reported. The authors noted that some of the results indicated there might be decreased comfort with aspects of group treatment delivered via telemental health.[4] The first comprehensive review discussed above also found that patients were satisfied with a single assessment/consultation or follow-up using telehealth.[1]

Based on the reviews and meta-analysis above, telehealth psychotherapy seems to be equivalent to and as effective as in-person psychotherapy. Many advantages were highlighted, including accessibility and improved outcomes for certain populations. Patient satisfaction with telehealth for psychotherapy was also comparable with in-person psychotherapy. These results are limited by the fact that two of the reviews did not include statistical analysis of their included studies.[1,4] More high-quality studies of the data with attention to demographic and cultural diversity are needed to ensure that these conclusions are accurate and generalizable. EBP

Laurie C. Ivey, PsyD
University of Denver, Denver, CO

The author declares no conflicts of interest.

References

1. Hilty DM, Ferrer DC, Parish MB, Johnston B, Callahan EJ, Yellowlees PM. The effectiveness of telemental health: a 2013 review. *Telemed J E Health*. 2013;19(6):444-454. [STEP 2]

2. Berryhill MB, Culmer N, Williams N, et al. Videoconferencing psychotherapy and depression: a systematic review. *Telemed J E Health.* 2019;25(6):435-446. [STEP 2]
3. Osenbach JE, O'Brien KM, Mishkind M, Smolenski DJ. Synchronous telehealth technologies in psychotherapy for depression: a meta-analysis: research article: meta-analysis of synchronous telehealth for depression. *Depress Anxiety*. 2013;30(11):1058-1067. [STEP 2]
4. Jenkins-Guarnieri MA, Pruitt LD, Luxton DD, Johnson K. Patient perceptions of telemental health: systematic review of direct comparisons to in-person psychotherapeutic treatments. *Telemed J E Health*. 2015;21(8):652-660. [STEP 2]

In patients with chronic hypercapnic respiratory failure from COPD, does use of home BiPAP improve outcomes?

EVIDENCE-BASED ANSWER

Overall, home bilevel positive airway pressure (BiPAP) has inconsistent effects on mortality and risk of intubation in patients with chronic hypercapnia from chronic obstructive pulmonary disease (COPD) and is not associated with improved quality of life or decreased hospitalizations (SOR: **B**, meta-analysis of randomized controlled trials [RCTs] and cohort studies). However, when treatment goal is normal partial pressure of carbon dioxide ($PaCO_2$) levels, BiPAP reduces mortality by 65% (SOR: **A**, meta-analysis of RCTs). Nocturnal BiPAP can be used in patients with chronic COPD and resting $PaCO_2$ more than 45 mmHg with the goal of normalizing $PaCO_2$ levels (SOR: **C**, expert opinion).

DOI 10.1097/EBP.0000000000001247

A 2020 meta-analysis of 21 RCTs and 12 cohort studies evaluated home noninvasive positive pressure ventilation (NPPV) in 51,085 adults with chronic stable COPD and hypercapnia ($PaCO_2$ of 45–56 mmHg) but without obstructive sleep apnea.[1] BiPAP was the mode of NPPV that was compared with home oxygen therapy, exercise programs, or standard care in 15 RCTs and six cohort studies (N=1,700). Patients in these studies received BiPAP for at least 1 month at home with inspiratory pressure settings ranging from 10 to 28 cmH_2O, expiratory pressure settings ranging from 2 to 8 cmH_2O, and rates ranging from 13 to 18 breaths per minute. Daily usage was inconsistently reported but ranged from five to nine hours per day. BiPAP was associated with about 6% lower risk of mortality compared with control interventions over an unreported time frame (22.3% vs 28.6%; 13 studies, N=1,423; weighted risk difference [WRD], 5.5%; 95% CI, –10 to –0.76; number needed to treat [NNT]=19). Similarly, BiPAP was associated with an 8% lower risk of intubation (5.3% vs 14%; three studies, N=267; WRD, –8.0%; 95% CI, –15 to –1.3; NNT=13). However, BiPAP was not associated with better quality of life as measured on various scales with results reported as standardized mean differences (nine studies, N=833; standardized mean difference, 0.16; 95% CI, –0.06 to 0.39). Nor did it decrease the rate of all-cause hospital admission (five studies, N=326; rate ratio, 0.91; 95% CI, 0.71–1.2). In subgroup analysis of only the RCTs, no difference between BiPAP and control was found for any of these outcomes.

In 2017, a meta-analysis of seven RCTs (N=810) compared the efficacy of long-term BiPAP with usual care for the treatment of adults with stable COPD and chronic hypercapnia.[2] The studies in this meta-analysis were all included in the meta-analysis above, but this meta-analysis had more restrictive study inclusion criteria (only RCTs, PaO_2 less than 60 mmHg, and NPPV use of at least five hours per day for at least three months), so it represents higher level evidence. Patients with $PaCO_2$ greater than 50 mmHg and FEV1 less than 50% predicted or less than 1.5 L received BiPAP with inspiratory pressures ranging from 10 to 21.6 cm of H_2O and expiratory pressures from 2 to 5.1 cmH_2O. The control groups received usual care including oxygen therapy. Overall, BiPAP did not decrease overall mortality compared with control (six studies, N=401; risk ratio [RR], 0.78; 95% CI, 0.54–1.1). However, when including only the 310 patients from two trials in whom BiPAP use was driven by the goal of normalizing $PaCO_2$ with an average inspiratory pressure of 20 cmH_2O, BiPAP did decrease mortality compared with control (RR, 0.35; 95% CI, 0.19–0.64). Although meta-analysis was not possible, four of five trials evaluating hospitalization rates and three of four trials evaluating exercise capacity with the six-minute walk distance showed no significant difference between patients on BiPAP versus control. Two of the trials were rated as low risk of bias (including one of the trials with goal of normalizing $PaCO_2$), but the other five trials had unclear concealment of allocation and unclear blinding of outcome assessors.

A 2020 clinical practice guideline from a panel of physician and respiratory experts convened by the American Thoracic Society made recommendations

on long-term noninvasive ventilation in patients with chronic stable COPD and hypercapnia.[3] Recommendations were based on a review of the literature and categorized as strong (patients should receive the therapy) or conditional (different choices are likely to be appropriate and therapy should be tailored to patient's situation). The guideline suggested using nocturnal NPPV in patients with chronic COPD and resting $PaCO_2$ more than 45 mmHg (conditional recommendation). It is also suggested that the target of NPPV be normalization of $PaCO_2$ (conditional recommendation). EBP

Tom Satre, MD
Gamaliel Ferrer, MD
University of Minnesota—CentraCare St. Cloud FMR, St. Cloud, MN

The authors declare no conflicts of interest.

References

1. Wilson ME, Dobler CC, Morrow AS, et al. Association of home noninvasive positive pressure ventilation with clinical outcomes in chronic obstructive pulmonary disease: a systematic review and meta-analysis. *JAMA* 2020; 323(5): 455–465. [STEP 2]
2. Liao H, Pei W, Li H, et al. Efficacy of long-term noninvasive positive pressure ventilation in stable hypercapnic COPD patients with respiratory failure: a meta-analysis of RCTs. *Int J Chron Obstruct Pulmon Dis.* 2017; 12:2977–2985. [STEP 1]
3. Macrea M, Oczkowski S, Rochwerg B, et al. Long-term noninvasive ventilation in chronic stable hypercapnic chronic obstructive pulmonary disease: a clinical practice guideline. *Am J Respir Crit Care Med.* 2020; 202(4): e74–e78. [STEP 5]

Does maternal consumption of probiotics while pregnant decrease the incidence of eczema in infants?

EVIDENCE-BASED ANSWER

Probiotics given to mothers while pregnant and breastfeeding may reduce the risk of atopic dermatitis in children (SOR: **B**, systematic review of moderate- to low-quality randomized controlled trials [RCTs]). However, supplementation in pregnancy alone does not appear to significantly reduce the risk of atopic dermatitis in children (SOR: **B**, RCT).

DOI 10.1097/EBP.0000000000001252

A 2019 meta-analysis of 21 studies (18 RCTs, two cohort studies and one open trial with 31,252 children) examined the effect of probiotic supplementation in pregnancy, breastfeeding, and infancy on the risk of developing atopic dermatitis (AD) in childhood.[1] Their subanalysis assessed eight estimates from five RCTs (total N not provided) evaluating the effect of probiotics given to pregnant and breastfeeding mothers on the outcome of AD incidence between the ages of 6 months and 6 years. There was an overall decreased risk of AD in infants of mothers who received probiotics in pregnancy and while breastfeeding compared with the placebo (risk ratio [RR], 0.72; 95% CI, 0.52–1.0). The authors noted that single strains were not as effective as probiotic mixtures.

A 2015 meta-analysis of 15 RCTs (8 RCTs overlap with the meta-analysis above) with 3,509 patients also examined whether administration of probiotics to pregnant mothers, breastfeeding mothers, or their infants reduced the risk of childhood eczema compared with placebo.[2] Their analysis divided the trials they reviewed by the groups who received the probiotic supplementation, which included pregnant women and their infants (group 1), pregnant women and breastfeeding mothers (group 2), and pregnant women, breast-feeding mothers, and their infants (group 3). Overall, probiotic supplementation reduced the risk of childhood eczema over a follow-up period of 6 to 36 months (RR, 0.72; 95% CI, 0.61–0.85). The 10 RCTs in which probiotics were given only to groups 2 and 3 included 1,595 patients and found a similar reduction in eczema (RR, 0.61; 95% CI, 0.50–0.74).

A 2011 RCT (N=250) that was included in the above two meta-analyses was the only trial which specifically examined the question of probiotic supplementation exclusively during pregnancy.[3] Pregnant women with infants at high risk for allergic disease were randomized to receive either probiotic supplementation with lactobacillus GG or placebo from 36 weeks gestation through delivery. No significant difference was seen between the groups in cumulative incidence of eczema in the first year of life (RR, 0.88; 95% CI, 0.63–1.2). EBP

Haley Ringwood, MD, MPH
Jessica Bull, MD
Susana Peralta, DO
University of Colorado Family Medicine Residency
Denver, CO

The authors declare no conflicts of interest.

References

1. Amalia N, Orchard D, Francis KL, et al. Systematic review and meta-analysis on the use of probiotic supplementation in pregnant mother, breastfeeding mother and infant for the prevention of atopic dermatitis in children. *Australas J Dermatol.* 2020; 61(2):e158–e173. [STEP 1]
2. Cuello-Garcia CA, Grindlay DJC, Williams HC. Probiotics for the prevention of allergy: a systematic review and meta-analysis of randomized controlled trials. *J Allergy Clin Immunol.* 2015; 136(4):952–961. [STEP 1]
3. Boyle RJ, Ismail IH, Kikivuori S, et al. Lactobacillus GG treatment during pregnancy for the prevention of eczema: a randomized controlled trial. *Allergy*. 2011; 66(4):509–516. [STEP 2]

Do early epidurals increase the rate of fetal malposition (or specifically occiput posterior [OP]) in the second stage of labor?

EVIDENCE-BASED ANSWER

Probably not. Early epidural compared with late epidural analgesia in labor does not increase the risk of malposition at delivery. In addition, epidural analgesia overall in labor does not increase risk of malposition compared with opioid analgesia. Epidural analgesia using low concentrations of local anesthesia at any part of labor, including early epidurals, compared with nonepidural analgesia does not increase the risk of cesarean delivery, instrumented vaginal delivery, or length of second stage of labor (SOR: **A**, systematic reviews of consistent randomized controlled trials).

DOI 10.1097/EBP.0000000000001253

A 2014 systematic review of nine randomized controlled trials (RCTs) with 15,752 pregnant women examined the effect of early or late epidurals in labor. A variety of maternal and fetal outcomes were evaluated with early initiation (defined as cervical dilation of less than 4–5 cm) or late initiation (defined as 4–5 cm or more) of epidural. Two of the nine trials with 483 women evaluated the outcome of fetal malposition. There was no significant difference between the rate of malposition at delivery between the early epidural and the late epidural groups (15% vs 15%; risk ratio [RR], 0.99; 95% CI, 0.65–1.51).

A 2018 systematic review of 40 RCTs compared effectiveness and safety of epidural versus nonepidural or no analgesia for pain management in labor in more than 11,000 laboring women. Although the review did not specifically evaluate the effect of epidural timing, many of the included RCTs involved epidurals given before 4- to 5-cm dilation. Four of the RCTs (N=673) specifically addressed malposition in the setting of epidural versus opioids for labor pain management. No significant difference was noted in risk of malposition in the epidural group versus opioid group (18% vs 13%; RR, 1.4, 95% CI, 0.98–2.0). This review did not address the timing of assessment for malposition (during labor or at delivery).

A 2017 meta-analysis of 10 RCTs including 1,809 pregnant women compared five obstetrical outcomes in women receiving epidural analgesia using low concentrations of local anesthetic with women receiving nonepidural analgesia. Many of the included trials evaluated epidurals given early in labor (2–4 cm dilation). No significant difference was observed between the two groups in duration of the second stage of labor (eight trials; N=1,445; mean difference [MD], 5.7 min; 95% CI, –6.1 to 18), rate of instrumented birth (eight trials; N=1,442; RR, 1.5; 95% CI, 0.97–2.4), cesarean section rate (nine trials; N=1,681; RR, 0.8; 95% CI, 0.6–1.1), rate of spontaneous vaginal delivery (six trials; N=1,456; RR, 0.98; 95% CI, 0.91–1.1), or duration of first stage of labor (four trials; N=438; MD, 17 min; 95% CI, –5.9 to 41). Although fetal malposition was not specifically evaluated, many of the reported obstetric outcomes are associated with fetal malposition. EBP

Charity Lehn, MD
Raeanna Simcoe, MD
Sara Foster Fabiano, MD
Kurt Cook, MD
Denver Health, University of Colorado Family Medicine Residency, Denver, CO

The authors declare no conflicts of interest.

References

1. Sng BL, Leong WL, Zeng Y, et al. Early versus late initiation of epidural analgesia for labour. *Cochrane Database Syst Rev.* 2014; (10):CD007238. [STEP 1]
2. Anim-Somuah M, Smyth RM, Cyna AM, Cuthbert A. Epidural versus non-epidural or no analgesia for pain management in labour. *Cochrane Database Syst Rev.* 2018; 5(5): CD000331. [STEP 1]
3. Wang T-T, Sun S, Huang S-Q. Effects of epidural labor analgesia with low concentrations of local anesthetics on obstetric outcomes. *Anesth Analg*. 2017; 124(5):1571–1580. [STEP 1]

Is EMDR effective in treatment of PTSD?

EVIDENCE-BASED ANSWER

In adults with posttraumatic stress disorder (PTSD), eye movement desensitization and reprocessing (EMDR) appears as effective as the standard of care (trauma-based cognitive behavioral therapy [CBT]) in improving PTSD. EMDR may be minimally to moderately superior to CBT in decreasing intrusion, arousal, anxiety, and posttraumatic symptoms of PTSD (SOR: **B**, meta-analyses of small randomized controlled trials [RCTs]). EMDR is more effective than fluoxetine in achieving long-term PTSD symptom reduction at six months for adult-onset compared with child-onset trauma survivors (SOR: **C**, small RCT).

DOI 10.1097/EBP.0000000000001258

Eye movement desensitization and reprocessing (EMDR) is a form of psychotherapy involving an eight-phase standard protocol using bilateral stimulation (most commonly horizontal saccadic eye movements) with the goal to reprocess traumatic memories and the stress symptoms associated with them.

A 2018 meta-analysis of 14 RCTs (N=675) compared the efficacy of cognitive behavioral therapy (CBT) versus EMDR in relieving posttraumatic symptoms and alleviating anxiety and depression in patients with posttraumatic stress disorder (PTSD).[1] The trials included children and adults (mean ages 10–63 years old) diagnosed with PTSD per Diagnostic and Statistical Manual of Mental Disorders (DSM)-III, -IV, or -V criteria. Trials involving treatment for non-PTSD conditions (ie, depression, bipolar disorder, behavioral problems, substance abuse) were excluded. PTSD severity was evaluated pretreatment and posttreatment via various standardized clinician and self-rated assessment tools with follow-up ranging from 3 to 15 months. Compared with CBT, EMDR reduced PTSD symptoms (11 trials, n=547, standard mean difference [SMD] –0.43; 95% CI, –0.73 to –0.12) and anxiety symptoms (five trials, n=239, SMD –0.71; 95% CI, –1.21 to –0.21) immediately posttreatment. At 3 months posttreatment, however, EMDR was not superior to CBT in reducing PTSD symptoms (four trials, n=186, SMD –0.21; 95% CI, –0.50 to 0.08) or depression symptoms (eight trials, n=295, SMD –0.21; 95% CI, –0.44 to 0.02).

Limitations include only short-term outcomes measured, and significant biases in more than one domain of included trials.

A 2015 meta-analysis of 11 RCTs (N=424) evaluated whether EMDR or CBT was superior for the treatment of adult PTSD.[2] The trials included adults 18 years old and older with PTSD defined per DSM-III or -IV criteria. PTSD severity was measured using various standardized clinician and self-rated assessment tools pretreatment and post-treatment with EMDR and CBT (follow-up periods undefined). Although this meta-analysis included some of the 2018 meta-analysis trials, specific PTSD symptoms were investigated. Overall, EMDR minimally improved symptom scores compared with CBT (SMD –0.43; 95% CI, –0.86 to –0.01). Specifically, a meta-analysis for subscaled scores of PTSD symptoms indicated that EMDR was superior for decreasing intrusion (six trials, n=170; SMD –0.37; 95% CI, –0.68 to –0.06) and arousal severity (five trials, n=146; SMD –0.34; 95% CI, –0.68 to –0.01) compared with CBT. Avoidance severity was not significantly different between groups (n=170; SMD –0.42; 95% CI, –0.92 to 0.08). Limitations include difficulty with double blinding due to nature of psychotherapy, loss to follow-up >20% in five trials, potential biases, and lack of fidelity checks for treatment sessions.

A 2007 RCT (N=76) compared the efficacy of fluoxetine to EMDR in the treatment of PTSD.[3] The trial included adults 18 to 65 years old with PTSD per DSM-IV criteria with trauma exposure one year or more before intake. Patients were randomly assigned to receive eight weekly treatment sessions with EMDR (n=24), fluoxetine (n=26), or pill placebo (n=26). The primary outcome (PTSD symptom severity) was measured using the Clinician-Administered PTSD Scale (CAPS, total severity score >50, asymptomatic end-state function <20). CAPS scores were assessed at pretreatment, posttreatment, and six months posttreatment. The Structured Clinical Interview for DMS-IV was used for determination of PTSD and comorbid diagnoses. Posttreatment, 88% of EMDR, 81% of fluoxetine, and 65% of placebo completers no longer met PTSD diagnosis criteria. Of those, 29%, 15%, and 12% of EMDR, fluoxetine, and placebo completers became asymptomatic, respectively. Between-group comparisons immediately after treatment were not statistically significant. After an intent-to-treat analysis, EMDR produced moderate improvement in symptom severity posttreatment and at six-month follow-up for adults (effect size, 0.56, 0.65 respectively) compared with fluoxetine. EMDR achieved small improvements in symptoms only at six months in children compared with fluoxetine (effect size, 0.47). Comparison statistics were not available for this reported finding. EBP

Matthew Snyder, DO
Diana Trang, DO
Nellis Family Medicine Residency
Las Vegas, NV

The authors declare no conflicts of interest.
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 U.S. Air Force Medical Department, the Air Force at large or the Department of Defense.

References

1. Chen L, Zhang G, Hu M, Liang X, Eye movement desensitization and reprocessing versus cognitive-behavioral therapy for adult posttraumatic stress disorder. *J Nervous Ment Dis*, 2015; 203(6):443–451. [STEP 1]
2. Khan AM, Dar S, Ahmed R, Bachu R, Adnan M, Kotapati VP, Cognitive behavioral therapy versus eye movement desensitization and reprocessing in patients with post-traumatic stress disorder: systematic review and meta-analysis of randomized clinical trials. *Cureus*. 2018; 10(9):e3250. [STEP 1]
3. Van der Kolk B, Spinazzola J, Blaustein M, et al, A randomized clinical trial of eye movement desensitization and reprocessing (EMDR), fluoxetine, and pill placebo in the treatment of posttraumatic stress disorder. *J Clin Psychiatry*, 2007; 68(1):37–46. [STEP 2]

Do repetitive concussions lead to chronic traumatic encephalopathy?

EVIDENCE-BASED ANSWER

Repetitive concussions may lead to symptoms that are related to the pathologic findings of chronic traumatic encephalopathy (CTE), such as depression, emotional distress, poor sleep, and impulsivity (SOR: **B**, observational studies). Among former football players, the number of years played correlates with the diagnosis CTE on brain biopsy (SOR: **C**, disease-oriented outcome).

DOI 10.1097/EBP.0000000000001255

CTE is a neurodegenerative disease that remains an exclusively postmortem diagnosis characterized by the accumulation of hyperphosphorylated tau proteins in

neurons and astrocytes. Clinical features consistent with CTE include mood changes, violent behavior, impaired cognition, and abnormal motor function.

A 2012 cohort study of 1,044 retired professional football players examined the relationship between number of concussions sustained and risk of depression.[1] Participants were surveyed in 2001 and 2010, and 106 participants (10%) reported a new depression diagnosis during the study period. Compared with those with no reported concussions, participants with one or more concussions were more likely to be diagnosed with depression during the study period (adjusted risk ratio [aRR], 2.3; 95% CI, 1.1–4.7 for 1–2 concussions; aRR, 3.3; 95% CI, 1.7–6.7 for 3–4 concussions; aRR, 4.1; 95% CI, 2.0–8.4 for 5–9 concussions; aRR, 5.8; 95% CI, 2.8–12.2 for 10+ concussions).

A 2017 cross-sectional study examined mental health in former professional athletes with a history of concussion.[2] Athletes retired from soccer, ice hockey, and rugby were recruited from a range of countries, and 576 participants completed the questionnaire. Athletes with six or more concussions reported more emotional distress symptoms (odds ratio [OR], 3.2; 95% CI, 1.5–6.9), sleep disturbance (OR, 5.2; 95% CI, 2.4–11), anxiety and depression (OR, 2.4; 95% CI, 1.2–4.9), and adverse alcohol use (OR, 2.9; 95% CI, 1.3–6.2) compared with those with no concussions. No significant difference was found for athletes with five or fewer concussions.

A cross-sectional study surveying 792 former collegiate athletes evaluated concussions and risk for depression, impulsivity, and aggression.[3] Athletes who reported three or more concussions were more likely than athletes reporting zero concussions to have moderate to severe depression based on Patient Health Questionnaire-9 scores (adjusted prevalence ratio, 2.4; 95% CI 1.0–5.7). Athletes reporting two concussions had greater impulsivity as measured by the 120-point Barratt Impulsiveness scale compared with athletes reporting zero concussions (adjusted mean difference [MD], 2.7; 95% CI, 1.2–4.1), although the correlation was less robust for athletes reporting three or more concussions (MD, 1.9; 95% CI, 0.6–3.2). Athletes reporting three or more concussions had a higher mean score for aggression as measured by the 145-point Buss-Perry Aggression Questionnaire (MD, 3.0; 95% CI, 1.4–4.7) compared with those reporting zero concussions.

A 2019 retrospective cohort study of brain tissue of former football players evaluated the correlation between years of football played and CTE.[4] The brains of 266 former football players aged 20 years and older were donated to a research bank, and neuropathologic grading for CTE was performed using a severity scale. Of the specimens examined, 223 (84%) had CTE. Compared with specimens without CTE, CTE diagnosis was positively associated with the number of years of football played (OR, 1.3 per year played; 95% CI, 1.2–1.4), and severe CTE was also associated with increasing years of football played (OR, 1.1 per year played; 95% CI, 1.1–1.2). Participants with CTE were 10 times as likely to have played >14.5 years of football (positive likelihood ratio ,10.2; 95% CI, 9.8–10.7) compared with those without CTE. EBP

Madhav Narayan, MD
Daffer Ghanim, MD
Emily Spencer, MD
Claire Bovet, MD
University of Colorado Family Medicine Residency
Denver, CO

The authors declare no conflicts of interest.

References

1. Kerr Z, Marshall S, Harding H, Guskiewicz K. Nine-year risk of depression diagnosis increases with increasing self-reported concussions in retired professional football players. *Am J Sports Med.* 2012; 40(10):2206–2212. [STEP 2]
2. Gouttebarge V, Aoki H, Lambert M, Stewart W, Kerkhoffs G. A history of concussions is associated with symptoms of common mental disorders in former male professional athletes across a range of sports. *Phys Sportsmed.* 2017; 45(4): 443–449. [STEP 3]
3. Kerr Z, Evenson K, Rosamond W, Mihalik J, Guskiewicz K. Association between concussion and mental health in former collegiate athletes. *Inj Epidemiol.* 2014; 1(1):28. [STEP 3]
4. Mez J, Daneshvar D, Abdolmohammadi B, et al. Duration of American football play and chronic traumatic encephalopathy. *Ann Neurol.* 2019; 87(1):116–131. [STEP 3]

Is anticoagulation appropriate for the treatment of superficial venous thrombosis?

EVIDENCE-BASED ANSWER

Fondaparinux has the lowest rate of pulmonary embolism/deep vein thrombosis and superficial vein thrombosis (SVT) recurrence. Rivaroxaban is noninferior to fondaparinux in the treatment of high-risk SVT (SOR: **B**, meta-analysis of randomized controlled trials and cohort trials). Anticoagulation using low-molecular-weight heparin, fondaparinux, or rivaroxaban is appropriate and warranted for use in certain clinical scenarios of SVT (SOR: **C**, clinical guideline).

DOI 10.1097/EBP.0000000000001262

A 2019 meta-analysis of 17 articles (11 randomized controlled trials and six cohort studies with N=6,862) evaluated the optimal first-line treatments for patients with isolated lower extremity superficial vein thrombosis (SVT).[1] Patients were mean age of 59 years old, 59% female, 66% with varicose veins, and a mean symptom duration of 5 days before treatment. Pharmacological treatments included fondaparinux (2.5 mg once daily, low dose, high dose, or not specified) for median duration 44 days, LMWH (low prophylactic, modified prophylactic, intermediate, or high dose) for median duration 30 days, rivaroxaban at prophylactic dosing for 45 days, unfractionated heparin ([UFH] low dose or unspecified dose) for 28 days or not specified, and warfarin (unspecified dose or duration) compared with placebo. There was one noninferiority trial of rivaroxaban versus fondaparinux at low/prophylactic doses. Outcomes were reported in events per patient-years of follow-up, with follow-up ranging from 56 to 183 days. The primary endpoint was occurrence of pulmonary embolism (PE) or deep venous thrombosis (DVT), and secondary endpoints included SVT recurrence, progression, symptomatic improvement, bleeding, and death.

Fondaparinux showed the lowest rate of DVT and PE (three studies, n=4,209; event rate 1.4 events per 100 patient years; 95% CI, 0.5–2.8). The incidence of DVT and PE (in events per 100 person-years) was 11 for rivaroxaban at prophylactic doses (one study, n=472; 95% CI, 4.3–20), 12 for warfarin (two studies, n=576; 95% CI, 3.3–60), 12 for LMWH at intermediate/full doses (seven studies, n=1,596; 95% CI, 6.8–18), 12 for LMWH at low/prophylactic doses (seven studies, n=2,014; 95% CI, 6.2–20), and 11 (two studies, n=3,429; 95% CI, 0.5–2.8) for observation/no treatment.

Fondaparinux had the lowest rate of SVT recurrence or extension at 7.7 events per 100 person-years (three studies, n=4,209; 95% CI, 1.9–17) and no therapy at 63 (two studies, n=3,429; 95% CI, 2.2–197). Bleeding and death were similar across all treatments, with UFH of any dose having the highest rates for both at 1.6 events per 100 person-years (three studies, n=636; 95% CI, 0.25–9). Rivaroxaban at low doses for 45 days was noninferior to low-dose fondaparinux for high-risk patients (male sex, over age 65 years old, previous VTE, autoimmune disease, or thrombosis of non-varicose vein) and was more cost-effective. The evidence for fondaparinux was heavily dependent on a single large publication.

In 2020, Canada updated clinical guidelines for the diagnosis and management of SVT in adult, pregnant, and pediatric populations.[2] According to these guidelines, SVTs within 3 cm of the saphenofemoral junction (SFJ) or saphenopopliteal junction (SPJ) should be treated with 3 months of therapeutic anticoagulation. Acceptable anticoagulation medications included 45 days of LMWH (eg, dalteparin 5,000-1,000 units subcutaneous (SQ) daily, enoxaparin 40–60 mg SQ daily, nadroparin 2,850–5,700 units SQ daily, tinzaparin 4,500–10,000 units SQ daily, fondaparinux 2.5 mg SQ daily, or rivaroxaban 10 mg PO daily). Further, if the SVT is equal to or greater than 5 cm in length and more than 3 cm away from the SFJ, then prophylactic treatment should be provided with fondaparinux, rivaroxaban, or LMWH for 45 days. If the SVT is more than 3 cm away from the SFJ or SPJ but less than 5 cm in length, no anticoagulation was recommended but prophylactic dosing could be considered if greater than 3 cm from deep venous system with severe symptoms or risk factors are present. Risk factors were listed as history of DVT and/or SVT, cancer, pregnancy, hormone management, or recent trauma or therapy. No strength of recommendation was provided. EBP

Jonathon Gross, MD, MBA
Charles Smith, MD
Sarah Daly, DO, FAAFP
Utah Valley Family Medicine, Provo, UT

The authors declare no conflicts of interest.

References

1. Duffet L, Kearon C, Rodger M, Carrier M. Treatment of superficial vein thrombosis: a systematic review and meta-analysis. *Thromb Haemost.* 2019; 119(3): 479–489. [STEP 2]
2. *Superficial Thrombophlebitis, Superficial Vein Thrombosis. Thrombosis Canada*. https://thrombosiscanada.ca/wp-content/uploads/2020/08/Superficial-Vein-Thrombosis_29Aug2020.pdf. Accessed Nov 11, 2020. [STEP 5]

In acute back pain, do topical NSAIDs relieve pain as well as oral NSAIDs?

EVIDENCE-BASED ANSWER

In patients with acute back pain, topical felbinac (not available in the United States) decreases pain to a similar degree as oral ibuprofen, but there is a dearth of evidence comparing other topical NSAIDs with oral NSAIDs (SOR: **B**, single RCT from systematic review). Recommendations on the use of topical NSAIDs for acute back pain are inconsistent, whereas recommendations on the use of oral NSAIDs are more consistently in favor of therapy (SOR: **C**, expert opinion from systematic review of clinical practice guidelines).

DOI 10.1097/EBP.0000000000001271

A 2015 systematic review and meta-analysis of 61 double-blinded RCTs (N=8,386) evaluated the efficacy and safety of topical NSAIDs in adults with acute musculoskeletal pain.[1] Of the 61 studies only three compared topical NSAIDs with oral NSAIDs; of which, only one (n=287) evaluated patients with acute back pain. This study is summarized in greater detail below. None of the studies comparing topical NSAIDs with placebo included patients with acute back pain. The review mentions a large number of unpublished studies, which were not included: some potentially including comparisons to oral NSAIDs.

A 1993 double-blinded RCT that included 287 patients compared topical felbinac with oral ibuprofen in adults with acute back pain.[2] Patients were 18 to 63 years old with a mean age of 37 years with pain in the lumbosacral region and duration of less than one month. Exclusion criteria were sciatic pain, two or more episodes of back pain in the past six months, and previous vertebral injury or disease. Patients received felbinac 3% foam (not available in the United States) 2 g applied to the back three times a day with an oral placebo tablet for seven days (n=140) or oral ibuprofen 400 mg three times a day with placebo foam for seven days (n=147). Response was defined as no pain or mild pain (lowest two categories on 5-point categorical scale) with movement and at rest at 1-week and 2-week follow-up visits. Patients reporting no pain at week 1 were not assessed at week 2. At week 1, the felbinac group had a significant increase in patients in the none and mild pain categories at rest (from 42% to 78%; $P<.001$) and with movement (from 3% to 64%; $P<.001$). The same was true in the ibuprofen group (rest from 30% to 81%; $P<.001$; movement from 8% to 72%; $P<.001$). No difference was observed between groups. At 2 weeks, the number of responders were similar ($P>.05$) in each group at rest (88% for felbinac and 89% for ibuprofen) and with movement (78% for felbinac and 82% for ibuprofen).

A 2020 systematic review of 17 clinical practice guidelines aimed to summarize treatment recommendations for neck and low back pain.[3] Only guidelines from European countries published since 2015 were included. Eleven guidelines involved low back pain, five neck pain, and one involved both. Seven of 17 were considered high-quality guidelines based on the Appraisal of Guidelines Research and Evaluation II checklist consisting of 23 items that address six different domains (scope and purpose, involvement of stakeholders, developmental rigor, presentation clarity, applicability, and editorial independence). Three guidelines made treatment recommendations regarding topical NSAIDs for acute low back pain. One high-quality guideline recommended "definitely do not" use topical NSAIDs for acute low back pain and two low-quality guidelines were "for" the use of topical NSAIDs in this condition. There were four high-quality recommendations in favor of oral NSAID use. Furthermore, there were five low-quality guidelines in favor of oral NSAIDs. The guidelines largely came from Northern and Western Europe. EBP

Anthony Hacker, MD
Thomas Satre, MD
University of Minnesota/CentraCare St. Cloud FMR
St. Cloud, MN

The authors declare no conflicts of interest.

References

1. Derry S, Moore RA, Gaskell H, McIntyre M, Wien PJ. Topical NSAIDs for acute musculoskeletal pain in adults. *Cochrane Database Syst Rev.* 2015; (6):CD007402. [STEP 1]
2. Hosie GA. The topical NSAID, felbinac, versus oral ibuprofen: a comparison of efficacy in the treatment of acute lower back injury. *Br J Clin Res.* 1993; 4:5–17. [STEP 2]
3. Corp N, Mansell G, Stynes S, et al. Evidence-based treatment recommendations for neck and low back pain across Europe: a systematic review of guidelines. *Eur J Pain*. 2021; 25:275–295. [STEP 5]

In diagnosing BV, are patient-collected samples as accurate as provider-collected samples?

EVIDENCE-BASED ANSWER

Yes. In symptomatic patients presenting with vaginal discharge, excellent agreement is found between the results of self-collected swabs and physician-collected swabs in the results of tests for bacterial vaginosis (BV; SOR: **B**, multiple small case-control studies). Among asymptomatic pregnant patients, also substantial agreement is found between self-collected and physician-collected swabs in the accuracy of diagnosing BV (SOR: **B**, small case-control study).

DOI 10.1097/EBP.0000000000001283

A 2015 case-control study (n=104) compared the validity of self-taken to clinician-taken vaginal swabs in the detection of bacterial vaginosis (BV).[1] Female patients 17 to 49 years old in the United Kingdom who presented with symptoms of vaginal discharge, genital irritation, or offensive genital smell were included in the study. Patients already diagnosed with vulvovaginal candidiasis or BV and those with immunodeficiency were excluded. Patients were seen by a physician or nurse and instructed on how to perform a self-taken low vaginal swab. The female patients then underwent a speculum examination and a clinician-taken high vaginal swab. Both self-taken and clinician-taken swabs were sent to a microbiology laboratory for microscopy and culture and assessed using Hay-Ison scoring for BV. Using the clinician-taken swabs as the reference standard, the sensitivity of self-taken swab for BV was 89% (95% CI, 69–97%) and specificity was 96% (95% CI, 87–99%) with a positive likelihood ratio (LR+) of 22.2 and negative likelihood ratio (LR–) of 0.11. Limitations of the study included small sample size and lack of blinding by laboratory staff to sample conditions.

A 2008 case-control study (n=50) examined the reliability of self-collected versus provider-collected vaginal swabs in the diagnosis of BV.[2] Sexually active women 20 to 45 years old in a gynecological outpatient clinic in India with a complaint of vaginal discharge first performed a self-collected swab and then the physician performed a speculum examination and collected a second sample. Samples were evaluated in the laboratory according to Nugent criteria. Compared with physician-collected samples, the self-obtained smear had a sensitivity of 70% and a specificity of 97% for BV (LR+ 23.3; LR– 0.31). No CIs were reported. Limitations of this study included small sample size and a patient population with a perceived lower health literacy.

A 2019 case-control study (n=550) evaluated the reliability of self-collected versus physician-collected vaginal swabs for the diagnosis of BV.[3] Sexually active women with vaginal discharge 18 to 45 years old who attended sexually transmitted infection clinics in India were included in the study. Patients with use of antibiotics or vaginal medication in the preceding two weeks, HIV, or pregnancy were excluded from the study. Patients received instruction and performed self-collection then underwent speculum examination with physician-collected swabs. Samples were evaluated in the laboratory according to Amsel's and Nugent criteria. Cohen's kappa indicated excellent agreement in the diagnosis of BV (k=0.95; 95% CI, 0.99–0.99). Study limitations include lack of blinding to specimen condition.

A 2005 case-control study (n=129) evaluated the reliability of self-collected versus physician-collected vaginal swabs for the diagnosis of BV among pregnant women.[4] Included patients were 24 to 29 weeks pregnant, at least 16 years old, and received prenatal care through a local hospital system in the United States. Patients were instructed and obtained self-collected samples and then physician-collected swabs were obtained during speculum examination. Samples were evaluated in the laboratory using Nugent criteria. Compared with physician-collected swabs, the self-collected swabs had sensitivity of 77%, specificity of

97% (LR+ 25.7; LR– 0.24). CIs were not reported. Study limitations included small sample size and limited extrapolation to nonpregnant patients. EBP

Adam H. Rosenfeld, BS
Nikole J. Cronk, PhD
Department of Family Medicine, University of Missouri
Columbia, MO

The authors declare no conflict of interest.

References

1. Barnes P, Vieira R, Harwood J, Chauhan M. Self-taken vaginal swabs versus clinician-taken for detection of candida and bacterial vaginosis: a case-control study in primary care. *Br J Gen Pract*. 2017; 67(665):e824–e829. [STEP 4]
2. Kashyap B, Singh R, Bhalla P, Arora R, Aggarwal A. Reliability of self-collected versus provider-collected vaginal swabs for the diagnosis of bacterial vaginosis. *Int J STD AIDS*. 2008; 19(8):510–513. [STEP 4]
3. Khan Z, Bhargava A, Mittal P, et al. Evaluation of reliability of self-collected vaginal swabs over physician-collected samples for diagnosis of bacterial vaginosis, candidiasis, and trichomoniasis, in a resource-limited setting: a cross-sectional study in India. *BMJ Open*. 2019; 9:e025013. [STEP 4]
4. Strauss RA, Eucker B, Savitz DA, Thorp JM. Diagnosis of bacterial vaginosis from self-obtained vaginal swabs. *Infect Dis Obstet Gynecol*. 2005; 13(1):31–35. [STEP 4]

Are NSAIDs effective for acute low back pain?

EVIDENCE-BASED ANSWER

NSAIDs are slightly more effective than placebo in reducing pain (by 7.3%) and disability (by 8.3%) over three weeks, and 1.4 times more likely to result in global improvement; however, the impact on adverse events and return to work is inconclusive (SOR: **A**, systematic review of randomized controlled trials). A clinical practice guideline recommends NSAIDs or skeletal muscle relaxants as the first-line options for acute or subacute low back pain (SOR: **C**, expert opinion based on assessment of a systematic review).

DOI 10.1097/EBP.0000000000001285

A 2020 systematic review included nine randomized controlled trials (RCTs; N=2,693) evaluating the efficacy of NSAIDs compared with placebo for the treatment of acute low back pain.[1] Patients were 17 years old and older presenting to primary care and specialty settings with low back pain of less than 12 weeks' duration. The most commonly studied NSAID was diclofenac (five RCTs), with doses ranging from 25 to 150 mg PO daily; other NSAIDs were piroxicam 20 to 40 mg PO daily (two RCTs), ibuprofen 400 to 1,200 mg PO daily (one RCT), and tenoxicam (not available in the United States) 20 mg PO daily (one RCT). Treatment lasted from one day to four weeks. The primary outcomes were pain intensity, disability, global improvement, adverse events, and return to work. Follow-up ranged from one day to two months. NSAIDs provided short-term pain reduction from baseline compared with placebo during a three-week follow-up (see **TABLE**). NSAIDs were also associated with

TABLE. Effect of NSAIDs versus placebo on acute low back pain

Outcome[a]	No. of RCTs	No. of patients	Result (95% CI)	Evidence quality
Pain intensity (MD)[b]	4	815	–7.3 (–11.0 to –3.6)	Moderate
Disability (MD)[c]	2	471	–2.0 (–2.9 to –1.1)	High
Global improvement (RR)[d]	5	1,202	1.4 (1.1 to 1.7)	Low
Adverse events (RR)[e]	6	1,394	0.86 (0.63 to 1.2)	Very low
Return to work (RR)[e]	1	266	1.5 (0.98 to 2.2)	Very low

Data from a systematic review of RCTs.[1] [a] Pain intensity, disability, and global improvement measured the change from baseline over three weeks; adverse events were measured between one day and twelve weeks; return to work was measured at seven days. [b] Pain intensity was measured using a visual analogue scale (0–100; lower=better); a 10% decrease (ie, –10 points) or more was considered clinically relevant. [c] Disability was measured using Roland Morris Disability questionnaire (0–24; lower=better); a 10% decrease (ie, –2.4 points) or more was considered clinically relevant. [d] Global improvement measured using various dichotomized Likert scales (eg, recovered vs not recovered); an RR greater than 1 favored NSAIDs over placebo. [e] For adverse events, an RR less than 1 favored NSAIDs over placebo; for return to work, an RR greater than 1 favored NSAIDs over placebo. MD=mean difference; RR=relative risk.

reduced disability and increased global improvement from baseline to three weeks. Of note, the magnitude of the effect of NSAIDs over placebo on reduction of pain and disability did not meet the 10% threshold of clinical relevance. The impact of NSAIDs on adverse events and return to work was inconclusive due to very low-quality evidence. Limitations included possible industry bias since almost half of the included trials were supported by pharmaceutical companies.

A 2017 evidence- and consensus-based clinical practice guideline on noninvasive treatments for low back pain from the American College of Physicians recommended either NSAIDs or skeletal muscle relaxants as the first-line options for acute or subacute low back pain if pharmacologic treatment is desired (strong recommendation based on moderate-quality evidence from a 2017 systematic review).[2] EBP

Eduardo Jesus Garza, MD
Northwest Washington Family Medicine Residency
Bremerton, WA

Amanda Kim, MD, MPH
Valley Family Medicine Residency Program
Renton, WA

Alexa Lindley, MD, MPH
University of Washington Family Medicine Residency
Seattle, WA

The authors declare no conflict of interest.

References

1. van der Gaag WH, Roelofs PD, Enthoven WT, et al. Non-steroidal anti-inflammatory drugs for acute low back pain. *Cochrane Database Syst Rev*. 2020; 4(4):CD013581. [STEP 1]
2. Qaseem A, Wilt TJ, McLean RM, et al. Clinical Guidelines Committee of the American College of Physicians. Non-invasive treatments for acute, subacute, and chronic low back pain: a clinical practice guideline from the American College of Physicians. *Ann Intern Med*. 2017; 166(7):514–530. [STEP 5]

What are the benefits and harms of testosterone replacement therapy in men with age-related low testosterone?

EVIDENCE-BASED ANSWER

Testosterone replacement therapy (TRT) in men with age-related low testosterone yields small improvements in sexual function, erectile dysfunction, libido, sexual satisfaction, and quality of life without clear benefits on energy, mood, physical function, or cognitive function (SOR: **B**, systematic reviews of randomized controlled trials [RCTs]). TRT increases the risk of erythrocytosis eight-fold but does not increase adverse cardiac events (SOR: **B**; systematic reviews of RCTs). There are insufficient data to determine if TRT impacts the risk of death, venous thromboembolism, or prostate cancer. Experts suggest offering TRT to men with age-related low testosterone to improve sexual function but not to improve energy, mood, or cognition (SOR: **C**, evidence-based guidelines).

DOI 10.1097/EBP.0000000000001307

A 2020 systematic review of 38 randomized controlled trials (RCTs; N=4,455) examined the benefits and harms of testosterone replacement therapy (TRT) in men with low testosterone not attributed to an underlying organic cause.[1] On average, there were 118 patients in each trial (range 10–790). The mean age ranged from 52 to 77 years old (overall average 66 years old) and the mean baseline testosterone was <300 ng/dL in 20 trials. Testosterone was administered as a patch (seven RCTs, N=457), gel (12 RCTs, N=2,558), or IM injection (19 RCTs, N=1,440). The patch and gel formulations were given daily and the frequency of injections varied from weekly to every 12 weeks depending on formulation (cypionate, enanthate, or undecanoate) and dose (125–1,000 mg). The comparators were placebo (36 RCTs, N=4,069) and no treatment (two RCTs, N=386). To assess benefits of TRT, the trials used a variety of different questionnaires to measure outcomes, and the review authors calculated the standardized mean difference (SMD) in pooled analyses. These showed that TRT produced small improvements in overall sexual function, erectile dysfunction, and quality of life when compared with placebo or no treatment (see **TABLE**). There were very small improvements in vitality and mood but no change in physical function. Analysis of effects of TRT on cognitive function were inconclusive; all but two studies were underpowered to show a difference and the two adequately powered studies showed no effect of TRT on

TABLE. Benefits and harms of testosterone replacement therapy versus placebo or no treatment in men with age-related low testosterone

	No. of RCTs	No. of patients	Outcome measure[a]	Outcome (95% CI)	Reference
Benefit					
Overall sexual function	7	1,140	SMD	**0.35 (0.23–0.46)**	1
Erectile function[b]	7	1,299	SMD	**0.17 (0.09–0.44)**	1
	4	1,344	SMD	**0.16 (0.06–0.27)**	2
Libido	3	1,383	SMD	**0.16 (0.01–0.34)**	2
Sexual satisfaction	2	676	SMD	**0.16 (0.01–0.31)**	2
Quality of life	7	1,043	SMD	**0.33 (0.16–0.5)**	1
Vitality/energy[b]	3	665	SMD	**0.17 (0.01–0.32)**	1
	2	1,503	SMD	0.08 (–0.02 to 0.18)	2
Mood[b]	5	872	SMD	**0.19 (0.05–0.32)**	1
	2	1,179	SMD	0.08 (–0.03 to 0.2)	2
Physical function	5	1,029	SMD	0.15 (–0.19 to 0.5)	1
Harm					
Adverse cardiac events	14	2,415	OR	1.2 (0.66–2.2)	1
Serious adverse events	8	2,268	OR	0.94 (0.73–1.21)	1
Erythrocytosis[c]	3	1,579	RR	**8.1 (1.9–35.4)**	2

Statistically significant results in bold font. Data from two systematic reviews of RCTs.[1,2] [a] SMD effects were interpreted as small=0.2, medium=0.5, and large≥0.8. [b] Both systematic reviews reported on erectile function, vitality or energy, and mood. [c] Erythrocytosis was defined as a hematocrit >54% or hemoglobin >17.5 g/dL. OR=odds ratio; RCT = randomized controlled trial; RR=relative risk; SMD=standardized mean difference.

cognitive function. As for harms, there were no differences in the primary outcomes of adverse cardiac events and serious adverse events (see **TABLE**), and the studies were inadequately powered to detect differences in mortality, venous thromboembolisms, and prostate cancer. There was no difference between the intervention and control groups in treatment withdrawals. Limitations included significant heterogeneity between the RCTs and low- to moderate-quality evidence for most of the outcomes.

A 2018 systematic review of four placebo-controlled RCTs (N=1,779) evaluated the benefits and harms of TRT in men with hypogonadism, defined as having one morning total testosterone level ≤300 ng/dL and one or more symptoms of hypogonadism.[2] Average ages across studies ranged from 50.6 to 72.2 years old. The TRT formulations used were 60 mg of testosterone solution 2% daily and testosterone gel 1% either 50 or 100 mg daily. Trial durations ranged from 12 to 52 weeks. In pooling results, the researchers calculated SMDs because of the different scales used to measure outcomes. Compared with placebo, TRT slightly improved libido, erectile function, and sexual satisfaction (see **TABLE**). TRT had no effect on energy level or mood and it did not improve memory. The only significant adverse effect was an increased risk of erythrocytosis, defined as a hematocrit >54% or hemoglobin >17.5 g/dL (see **TABLE**). The trials were deemed to be at low risk of bias but the review was limited by the inclusion of men with hypogonadism of various causes and insufficient power and duration to assess the effects of TRT on the incidence of prostate cancer, cardiovascular events, and bone fractures. One trial with 790 patients was also included in the previously mentioned review.[1]

A 2020 evidence-based guideline from the American College of Physicians (ACP) on testosterone treatment in men with age-related low testosterone suggested discussing TRT for men with sexual dysfunction to improve sexual function (conditional recommendation based on low-certainty evidence).[3] The ACP recommended against testosterone treatment to improve energy, vitality, physical function, or cognition

(conditional recommendation based on low-certainty evidence).

An evidence-based 2018 Endocrine Society guideline on testosterone therapy in men with hypogonadism suggested offering TRT to men more than 65 years old if they have consistent and unequivocally low morning serum testosterone concentrations and symptoms consistent with testosterone deficiency (conditional recommendation based on low-quality evidence).[4] The guideline stated that TRT be offered only after a discussion of the risks and benefits and recommended against routinely prescribing TRT to men with low testosterone (strong recommendation based on low-quality evidence). EBP

Mikaela Alger, MD
Puyallup Tribal Health Authority Family Medicine Faculty, Tacoma, WA

Natalia A. Luera, MD FAAFP
Kadlec Family Medicine Residency, Richland, WA

Rachel Weiner, MD MPH
Kaiser Permanente Washington Family Medicine Residency, Seattle, WA

The authors declare no conflicts of interest.

References

1. Diem SJ, Greer NL, MacDonald R, et al. Efficacy and safety of testosterone treatment in men: an evidence report for a clinical practice guideline by the American College of Physicians. *Ann Intern Med.* 2020; 172(2): 105–118. [STEP 2]
2. Ponce OJ, Spencer-Bonilla G, Alvarez-Villalobos N, et al. The efficacy and adverse events of testosterone replacement therapy in hypogonadal men: a systematic review and meta-analysis of randomized, placebo-controlled trials. *J Clin Endocrinol Metab.* 2018; doi: 10.1210/jc.2018-00404. Epub ahead of print. [STEP 2]
3. Qaseem A, Horwitch CA, Vijan S, Etxeandia-Ikobaltzeta I, Kansagara D; Clinical guidelines committee of the American College of Physicians. Testosterone treatment in adult men with age-related low testosterone: a clinical guideline from the American College of Physicians. *Ann Intern Med.* 2020; 172(2):126–133. [STEP 5]
4. Bhasin S, Brito JP, Cunningham GR, et al. Testosterone therapy in men with hypogonadism: an Endocrine Society clinical practice guideline. *J Clin Endocrinol Metab.* 2018; 103(5):1715–1744. [STEP 5]

Is curcumin an effective treatment of delayed onset muscle soreness?

EVIDENCE-BASED ANSWER

Oral curcumin supplementation of 500 to 2,500 mg given for 2.5 to 28 days before strenuous exercise may result in a moderate decrease in patient-reported delayed onset muscle soreness (SOR: **B**, multiple small randomized controlled trials).

DOI 10.1097/EBP.0000000000001286

A 2020 randomized controlled trial (RCT; n=30) compared efficacy of curcumin with placebo on delayed-onset muscle soreness, defined as muscle pain appearing 48 hours after exercise.[1] Participants included 12 men and 18 women, mean age 36 years old, who exercised at least four hours a week and had no musculoskeletal pathology. Smokers and subjects using anti-inflammatory, analgesic, or antioxidant drugs in the prior month were excluded. Baseline measurements of VO_2 max, creatine kinase, and muscle pain were obtained seven days before intervention, after a downhill treadmill running eccentric exercise protocol that was identical to the intervention exercise protocol. The intervention group (n=15) then received Cureit® brand oral curcumin 500 mg daily for four days leading up to the intervention exercise, whereas controls (n=15) received starch placebo 500 mg daily. Primary outcomes were VO_2 max testing, creatine kinase levels, and pain reduction. Pain outcomes were assessed 48 hours after the exercise protocol using a 0 to 4 visual analog scale, where zero represented no pain and four represented disabling pain while descending or climbing stairs. Curcumin significantly reduced pain scores from post baseline exercise to post intervention exercise compared with placebo (visual analog scale 2.9 to 1.2 in intervention, 2.7 to 2.4 in control, $P \leq .001$). No significant differences were noted between creatine kinase levels or VO_2 max levels. No adverse events were reported.

A 2015 randomized, double-blind crossover trial (N=19) compared the efficacy of curcumin with placebo on delayed-onset muscle soreness.[2] Participants were

male, mean age 34 years old, and did regular "light to moderate physical activity." Subjects who participated in leg weight training in the previous three months, had current lower limb musculoskeletal injury or neurological disease, or currently used NSAIDs were excluded. The intervention group received 2.5 g of oral curcumin, whereas the control group received 2.5 g of placebo starch, twice daily for 2.5 days before and after exercise (five days total). A single-leg press eccentric exercise protocol was used to induce delayed-onset muscle soreness. Crossover was separated by a 14-day washout. Outcomes were measured at zero, 24, and 48 hours postexercise. The primary outcome was muscle pain after a series of exercises, assessed using a zero to 10 visual analog scale, with zero representing no pain and 10 representing severe pain. At 24 and 48 hours postexercise, curcumin intervention resulted in reductions in rise from immediate postexercise pain levels compared with placebo for single-leg squat pain: –1.4 (90% CI, –2.4 to –0.4) and –1.7 (90% CI, –2.7 to –0.7), gluteal stretch: –1.0 (90% CI, –1.9 to –0.1) and –1.9 (90% CI, –2.8 to –1.0), single-leg jump at 24 hours postexercise only: –1.5 (90% CI, –2.7 to –0.3), and pain on walking downstairs at 48 hours postexercise only: –1.3 (90% CI, –2.5 to –0.1). No significant effect of curcumin was noted on pain on walking downstairs at 24 hours post exercise, quadriceps stretching pain, or single-leg jump at 48 hours postexercise. The subjects taking curcumin experienced no side effects.

A 2020 RCT (N=20) compared the effect of curcumin to placebo on oxidative stress, inflammation, muscle damage, and muscle soreness.[3] Healthy males who participated in at least 150 minutes of moderate-intensity aerobic activity or 30 minutes of vigorous-intensive aerobic activity per week were included (mean age 22 years old). Subjects with known cardiometabolic or musculoskeletal disorders, allergy to curcumin, regular use of NSAIDs, or current use of blood thinners were excluded. The intervention group received 500 mg of oral curcumin capsules (CurcuFresh® brand) three times daily for 28 days; placebo group received an odor- and color-matched placebo. All subjects underwent an eccentric exercise protocol (repeated single-leg sit-and-stand repetitions), at baseline and after 25 days of intervention. The primary outcomes were blood concentrations of tumor necrosis factor-alpha, total antioxidant capacity, malondialdehyde, and creatine kinase. Creatine kinase levels were significantly reduced in the curcumin group compared with placebo (200 U/L compared with 287 U/L, $P \leq .0001$). There were no significant differences in tumor necrosis factor-alpha, total antioxidant capacity or malondialdehyde levels between curcumin and control groups. The secondary outcome was muscle soreness, assessed using a visual analog scale, with zero representing no soreness and the highest value (scale uncertain) representing unbearable soreness. Muscle soreness measurements were made immediately pre-exercise, and at 60 minutes, 24, and 48 hours postexercise. Compared with placebo, curcumin resulted in lower overall muscle soreness (visual analog scale 3.4 in placebo, 2.9 in intervention, $P=.012$). The study did not specify the period for the "overall" muscle soreness measurements. No comment was noted on side effects from curcumin treatment. EBP

Matthew Ludemann, MD
Taylor Hart, MD
St. Anthony North Family Medicine Residency
Westminster, CO

The authors declare no conflict of interest.

References

1. Amalraj A, Divya C, Gopi S. The effects of bioavailable curcumin (Cureit) on delayed onset muscle soreness induced by eccentric continuous exercise: a randomized, placebo-controlled, double-blind clinical study. *J Med Food*. 2020; 23(5):545–553. [STEP 2]
2. Nicol LM, Rowlands DS, Fazakerly R, Kellett J. Curcumin supplementation likely attenuates delayed onset muscle soreness (DOMS). *Eur J Appl Physiol*. 2015; 115: 1769–1777. [STEP 2]
3. Basham S, Waldman H, Krings B, et al. Effect of curcumin supplementation on exercise-induced oxidative stress, inflammation, muscle damage, and muscle soreness. *J Diet Suppl*. 2020; 17(4):401–414. [STEP 2]

Do physically active pregnant women experience less back pain compared with more sedentary pregnant women?

EVIDENCE-BASED ANSWER

Yes, participation in a defined exercise program during pregnancy is modestly protective against low back pain and reduces new sick leave associated with low back or lumbopelvic pain by 23% (SOR: **A**, meta-analysis of randomized controlled trials). Regular exercise before conception, especially high-impact activities, provides a small decrease in pelvic girdle pain during pregnancy (SOR: **B**, single cohort study).

DOI 10.1097/EBP.0000000000001238

A 2017 meta-analysis of 11 randomized controlled trials (RCTs) (N=2,347) evaluated the effects of exercise compared with routine daily activities on the primary prevention of low back pain, pelvic girdle pain, and related sick leave in pregnant women.[1] All studies were conducted outside the United States. Interventions varied and included water aerobics, sitting pelvic tilt exercise, strengthening exercises, low-impact gymnastics, energy expenditure exercise, and combinations of exercises. Exercise frequencies varied from one hour per week to at least 30 minutes daily, and duration from eight to 24 weeks. All studies began in the second trimester except one in the third trimester (n=67). The incidence of back and/or pelvic girdle pain was measured by self-report (10 RCTs, n=2,305) and the Roland-Morris Disability Questionnaire (one RCT, n=42) in either the second to the third trimesters (three RCTs, n=498) or in the third trimester alone (eight RCTs, n=1,849). Sick leave related to low back pain or pelvic girdle pain was measured in four RCTs (n=1,412). Exercise during pregnancy reduced low back pain by 9% compared with usual daily activity (seven RCTs, n=1,175; pooled relative risk [RR] 0.91; 95% CI, 0.83–0.99). However, no benefit to exercise during pregnancy over daily activity was noted for lumbopelvic pain (eight RCTs, n=1,737; pooled RR 0.96; 95% CI, 0.90–1.02). Exercise during pregnancy decreased sick leave related to low back pain or lumbopelvic pain by 23% (four RCTs, n=1,412; pooled RR 0.77; 95% CI, 0.63–0.94). Studies were limited by the heterogeneity of exercise types, exercise durations, and timeline of outcome assessments.

A 2015 prospective population-based cohort study (N=39,184) assessed the association between exercise levels prepregnancy and pelvic girdle pain in pregnancy in nulliparous women with singleton pregnancies.[2] Enrolled women completed questionnaires that included items about current and prepregnancy exercise frequency and type, maternal health, current living status, education, smoking status, other lifestyle behaviors, and medical history at pregnancy weeks 17 and 30. The primary outcome was self-reported pelvic girdle pain, defined as combined pain in the anterior pelvis and in the posterior pelvis bilaterally, in pregnancy week 30. Incidence of the primary outcome was 10%. In those who did not develop pelvic girdle pain, it was more common to report an exercise frequency of three to five times per week compared with women who developed pelvic girdle pain (no measure of effect given, $P<.001$). Participation in high-impact exercises such as running, jogging, orienteering, ballgames, and high-impact aerobics was more commonly reported in women without pelvic girdle pain, than in those who had pelvic girdle pain (no measure of effect given, $P<.001$). Women who exercised three to five times weekly before pregnancy had a 14% reduced risk of pelvic girdle pain in pregnancy compared with no exercise (adjusted RR [aRR] 0.86; 95% CI, 0.77–0.96). Participation in high-impact exercises before pregnancy was associated with a similar risk reduction of pelvic girdle pain (aRR 0.86; 95% CI, 0.76–0.96). Limitations included that both exposure and outcome were self-reported. Generalizability was limited by a low participation rate and healthy participant bias. EBP

Nicole Formhals
Charles Baumann
Natalie Long, MD
Department of Family Medicine, University of Missouri, Columbia, MO

The authors declare no conflicts of interest.

References

1. Shiri R, Coggon D, Falah-Hassani K. Exercise for the prevention of low back pain: systematic review and meta-analysis of controlled trials. *Am J Epidemiol*. 2018; 187(5): 1093–1101. [STEP 1]
2. Owe KM, Bjelland EK, Stuge B, Orsini N, Eberhard-Gran M, Vangen S. Exercise level before pregnancy and engaging in high-impact sports reduce the risk of pelvic girdle pain: a population-based cohort study of 39 184 women. *Br J Sports Med*. 2016; 50(13):817–822. [STEP 3]

In patients with rotator cuff disease, does shock wave therapy improve pain and function?

EVIDENCE-BASED ANSWER

No. Shock wave therapy is unlikely to offer any clinically relevant improvement in pain or function in patients with rotator cuff disease (SOR: **B**; meta-analysis of low-quality randomized controlled trials [RCTs] and single small RCT).

DOI 10.1097/EBP.0000000000001359

A 2020 systemic review and meta-analysis of 32 randomized controlled trials (RCTs) (N=2,281) examined the efficacy of shock wave therapy in rotator cuff disease.[1] Any RCT comparing shockwave therapy with placebo or other treatment in patients with rotator cuff disease was considered for inclusion. Rotator cuff pathology varied by study, including patients with calcific or noncalcific tenonitis, tendinopathy, subacromial bursitis, or tears. Patients with significant injury (eg, complete rotator cuff tear), shoulder arthritis, or systemic inflammatory conditions were generally excluded. Mean participant ages ranged from 48 to 56 years old, 61% were female, and average duration of rotator cuff condition ranged from 7 to 60 months. The intervention group received extracorporeal or radial shock wave therapy that varied from 0.1 to 0.4 mJ/mm^2 energy shock waves over 1 to 6 sessions, seven to 16 days apart. Twelve studies compared shock wave therapy versus placebo sham therapy. The primary outcomes were mean pain and function scores measured using a visual analogue scale (VAS) and the Constant Score, respectively. Scores ranged from 0 to 10 on the VAS, with higher scores indicating worsening pain; a change of 1.5 points was considered clinically relevant. Scores ranged from 0 to 100 on the Constant Score, with higher scores indicating improving function; a difference of 10 points was considered clinically relevant. If pooled studies used different scales, results were reported as standardized mean differences (SMDs). A statistically significant but clinically irrelevant improvement was noted in pain with shock wave therapy compared with placebo (9 trials, N=608; SMD –0.49; 95% CI, –0.88 to –0.11) at 6-week to 3-month interval, which is about a 0.78-point difference on the VAS scale. At the 6- to 12-month interval, no significant difference was noted on the VAS pain scale (3 trials, N=155; mean difference [MD] –2.4; 95% CI, –5.8 to 0.95) in pain improvement with shock wave therapy compared with placebo. Similarly, a statistically significant but clinically irrelevant improvement in function was noted with shock wave therapy compared with placebo (9 trials, N=612; SMD 0.62; 95% CI, 0.13–1.1) at the 6-week to 3-month interval, which is a 7.93-point difference on the Constant Score. At the 6- to 12-month interval, no significant difference in function was noted as measured on the Constant Score (3 trials, N=155; MD 15; 95% CI, –2.6 to 33) with shock wave therapy compared with placebo. Subgroup analysis also found no difference in function between patients with calcific deposits (5 trials, N=260; SMD 0.84; 95% CI, –0.20 to 1.9) or without calcific deposits (5 trials, N=253; SMD 0.29; 95% CI, –0.04 to 0.61) at 6-week to 3-month interval. Side effects were reported as minimal and included treatment-related pain and self-limiting bruising and bleeding. This systemic review was limited by lack of standardization among interventions and presence of selection bias (75% of studies), performance bias (62% of studies), detection bias (62% of studies), and reporting bias (45% of studies).

A 2020 double-blinded RCT (N=37) published after the review above examined the therapeutic effects of shock wave therapy on rotator cuff lesions with shoulder stiffness compared with placebo sham therapy in patients 35 to 80 years old.[2] Patients had shoulder pain (0-10 VAS scores >3), stiffness (greater than 50% reduction in range of motion), and positive impingement sign that was unresponsive to physical therapy or activity modification for at least three months with supraspinatus tendinopathy or partial tear on MRI. Patients with rheumatic disease, glenohumeral osteoarthritis, full-thickness cuff tears, fractures, or a subacromial injection within three weeks were

excluded. The intervention group (n=16) received 0.32 mJ/mm^2 extracorporeal shock wave therapy (ESWT) during a single one-hour session while the placebo control group (n=15) received placebo (ie, sham therapy) for one hour. Researchers assessed outcomes using a 0 to 10 VAS pain scale and reported the median score. At the three-month interval postintervention, VAS median scores were not significantly different between groups: 1.0 with ESWT compared with 3.0 with sham therapy (P=.089). However, at the 12-month interval, median VAS scores were statistically different but likely not clinically different between groups: 0.0 with ESWT compared with 0.5 with sham therapy (P=.025). Side effects were minimal and included transient erythematous swelling at treatment site (10.5%) and self-limiting petechial bleeding (5.3%). This RCT does not change the conclusion of the above systemic review given the small sample size and small magnitude of effect. EBP

Prakhya Bhatnagar, MD
Amy Olmschenk, DO
University of Minnesota, CentraCare St. Cloud Family Medicine Residency, St. Cloud, MN

The authors declare no conflicts of interest.

References

1. Surace SJ, Deitch J, Johnston RV, Buchbinder R. Shock wave therapy for rotator cuff disease with or without calcification. *Cochrane Database Syst Rev.* 2020;(3):CD008962. [STEP 1]
2. Ko J-Y, Siu K-K, Wang F-S, et al. The therapeutic effects of extracorporeal shock wave therapy (ESWT) on the rotator cuff lesions with shoulder stiffness: a prospective randomized study. *Biomed Res Int.* 2020;2020:1-9. [STEP 3]

After incision and drainage for skin abscesses, does packing of the wound with gauze lead to better healing outcomes?

EVIDENCE-BASED ANSWER

No. Packing of the wound after incision and drainage of a skin abscess does not decrease the recurrence of abscess or the development of fistula in-ano after perianal procedures. Packing may lead to an increased need for a second intervention at 48 hours (SOR: **B**, meta-analysis of small randomized controlled trials). No specific guideline recommendations are found for or against packing of skin abscesses.

DOI 10.1097/EBP.0000000000001327

A 2020 meta-analysis of eight randomized controlled trials (N=485) compared incision and drainage of cutaneous abscesses with (n=243) and without (n=242) packing[1]. The median ages of patients in the trials ranged from 17 to 48 years old. Most studies excluded abscesses associated with Crohn's disease, immunocompromised patients, high-risk locations such as on the face or neck, and recurrent abscesses. Three studies exclusively involved anorectal abscesses with incision and drainage performed in the operating room under general or spinal anesthesia. Five studies included non-anorectal cutaneous abscesses, mostly located on the extremities, trunk, or breast, with incision and drainage performed in an emergency room or outpatient surgical center. Ribbon gauze was typically used for packing, with a change at 24 to 48 hours by a health care professional. Oral antibiotics were routinely used in four studies and selectively used in one. The remaining three studies were unclear on antibiotic use. No study reported blinding of participants and personnel, whereas blinding of outcome assessment was attempted in three studies. The primary outcome was recurrence of abscess at maximum study follow-up period (2 to 4 weeks). Secondary outcomes were development of fistula in-ano and need for a second intervention within the first 48 hours postoperatively (extension of incision, packing, irrigation, loculation break up, and admission). The overall risk of recurrence of the abscess was 5.9% and fistula formation was 17%. No significant difference was noted between nonpacking and packing in the risk of abscess recurrence (6 trials, N=289; relative risk [RR] 1.3; 95% CI, 0.53–3.2) and development of a fistula in-ano (3 trials, N=107; RR 0.61; 95% CI, 0.25–1.5). However, the need for second intervention was lower in the nonpacking group (4 trials, N=293; RR 0.70; 95% CI, 0.49–0.99). Analysis of

subgroups of anorectal abscesses, children, adults, and antibiotic use revealed no difference in abscess recurrence with nonpacking versus packing. However, a trial sequential analysis showed that the results of the analysis were not conclusive, likely related to small sample sizes and overall low pooled rate of abscess recurrence.

The Infectious Disease Society of America released updated evidence-based practice guidelines for the diagnosis and management of skin and soft tissue infections in 2014.[2] The guideline made no formal recommendation for or against packing after incision and drainage of abscess. However, the guideline pointed to one small randomized controlled trial, included in the meta-analysis above, that did not find any difference in healing outcome with packing but an increase in pain. No level of evidence was reported. EBP

John Evenocheck, MD
Amy Olmschenk, DO
UMN/St. Cloud Family Medicine Residency
St. Cloud, MN

The authors declare no conflicts of interest.

References

1. Mohamedahmed AYY, Zaman S, Stonelake S, et al. Incision and drainage of cutaneous abscess with or without cavity packing: a systematic review, meta-analysis, and trial sequential analysis of randomised controlled trials. *Langenbecks Arch Surg Epub 2020 Aug 1*. doi: 10.1007/s00423-020-01941-9. [STEP 1]
2. Stevens D, Bisno A, Patchen C, et al. Practice guidelines for the diagnosis and management of skin and soft tissue infections: 2014 update by the Infectious Diseases Society of America. *Clin Infect Dis.* 2014;59(2):e10-e52. [STEP 5]

Does neutral protamine hagedorn result in better neonatal outcomes compared with long-acting insulin (levemir/lantus) in pregnant patients with gestational diabetes?

EVIDENCE-BASED ANSWER

There does not seem to be any difference in clinical outcomes between glucose management with neutral protamine hagedorn or long-acting insulins (levemir/lantus) in pregnant patients with both gestational and pregestational diabetes (pregestational diabetes mellitus; SOR: **B**, meta-analysis of observational cohort studies, retrospective cohort study, randomized control trial).

DOI 10.1097/EBP.0000000000001367

A 2012 meta-analysis of eight observational cohort studies (N=702) examined the maternal and neonatal outcomes of pregnant women with gestational diabetes mellitus (GDM) or pregestational diabetes mellitus (pregestational DM) who were treated with either neutral protamine hagedorn (NPH) or lantus.[1] All the studies specifically examined neonatal outcomes, although not all studies examined the same outcomes. Some individual studies found differences in neonatal outcomes between treatment arms, and in all cases, the negative outcomes were more likely in the NPH group. The meta-analysis, however, found no statistically significant differences in any of the neonatal outcomes, including large for gestational age (four studies; n=381; odds ratio [OR] 1.1; 95% CI, 0.68–1.6), neonatal hypoglycemia (seven studies; n=650; OR 0.99; 95% CI, 0.63–1.6), neonatal intensive care unit (NICU) admissions (six studies; n=581; OR 0.79; 95% CI, 0.45–1.4), congenital malformations (five studies; n=508; OR 0.78; 95% CI, 0.39–1.6), hyperbilirubinemia (six studies; n=586; OR 0.93; 95% CI, 0.49–1.8), respiratory distress syndrome (six studies; n=549; OR 1.6; 95% CI, 0.82–3.2), Apgar score at five minutes <7 (four studies; n=332; OR 1.4; 95% CI, 0.26–7.1), macrosomia (four studies; n=355; OR 1.2; 95% CI, 0.71–2.0), birth weight (seven studies; n=601; mean difference [MD] 0.09 g; 95% CI, –19 to 45), and gestational age at delivery (seven studies; n=611; MD 0.09 weeks; 95% CI, –0.43 to 0.61). The major limitation of this study was its inclusion of only observational cohort studies. Also, the average HbA1C of both populations was >8%, so the conclusions of the study may be more applicable to pregestational DM than GDM.

A 2020 retrospective observational cohort study examined women with GDM or pregestational DM who had received either NPH (n=19) or long-acting insulin (levemir or lantus) (n=44).[2] The clinic treatment algorithm favored use of latntus or levemir while NPH was prescribed as an alternative. The study did not describe why prescribers chose one regimen over another. Patients with type 2 DM and GDM were required to check fasting blood sugars at least four times a day, whereas those with type 1 DM were required to check more frequently. No differences in the rate of any of the following neonatal outcomes were noted between the long-acting insulin group compared with the NPH group: macrosomia (18% vs 21%, *P*=.69), neonatal hypoglycemia (85% vs 43%, *P*=.052), hyperbilirubinemia (27% vs 36%, *P*=.48), respiratory distress syndrome (27% vs 36%, *P*=.48), and NICU admission (55% vs 64%, *P*=.28). A large limitation of this study is its small size and retrospective observational design.

A 2015 unblinded randomized controlled trial sought to determine if treatment with levemir was noninferior to NPH in 105 women with GDM or pregestational DM, with a primary outcome of maternal glycemic control.[3] Patients who failed medical nutritional therapy or oral hypoglycemic therapies were randomized to receive levemir (n=42) or NPH (n=45) between 14 and 33 weeks of gestation. The study evaluated neonatal outcomes secondarily and found no difference between the levemir and NPH groups, respectively, in any of the following domains: median birth weight (3,230 vs 3,235 g, *P*=.87), median gestational age at delivery (38.9 vs 38.8 weeks, *P*=.72), Apgar score at five minutes <7 (3% vs 0%, *P*=.48), NICU admission (7.5% vs 14%, *P*=.48), and neonatal hypoglycemia <40 mg/dL (0% vs 5%, *P*=.49). The most notable limitation of the study was that it was not adequately powered to demonstrate differences in neonatal outcomes, so noninferiority conclusions cannot be drawn with respect to these outcomes. EBP

Allison Costello, MD
Haley Ringwood, MD
Emily Lines, MD
University of Colorado Family Medicine Residency
Aurora, CO

The authors declare no conflicts of interest.

References

1. Lepercq J, Lin J, Hall GC, et al. Meta-analysis of maternal and neonatal outcomes associated with the use of insulin glargine versus NPH insulin during pregnancy. *Obstet Gynecol Int*. 2012;2012:649070. [STEP 3]
2. Sleeman A, Odom J, Schellinger M. Comparison of hypoglycemia and safety outcomes with long-acting insulins versus insulin NPH in pregestational and gestational diabetes. *Ann Pharmacother*. 2020;54(7):669-675. [STEP 3]
3. Herrera KM, Rosenn BM, Foroutan J, et al. Randomized controlled trial of insulin detemir versus NPH for the treatment of pregnant women with diabetes. *Am J Obstet Gynecol*. 2015;213(3):426.e1-426.e7. [STEP 2]

What are the nonpharmacologic methods to increase milk supply?

EVIDENCE-BASED ANSWER

In mothers of preterm newborns, breast milk production may be increased with maternal relaxation (meditation with or without music and images), increasing pumping frequencies from 5 to 7 times per day (300 g per day), and using a hands-on technique in combination with electronic pumping (280 mL per day; SOR: **B**, individual randomized controlled trial [RCT] and observational trials). In mothers of term infants, breast milk production may be increased by on-demand feeding compared with scheduled feeding every 3 to 4 hours (220 mL per day) and higher pumping pressures (70 g per day; SOR: **B**, individual RCTs).

DOI 10.1097/EBP.0000000000001358

A 2012 randomized controlled trial (RCT) of exclusively pumping mothers with infants admitted to the neonatal intensive care unit (NICU) for prematurity or critical illness evaluated the effect of relaxation on milk production.[1] Women were assigned to the control of standard

pumping conditions (group A, n=43) or one of three experimental conditions while pumping: guided meditation recordings (group B, n=42), group B plus guitar music and lullabies (group C, n=40), or group C plus images of the infant (group D, n=43). Mothers in all relaxation intervention groups produced more milk than the control group on days 2 to 14 of the study period. On day two, breast milk production in the intervention groups as a percentage of the control group were as follows: (B) 184% (P=.0004), (C) 227% (P=.0001), and (D) 186% (P=.0005). By day 14, breast milk production in the intervention groups was statistically significant more than the group A control of 318 mL: (B) 591 mL (186% of control), (C) 1,028 mL (323% of control), and (D) 861 mL (271% of control), each with P<.0001 compared with the control group. External validity for this study was limited by inclusion of only pumping-dependent mothers with infants in the NICU.

A 2001 secondary data analysis from a prospective observational study assessed the milk production and pumping frequency of 39 pump-dependent mothers of low-birth-weight (<1,500 g) and preterm (<32 weeks) infants.[2] Patients were separated into groups based on self-reported pumping frequency for five weeks postpartum: high-frequency pumping (average 7.0 sessions per day) versus low-frequency pumping (average 4.9 sessions per day) and early breastfeeding initiation (≤48 hours) versus late initiation (>48 hours). Mothers in the high-frequency pumping group had higher mean daily milk weights (632 vs 319 g; P<.01). Overall milk production positively correlated with frequency of pumping in weeks 2 to 5 (17,831 vs 9,098 g; P=.03). Limitations of this study included small size, reliance on self-reported behaviors, and including only low-birth-weight and preterm infants.

A 2009 observational study examined the effectiveness of hands-on pumping while using an electronic pump for 52 mothers of infants less than 31 weeks and less than 1,500 g.[3] At three weeks postpartum, all mothers were taught hands-on pumping technique defined as manual massage of the breasts while pumping. Milk volumes at eight weeks were compared with preintervention volumes and demonstrated increased milk yield of 48% (583 vs 863 mL, P<.003). The limitations of this study included lack of control group, use of a convenience sample of participants, and including only preterm, low-birth-weight infants and pump-dependent mothers.

A 1983 RCT evaluated the effect of feeding on demand on early milk production and infant weight gain among mothers with term infants and normal deliveries.[4] The intervention group were taught infant hunger cues and instructed to feed on demand (n=28 at enrollment; n=12 at day 35). The control group received standard care for this hospital, which instructed mothers to feed every 3 to 4 hours (n=47 at enrollment; n=16 at day 35). The intervention group had significantly more feedings per day (9.9 vs 7.3; P<.001) on days 1 to 14, increased milk production (725 vs 502 mL; P<.001), and increased infant weight gain (561 vs 347 g; P<.02). At day 35, the intervention group continued to have increased number of feedings; however, milk production and weights were not significantly different. The validity of this study was limited by high dropout rates.

A 2017 RCT of 148 first-time mothers of term infants after cesarean section evaluated the effect of pumping pressure on the onset of lactation and milk volume.[5] Patients were assigned to three groups: pumping using high pressure (–150 mmHg, n=50), pumping using low pressure (–100 mmHg, n=48), or control group without pumping (n=50). Mothers in the pumping groups started to pump within two hours after delivery, pumped every three hours while awake, and every 4 to 6 hours overnight for 30 minutes in addition to breastfeeding their infants. Compared with the no pumping group, time to onset of lactation was significantly earlier in both the high pressure (53 vs 70 hours; P<.0001) and the low pressure (60 vs 70 hours; P<.0003) groups. The intervention groups also produced significantly higher milk volumes (high-pressure group: 157 g, low-pressure group: 103 g, control group: 84 g; analysis of variance P<.001). However, the low-pressure group was not statistically different than the control group (P=.84). Mothers in the high-pressure pumping groups reported increased nipple pain and trauma. EBP

Chelsea Roberts, MD
Heather Bleacher, MD MPH
Roxanne Radi, MD MPH
University of Colorado Family Medicine Residency
Denver, CO

The authors declare no conflicts of interest.

References

1. Keith DR, Weaver BS, Vogel RL. The effect of music-based listening interventions on the volume, fat content, and caloric content of breast milk-produced by mothers of premature

and critically ill infants. *Adv Neonatal Care*. 2012;12(2):112-119. [STEP 2]

2. Hill PD, Aldag JC, Chatterton RT. Initiation and frequency of pumping and milk production in mothers of non-nursing preterm infants. *J Hum Lact*. 2001;17(1):9-13. [STEP 3]
3. Morton J, Hall JY, Wong RJ, Thairu L, Benitz WE, Rhine WD. Combining hand techniques with electric pumping increases milk production in mothers of preterm infants. *J Perinatol*. 2009;29(11):757-764. [STEP 3]
4. De Carvalho M, Robertson S, Friedman A, Klaus M. Effect of frequent breast-feeding on early milk production and infant weight gain. *Pediatrics*. 1983;72(3):307-311. [STEP 3]
5. Zhang F, Yang Y, Bai T, et al. Effect of pumping pressure on onset of lactation after caesarean section: a randomized controlled study. *Matern Child Nutr*. 2018;14(1):e12486. [STEP 2]

In adults with vitamin B_{12} deficiency, is oral vitamin B_{12} as effective as intramuscular vitamin B_{12} for normalizing B_{12} levels?

EVIDENCE-BASED ANSWER

In adults with vitamin B_{12} deficiency, oral vitamin B_{12} is as effective as intramuscular vitamin B_{12} for normalizing vitamin B_{12} levels (SOR: **C**, based on a systematic review with low-quality evidence, and 1 randomized controlled trial [RCT]). No significant difference was noted between oral and intramuscular vitamin B_{12} in quality of life, signs of B_{12} deficiency, or treatment satisfaction (SOR: **B**, based on 1 RCT).

DOI 10.1097/EBP.0000000000001392

A 2018 Cochrane review of three randomized controlled trials (n=153) compared vitamin B_{12} levels in oral versus intramuscular treatment groups in adults with vitamin B_{12} deficiency.[1] In a Turkish study of patients with megaloblastic anemia (n=60, average age 62 years old), patients received an oral or intramuscular vitamin B_{12} dose of 1,000 μg daily for 10 days, then weekly for four weeks, then monthly, for a total of 15 mg within 90 days on either treatment (1:1 oral-to-intramuscular ratio). Mean vitamin B_{12} levels improved similarly with oral and intramuscular therapy: from 72.9 to 214 pg/mL (oral) and from 70.2 to 226 pg/mL (intramuscular) (mean difference [MD] −11.7; 95% CI, −29.5 to 6.1). A New York study of patients from ambulatory care centers (n=33, average age 71 years old) compared oral vitamin B_{12} 2,000 μg per day for 120 days with IM vitamin B_{12} 1,000 μg given nine times over 90 days (a 27:1 total oral-to-intramuscular ratio). Mean vitamin B_{12} improved more with oral compared with intramuscular therapy, 93 to 1,005 pg/mL versus 95 to 325 mg/mL (MD 680 pg/mL; 95% CI, 393–967). An Indian study of patients at a tertiary care hospital (n=60, average age 42) compared oral vitamin B_{12} 1,000 μg per day for three months (total dose 90 mg) with an intramuscular vitamin B_{12} dose of 1,000 μg (total dose 15 mg within 3 months of treatment), a 6:1 total oral-to-intramuscular ratio. Oral treatment normalized B_{12} levels (defined as ≥200 pg/mL) less compared with intramuscular treatment (67% vs 90% of patients, P=.06), but this difference was not significant. The reviewers determined that low-quality evidence showed oral and intramuscular treatment had similar effects in normalizing vitamin B_{12} levels. Only the Turkish study commented on adverse outcomes and reported none. The clinical applicability of this review was limited by the small number of studies and small number of participants, short treatment durations, and variations in definitions of deficiency, ages of patients, and dosing regimens. Because of the heterogeneity of the studies, the researchers could not pool the outcomes.

A 2019 randomized noninferiority trial in 22 primary care clinics in Spain (n=283) compared intramuscular to oral vitamin B_{12} for 52 weeks.[2] Patients were ≥65 years old (average age 75.2 years), 58.3% female, and had vitamin B_{12} levels <211 pg/mL. Exclusion criteria included treatment of vitamin B_{12} deficiency within the past five years, malabsorption, serious neurological or psychiatric symptoms, blood diseases, serum folic acid <2.3 ng/mL, optic nerve atrophy, and stage 4 or greater chronic kidney disease. Patients in the intramuscular group received a nurse injection of 1 mg IM vitamin B_{12} on alternate days for two weeks, then 1 mg per week for six weeks, then 1 mg per month for weeks 9 to 52. The oral group self-administered 1 mg vitamin B_{12} daily for eight weeks then 1 mg per week for weeks 9 to 52. The total oral-to-intramuscular dosing ratio was about 5:1. The primary outcome was normalization of vitamin B_{12} levels, defined as a vitamin B_{12} concentration ≥211 pg/mL. Noninferiority was present if the CI was not outside 10%. At eight weeks, vitamin B_{12} levels normalized

in 95.0% with oral versus 90.2% with IM treatment (difference of 4.8%; 95% CI, −1.3 to 10.9). At 52 weeks, vitamin B_{12} levels normalized in 73.6% with oral versus 80.4% with intramuscular treatment (difference of −6.8%; 95% CI, −16.6 to 2.9). Because the sample was 88.4% of the calculated necessary size, a Bayesian analysis was performed, which determined that the probability of a >10% difference in treatment effectiveness was 0.001 and 0.060 at 8 and 52 weeks, respectively. The authors concluded that oral treatment was as effective as IM at eight weeks and that the likelihood of differences at 52 weeks below the noninferiority threshold was very low. EBP

Thomas P. Garigan, MD, MA, FAAFP
Gregory J. C. Sue, DO
Dwight D. Eisenhower Army Medical Center, Fort Gordon, GA

The authors declare no conflicts of interest.
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 U.S. Army Medical Department, the Army at large, or the Department of Defense.

References

1. Wang H, Li L, Qin L, Song Y, Vidal-Alaball J, Liu T. Oral vitamin B versus intramuscular vitamin B for vitamin B deficiency. *Cochrane Database Syst Rev*. 2018;3(3): CD004655. [STEP 2]
2. Sanz-Cuesta T, Escortell-Mayor E, Cura-Gonzalez I, et al. Oral versus intramuscular administration of vitamin B12 for vitamin B12 deficiency in primary care: a pragmatic, randomised, non-inferiority clinical trial (OB12). *BMJ Open*. 2020; 10(8):e033687. [STEP 2]

At what HbA1c level should patients with newly diagnosed type II diabetes mellitus be started on insulin therapy rather than oral medications?

EVIDENCE-BASED ANSWER

Two to three weeks of short-term intense insulin therapy, for patients with newly diagnosed type 2 diabetes with an average HbA1c>9.5%, can induce remission in up to half of patients (SOR: **B**, limited quality patient and disease-oriented evidence). Consensus guidelines recommend initiating insulin therapy in patients with HbA1c>9.0% who present with hyperglycemic symptoms (polyuria, polydipsia, or polyphagia; SOR: **C**, expert opinion).

DOI 10.1097/EBP.0000000000001324

A systematic review and meta-analysis (2 randomized controlled trials [RCTs], 5 single-arm interventional trials; N=839) investigated the effect of short-term insulin therapy (2–3 weeks) on β-cell function, insulin resistance, and glycemic remission.[1] Studies included adults (18 years old or older) with newly diagnosed type 2 diabetes. The RCTs used oral hypoglycemic treatments as the controls, to which the insulin therapies were compared. Studies were excluded if they included patients who were not newly diagnosed with type 2 diabetes or if the data were the result of dissimilar assessment models. This meta-analysis used the Homeostasis Model Assessment of β-cell function, or Insulin Resistance (HOMA-B, HOMA-IR), to reflect changes in β-cell function, insulin resistance, or disease remission, resulting from short-term intensive insulin therapy. HOMA-B and HOMA-IR are widely validated assessments based on plasma fasting glucose and insulin levels. An increase in β-cell function is reflected by HOMA-B results greater than zero, and a decrease in insulin resistance is reflected by HOMA-IR less than zero. The analysis showed a 1.13% (95% CI, 1.02–1.25) change in β-cell function from baseline and a –0.57% (95% CI, —0.84 to –0.29) change in insulin resistance. Additionally, in the 4 studies (1 RCT and 3 single-arm studies) that assessed glycemic remission (N=559), 66% of patients (292 of 441) did not require glycemic medication after three months, 59% (222 of 377) after six months, 46% (229 of 495) after 12 months, and 42% (53 of 126) after 24 months. All studies were conducted in Asian and might not be generalizable.

A multicenter randomized parallel-group clinical trial between 2004 and 2006 in China (n=382) assessed the effectiveness of transient intensive insulin therapy in

patients newly diagnosed with diabetes compared with oral hypoglycemic agents on β-cell function and diabetes remission rate.[2] Patients with an HbA1c of 9.5% to 9.8% were enrolled in the study and randomly assigned to one of three groups: continuous subcutaneous insulin infusion (CSII), multiple daily injections (MDI), or oral hypoglycemic agents (OHA). In both insulin groups, the total daily dose of insulin was 0.4 to 0.5 IU/kg—50% basal and 50% bolus doses in the insulin infusion group, and 30% to 20%, and 30%, 20%, 20% pre-meal, and 30% nighttime insulin in the MDI group. The oral hypoglycemic group, with body mass index (BMI) 20 to 25 kg/m^2, was initially treated with gliclazide 80 mg twice a day and titrated up to a maximum dose of 160 mg twice a day to achieve glycemic control. Those with BMI 25 to 35 kg/m^2 were initially treated with metformin 500 mg twice a day and titrated up to a maximum dose of 2,000 mg daily. A combination of gliclazide and metformin was used in patients who could not reach the glycemic control goal with one oral agent or who had HbA1c >8.6%. More patients achieved glycemic control (HbA1c ≤8.0%) in the insulin groups (97.1% in CSII and 95.2% in MDI) in less time (4.0 days with infusion and 5.6 days with multiple daily injections) than those treated with oral agents (82.6% in 9.3 days; $P<.0001$ vs infusion and $P=.01$ vs multiple injections). Remission rates, defined as patients postintervention who maintained optimum glycemic control (fasting plasma glucose <90 mg/dL, and 2-hour postprandial glucose <180 mg/dL) for at least 12 months without medication, were significantly higher in insulin groups (51.1% and 44.9% in CSII and MDI, respectively) than in the oral agent group (26.7%; $P=.0012$). This clinical trial was used in the meta-analysis above, however, was included here because it detailed the difference in outcomes between the two groups: insulin therapies compared with oral agents.

Since 2013, the American Academy of Clinical Endocrinologists and American College of Endocrinology have periodically released recommendations for comprehensive management of patients with type 2 diabetes.[3] For newly diagnosed patients with HbA1c >9.0% (>11.7 mmol/L) who are symptomatic (presenting with polyuria, polydipsia, or polyphagia), the 2020 update recommended initiation of insulin with or without additional diabetes medication. EBP

James Mikolajczak, DO
Joel Herness, MD
Nellis FMR, Las Vegas, NV

The authors declare no conflicts of interest.

References

1. Kaercher-Kramer C, Zinman B, Renakaran R. Short-term intensive insulin therapy in type 2 diabetes mellitus: a systematic review and meta-analysis. *Lancet Diabetes Endocrinol.* 2013;1:28-34. [STEP 2]
2. Weng J, Li Y, Xu W, et al. Effect of intensive insulin therapy on ß-cell function and glycaemic control in patients with newly diagnosed type 2 diabetes: a multicentre randomised parallel-group trial. *Lancet.* 2008;371:1753-1760. [STEP 2]
3. Garber A, Handelsman Y, Grunberger G, et al. Consensus Statement by the American Association of Clinical Endocrinologists and American College of Endocrinology on the Comprehensive Type 2 Diabetes Management Algorithm - 2020 Executive Summary. *Endocr Pract*. 2020;26(1):107-139.

Is vaginal vitamin E effective in reducing the symptoms of genitourinary syndrome of menopause?

EVIDENCE-BASED ANSWER

Vaginal vitamin E may be more effective in reducing the genitourinary symptoms of menopause than placebo (SOR: **C**, small randomized controlled trial [RCT]) but seems no better than vaginal estrogen (SOR: **B**, 2 small RCTs) and is perhaps less effective than hyaluronic acid (SOR: **C**, small RCT).

DOI 10.1097/EBP.0000000000001412

A 2013 double-blinded randomized controlled trial (RCT; n=42) evaluated the effectiveness of vitamin E vaginal suppositories for treating atrophic vaginitis.[1] The trial included sexually active, postmenopausal women (45–65 years old) from a single center in Iran with amenorrhea for 12 months or follicle stimulating hormone (FSH) level greater than 40, normal Pap smear in past three years, symptoms of vaginal atrophy, and vaginal pH greater than 5. Patients received either suppositories containing a semisynthetic fatty acid

triglyceride and 1 mg (2.22 IU) of vitamin E (n=20) or placebo suppositories containing only the semisynthetic fatty acid triglyceride (n=22). Suppositories in both groups were administered daily for two weeks, followed by every other day for six weeks. The primary outcome was the change in patient-rated vaginal symptoms, including irritation, itching, dryness, and dyspareunia. Each symptom was rated on a scale of 1 to 4 (more severe symptoms), and the composite score was compared before and after treatment. At eight weeks, the vitamin E group had a significantly lower symptom score with an intergroup mean difference [MD] of 3.0 (95% CI, 2.7–3.3). No adverse effects were reported. Limitations included recruitment from a single center, potentially limiting generalization of outcomes.

A 2012 double-blinded RCT (n=40) compared effectiveness of vitamin E to hyaluronic acid (HA) in sexually active postmenopausal women (45–65 years old) from a single center in Iran.[2] Inclusion criteria were the same as the above referenced study. The HA group (n=20) vaginal suppositories contained 5 mg HA, and the vitamin E group (n=20) vaginal suppositories contained 1 mg vitamin E. Suppositories in both groups were administered daily for two weeks, followed by three times weekly for the remaining six weeks. The primary outcome was the change in patient-rated symptoms via the same composite score as the above study, measured at baseline and after eight weeks. At eight weeks, the patient symptom scores improved more with HA than with vitamin E (intergroup MD 2.8; 95% CI, 2.2–3.3). No side effects were reported for either intervention. Limitations included recruitment from a single center potentially limiting generalization of outcomes.

A 2019 single-blinded RCT (n=52) compared effectiveness of vitamin E with vaginal estrogen cream for alleviation of genitourinary symptoms.[3] The trial included sexually active, postmenopausal women (40–65 years old) from multiple centers in Iran with amenorrhea for at least 12 months (24 months if age <50 years old), FSH serum level greater than 40 IU, a normal Pap smear in the last three years, and vaginal pH >5. Patients received either vitamin E 100 IU vaginal suppositories (n=26) or 0.5 g vaginal estrogen cream (n=26). Both groups received daily treatment for two weeks, followed by twice weekly for 10 weeks. Patient symptoms in areas of sexual desire, sensation, lubrication, and orgasm were assessed through the Abbreviated Sexual Function Questionnaire (ASFQ, range 2–36), with higher scores indicating better sexual function. This was measured at the baseline, four weeks, eight weeks, and 12 weeks. At completion of therapy, ASFQ score of the vitamin E group was no different than the estrogen group (MD –0.19; 95% CI, –4.4 to 4.0). The study was limited by lack of patient blinding.

A 2014 single-blinded RCT (n=52) compared effectiveness of vitamin E suppositories with conjugated vaginal estrogen cream for improving menopausal symptoms in healthy postmenopausal women (40–65 years old) from multiple centers in Iran.[4] Inclusion criteria were similar to the above referenced 2019 study. Patients received either 100 IU vitamin E suppositories (n=22) or 0.5 g vaginal estrogen cream (n=21). Both treatments were administered daily for two weeks, followed by twice a week for the next 10 weeks. Vasomotor, psychosocial, physical, and sexual symptoms were assessed through the Menopause-Specific Quality of Life questionnaire (scoring range 0–174), with lower total scores indicating better quality of life. At week 12, no significant difference was noted between vitamin E and estrogen groups (intergroup MD of –2.6; 95% CI, –7.0 to 1.7). No side effects were noted. The study was limited by lack of patient blinding. EBP

Laura Porterfield, MD
Zuleica Santiago Delgado, MD
Amanda Song, MD
Farha Syed, MD
Nyajuok Wur, MD
University of Texas Medical Branch, Galveston, TX

The authors declare no conflicts of interest.

References

1. Ziagham S, Abbaspoor Z, Safyari S, et al. Effect of vitamin E vaginal suppository on atrophic vaginitis among postmenopausal women. *Jundishapur J Chronic Dis Care*. 2013; 2:11-19. [STEP 2]
2. Ziaghma S, Abbaspoor Z, Abbaspoor M. Effect of hyaluronic acid and vitamin E vaginal tablets on atrophic vaginitis: a randomized controlled trial. *Afr J Pharmacol*. 2012;6(45): 3124-3129. [STEP 2]
3. Golmakani N, Parnan Emamverdikhan A, Zarifian A, et al. Vitamin E as alternative local treatment in genitourinary syndrome of menopause: a randomized controlled trial. *Int Urogynecol J*. 2019;30(5):831-837. [STEP 2]
4. Parnan Emamverdikhan A, Golmakani N, SharifiSistani N, et al. Comparing two treatment methods of vitamin E suppository and conjugated estrogen vaginal cream on the quality of Life in menopausal women with vaginal atrophy. *J Midwifery Reprod Health*. 2014;2(4):253-261. [STEP 2]

Is metformin effective for treating infertility associated with PCOS?

EVIDENCE-BASED ANSWER

Metformin may be better than placebo for increasing live-birth rates in infertile women with polycystic ovary syndrome (PCOS); however, metformin is probably less effective than clomiphene citrate (CC). Adding metformin to CC does not improve live birth success more than CC alone (SOR: **B**, meta-analysis of very low- to moderate-quality randomized controlled trials [RCTs]). Furthermore, gastrointestinal side effects are more common with metformin than placebo (SOR: **B**, meta-analysis of very low- to moderate-quality RCTs). Experts note that metformin increases ovulation rates in women with PCOS when compared with placebo but recommend against its use as first-line therapy for anovulation because ovulation induction agents such CC or letrozole are more effective in increasing ovulation, pregnancy, and live-birth rates (SOR: **C**, clinical practice guideline).

DOI 10.1097/EBP.0000000000001328

A systematic review and meta-analysis of 41 randomized controlled trials (RCTs; N=4,552) examined the effectiveness and safety of metformin in improving live-birth rates in women with polycystic ovary syndrome (PCOS) who were undergoing ovulation induction.[1] PCOS was defined by Rotterdam Consensus criteria by the presence of at least two of the following: (1) polycystic ovaries seen on ultrasonography, (2) oligo-ovulation or anovulation, and (3) clinical or biochemical signs of hyperandrogenism. The trials were conducted in 25 countries with a range of 18 to 626 women per study; the women were 24 to 33 years old with a body mass index of 21 to 39 kg/m^2. Three main interventions and comparisons were done: metformin compared with placebo or no treatment, metformin compared with clomiphene citrate (CC), and the combination of metformin and CC compared with CC alone. Study duration averaged 19.7 weeks (range 4–96 weeks). Metformin doses ranged from 850 to 2,000 mg per day; the most common dose (used in 21 trials) was 1,500 mg per day. Metformin yielded a borderline improvement in live-birth rates compared with placebo (4 trials; N=435; odds ratio [OR] 1.6; 95% CI, 1.0–2.5); however, metformin seemed to be less effective than CC alone (5 trials; N=741; OR 0.7; 95% CI, 0.5–1.0), and metformin combined with CC did not increase live-birth rates more than CC alone (10 trials; N=1,219; OR 1.3; 95% CI, 0.98–1.7). Furthermore, the addition of metformin increased gastrointestinal side effects when compared with placebo (6 trials; N=852; OR 4.3; 95% CI, 2.8–6.4) and CC (7 trials; N=713; OR 4.0; 95% CI, 2.6–6.1). The reviewers deemed the evidence quality to be very low to moderate. This systematic review reported limitations caused by imprecision, inconsistency, and risk of bias from poor reporting on methodology and incomplete outcome data.

A 2017 evidence-based guideline from the American Society of Reproductive Medicine on the role of metformin for ovulation induction in infertile women with PCOS found insufficient evidence to recommend metformin over placebo to improve pregnancy rates and live births (SOR: **C**, insufficient evidence supporting recommendation based on underpowered and inconsistent RCTs).[2] CC alone was regarded as superior to metformin for achieving ovulation induction, clinical pregnancy, and live births (SOR: **B**, fair evidence supporting recommendation based on one large well-designed RCT). The authors concluded that metformin combined with CC improved ovulation and pregnancy rates, but not live births, when compared with CC alone (SOR: **A**, good evidence supporting recommendation based on several RCTs). It was unclear if metformin alone was more or less effective than letrozole alone in improving pregnancy and live-birth rates (SOR: **C**, insufficient evidence supporting recommendation), but letrozole alone was noted to likely enhance ovulation (SOR: **B**, fair evidence supporting recommendation based on 1 large well-designed RCT). The guideline was based on a systematic review of the literature with explicit inclusion and exclusion criteria and included 73 studies that were and evaluated by a four-person independent task force. EBP

Robert Crouch, DO
Family Medicine Residency of Idaho, Nampa, ID

Julia Hamilton, DO
Providence St. Peter Family Medicine Residency Olympia, WA

Tyler Raymond, DO, MPH, FAAFP
Madigan Army Medical Center Family Medicine Residency, Tacoma, WA

The authors declare no conflicts of interest.
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 Army Medical Department, the Army at large, or the Department of Defense.

References

1. Sharpe A, Morley LC, Tang T, Norman RJ, Balen AH. Metformin for ovulation induction (excluding gonadotrophins) in women with polycystic ovary syndrome. *Cochrane Database Syst Rev.* 2019;(12):CD013505. [STEP 1]
2. American Society for Reproductive Medicine. Role of metformin for ovulation induction in infertile patients with polycystic ovarian syndrome (PCOS): a guideline. *Fertil Sterility*. 2017;108(3): 426-441. [STEP 5]

Is the levonorgestrel intrauterine system effective in decreasing menorrhagia caused by uterine leiomyoma (fibroids)?

EVIDENCE-BASED ANSWER

The levonorgestrel-releasing intrauterine system (LNG-IUS) reduces menorrhagia associated with non-submucosal uterine leiomyomas by 50% to 99.5% during a 48-month period compared with baseline (SOR: **B**, systematic review of primarily observational studies). LNG-IUS may yield a greater reduction in lost days of work/activity per month compared with baseline than does combined oral contraceptives (6.9 vs 2.0; SOR: **C**, small randomized controlled trial). In perimenopausal patients, LNG-IUS may reduce hysterectomies for menorrhagia by up to 90% and result in greater patient satisfaction than hysterectomy (SOR: **C**, small cohort study).

DOI 10.1097/EBP.0000000000001415

A 2014 systematic review evaluated the efficacy and safety of the levonorgestrel-releasing intrauterine system (LNG-IUS; Mirena: releasing 20 μg/d) as treatment for premenopausal women with symptomatic leiomyoma.[1] Menstrual blood loss (MBL) was evaluated in one randomized controlled trial (RCT) and seven of the included observational studies (N=413). The studies included women (20–54 years old) with uterus size similar to a 12 weeks' gestation size, a solitary leiomyoma up to 2.5 cm, or multiple leiomyoma up to 1.5 cm each. Patients were treated with LNG-IUS, combined oral contraceptive, or thermal balloon ablation. MBL was evaluated using the validated pictorial blood loss assessment charts (PBACs). The PBAC is semi-quantitative, in that women count the number of used pads or tampons each day and assign scores based on a visual estimate of soiling. Points are also given for clots based on size and for episodes of "flooding." Six of the observational studies (N=333) showed a 50% to 91% reduction in MBL at six months, and a 69% to 97% reduction at 12 months (compared with the same women before LNG-IUS insertion). One study comparing LNG-IUS for idiopathic versus leiomyoma-associated menorrhagia found similar reductions in MBL in both groups. Within one month, the leiomyoma group's MBL fell by 87% ($P<.0001$), and at three, 12, 24, 36, and 48 months, the MBL was reduced by 92%, 97%, 97%, 99.5%, and 99.5%, respectively. One prospective historical control study found LNG-IUS to be as effective as thermal ablation for leiomyoma-associated menorrhagia after 12 months. Limitations included small sample sizes, lack of blinding, or lack of control groups. The smallest study had a higher loss to follow-up rate, higher device expulsion rate, and a high degree of MBL before intervention.

A 2011 RCT (n=58), included in the above systematic review, is summarized separately here because this RCT is the only non-pilot RCT addressing this topic.[2] This single-center, open, parallel-treatment RCT compared LNG-IUS to combined oral contraceptive (30 μg of ethinyl estradiol and 150 μg of levonorgestrel) over 12 months. The trial included Egyptian women (20–50 years old) with menorrhagia, leiomyoma <5 cm, no uterine distortion, and a regular cycle. The primary outcome included change in MBL as measured by alkaline hematin and PBAC, and lost days (inability to work or perform daily activities in the last 30 days). LNG-IUS reduced MBL greater than oral contraceptives when directly measured by alkaline hematin at 12 months (91% vs 13%; $P<.001$) and when evaluated

by PBAC (88% vs 54%; $P=.02$). Hemoglobin levels increased in the LNG-IUS group from 9.7 to 12 g/dL ($P<.001$) but did not change in the oral contraceptive group. No difference in treatment failure rates was observed (23% for LNG-IUS; 38% for oral contraceptive; $P=.101$). Lost days improved in both groups (LNG-IUS 8.2–1.3 days; $P=.003$; oral contraceptive 8.3–6.3 days; $P<.001$). The sample size was powered for an anticipated 15% loss to follow-up, but loss rates were 24.1% in the LNG-IUS group and 27.6% in the oral contraceptive group. This RCT is also limited by absence of blinding, potential measurement bias that could inflate the reported efficacy of oral contraceptives, and the fact that the LNG-IUS and pads were provided by their manufacturers.

A 2013 prospective longitudinal cohort study (n=60) evaluated use of LNG-IUS in Brazilian perimenopausal women referred by primary care for hysterectomy to treat menorrhagia associated with leiomyoma.[3] Women with submucous leiomyoma were excluded. All participants were offered LNG-IUS in lieu of hysterectomy. Thirty-nine patients (65%) chose LNG-IUS. The remaining women had either abdominal or vaginal hysterectomy. No loss to follow-up was noted. The primary outcome was the percentage of hysterectomies avoided by use of LNG-IUS, with a mean follow-up of 28 months. In the first six months, four patients (10%) using LNG-IUS opted for hysterectomy; all for unacceptable bleeding, and one of those also had expulsion of the LNG-IUS. After 24 months, the remaining 35 patients (90%) still preferred LNG-IUS. At six months, three of 35 (9%) had amenorrhea; at 24 months, 22 of 35 (63%) had amenorrhea, and none had excessive bleeding. Patient satisfaction was assessed by intention-to-treat after six months using a five-point Likert-type scale. More patients using LNG-IUS were satisfied or very satisfied (90% vs 71%; $P=.02$). However, a higher percentage of hysterectomy patients were very satisfied (33% vs 13%; $P=.02$). At six-month follow-up, five of 21 patients (24%) in the hysterectomy group (and none in the LNG-IUS group) reported frequent abdominal pain. The patients who opted for hysterectomy were slightly younger (42 vs 45 years old; $P=.04$), had longer duration of irregular bleeding before treatment (22 vs 18 months; $P=.02$), and larger uterine volumes (238 vs 213 cm^3; $P=.04$). These factors may have influenced their decision to have a hysterectomy. No comparison was done between vaginal and abdominal surgical approaches. Given this study population was all perimenopausal, if the results are applicable to younger women remains unknown. EBP

Dillon Savard, MD, FAAFP
Ashley Brown, MD, MPH
Offutt Air Force Base Family Medicine Residency Program, Bellevue, NE

The authors declare no conflict of interest.
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 Air Force Medical Department, the Air Force at large, or the Department of Defense.

References

1. Wenxiao J, Qi S, Miaomiao C, et al. Levonorgestrel-releasing intrauterine system use in premenopausal women with symptomatic uterine leiomyoma: a systematic review. *Steroids*. 2014;86:69-78. [STEP 1]
2. Sayed G, Zakherah M, El-Nashar S, et al. A randomized clinical trial of a levonorgestrel-releasing intrauterine system and a low-dose combined oral contraceptive for fibroid-related menorrhagia. *Int J Gynecol Obstet*. 2011;112(2):126-130. [STEP 2]
3. Machado R, de Souza I, Beltrame A, et al. The levonorgestrel-releasing intrauterine system: its effect on the number of hysterectomies performed in perimenopausal women with uterine fibroids. *Gynecol Endocrinol*. 2013;29(5):492–495. [STEP 3]

Is a plant-based diet safe in pediatric populations?

EVIDENCE-BASED ANSWER

Children eating plant-based diets have similar height, weight, and nutrient/health status compared with omnivorous children (SOR: **A**, systematic review of observational studies and cross-sectional study). Well-planned vegetarian diets can meet nutritional requirements, support normal growth, and age-appropriate development (SOR: **C**, consensus guideline). Although most vegetarian-type diets do not have detrimental effects and are thus not unsafe, nutritional supplementation may be needed (SOR: **B**, systematic review of observational studies and consensus guideline).

DOI 10.1097/EBP.0000000000001425

A 2017 systematic review of 24 observational studies, consisting of 12 cohort studies and 12 cross-sectional

studies, examined biomarkers, anthropometry, and dietary intake of vegetarian children (N=11,435).[1] Studies included children from 0 to 18 years old with a median sample size of 35. Inclusion criteria were vegetarian-type diets (nonspecified vegetarian, Seventh Day Adventist, lacto-ovo-vegetarian, and vegan) and characteristics of nutritional or health status. Among the studies, nutrient intake was assessed by measurement of vitamin C, folate, dietary fiber, vitamin B_{12}, vitamin D, and iron intake. Physical development (measured by height and weight) was compared with omnivorous or nonvegetarian control groups. Nutrient and health status was measured by various serum measurements (fatty acid profile, bone formation markers, vitamin B_{12}, iron, and vitamin D), blood pressure, general fitness, sexual development, and age of menarche. In studies with nonspecified vegetarian diets (7 studies), vitamin B_{12}, calcium, and iron intake were less than or equal to recommended values, whereas vitamin C and folate intake were higher than recommended reference values. The studies with vegan diets (2 studies) found daily intake of vitamin B_{12}, vitamin D, and vitamin A less than control groups when not supplemented, but greater than or equal to control groups when supplemented. Of note, bone mineral content and bone mineral density were less than the recommended values in vegan diets. Height and weight journals showed children on nonspecified vegetarian (7 studies), vegan (2 studies), lacto-ovo-vegetarian (3 studies), and Seventh Day Adventist (4 studies) diets to be of equal height and equal or below weight compared with omnivorous control groups. Nutrient and health serum markers were at or above control ranges (except for serum vitamin D, which was lower than recommended reference). Data points were not provided. Limitations included small individual sample sizes of most studies, volunteer status of all participants, and heterogenous nature of most outcome measures.

A 2019 cross-sectional study examined energy intake, macronutrient intake, and anthropometrics of German vegetarian, vegan, and omnivorous children 1–3 years old (n=430).[2] A three-day weighed dietary record assessed diet, lifestyle, body weight, and height. The children's parents weighed and recorded all food and drink consumed, with those children who were breastfeeding getting their breast milk weighed. Dietary anthropometrics were measured by their parents or pediatricians during the latest check-up. After adjustment for confounders, omnivorous children had the highest median intake of protein when compared with vegetarian and vegan children (2.7 g/kg/body weight versus each 2.3 and 2.4 g/kg body weight, $P<.0001$). Added sugars were higher in omnivorous children compared with vegan children (5.3% energy intake vs 3.8%; $P=.002$). Vegan children had the higher intake of carbohydrates compared with omnivorous children (56% vs 50% daily energy intake, $P<.0001$). Both vegetarian and vegan children had higher fiber intake than omnivorous children (17% and 22% vs 12% g/kcal, $P<.0001$). Although more omnivorous children were classified as overweight, median weight-for-height, height-for-age, and weight-for-age z-scores did not differ between vegan, vegetarian, and omnivorous children.

A 2019 Consensus paper from the Nutrition Committee of the German Society of Pediatric and Adolescent Medicine reviewed the effects of vegetarian diets on infants, children, and adolescents.[3] The committee made several observations based on limited data from studies done in the 1980s and 1990s. Because of changes in fortification and increased availability of processed plant-based foods, they concluded that these studies were not applicable to the current plant-based diet population. According to the committee, inherent risk of nutrient deficiency was noted in children with plant-based diets, and this grows with an increasing level of dietary restriction. Caution was advised in vegetarian children when considering nutrients only found in animals (vitamin B_{12}), nutrients only found in small amounts in plants (calcium), and nutrients that may not be absorbed well (iron, zinc). They further stated that vegetarian children should eat a higher amount of iron to compensate for this discrepancy, and breastfeeding infants of vegan mothers should have vitamin B_{12} supplemented. No strength of recommendation was provided. EBP

Austin C. Davis, DO
Mitchell Duininck, MD
In His Image

The authors declare no conflicts of interest.

References

1. Schurmann S, Kersting M, Alexy U. Vegetarian diets in children: a systematic review. *Eur J Nutr*. 2017;56:1797-1817. [STEP 1]
2. Weder S, Hoffmann M, Becker K, Alexy U, Keller M. Energy, macronutrient intake, and anthropometrics of vegan, vegetarian, and omnivorous children (1-3 years) in Germany (VeChi Diet Study). *Nutrients*. 2019;11(4):832. [STEP 2]
3. Rudloff S, Buhrer C, Jocham F, et al. Vegetarian diets in childhood and adolescence. *Mol Cell Pediatr*. 2019;6:4. [STEP 5]

Is tinnitus retraining therapy more effective than standard of care treatment for tinnitus?

EVIDENCE-BASED ANSWER

Tinnitus retraining therapy is an effective treatment to reduce tinnitus but is not superior to standard of care for long-term (12 to 18 months) improvement of symptoms (SOR: **A**, consistent results from 2 randomized controlled trials).

DOI 10.1097/EBP.0000000000001440

A 2019 randomized controlled trial (RCT; n=151) compared the efficacy of tinnitus retraining therapy (TRT) versus standard of care or partial TRT in the treatment of tinnitus.[1] Patients were active duty and retired military personnel and their dependents with functionally adequate hearing sensitivity. Patients were randomized to receive TRT, which is tinnitus-specific educational counseling with low-level broadband sound generators, standard of care (patient-centered counseling only), or partial TRT, which is tinnitus-specific educational counseling with placebo sound generators. The participants' response to treatment was measured at three, six, 12, and 18 months using the Tinnitus Questionnaire. The Tinnitus Questionnaire is a 52-item survey, scored on a scale of 0 to 2, measuring tinnitus severity in five subscales: psychological distress, intrusiveness, auditory difficulties, sleep disturbances, and somatic symptoms. The scoring range is 0 to 102 points, with higher scores indicating more severe tinnitus. The scale scores were converted to effect sizes (ESs). A significant improvement was noted at 18 months in symptom scores in all treatment groups; however, no difference between TRT, partial TRT, or standard-of-care therapies (ES –1.3 vs –1.2 vs –1.0, respectively; $P>.05$).

A 2017 RCT (n=39) examined TRT versus standard of care in adults with chronic (1 year or longer) moderate-to-severe tinnitus (per Tinnitus Handicap Inventory score >36).[2] Those who received previous treatment for tinnitus or considered amenable by medical/surgical therapy, complained of concomitant hyperacusis, or depression were excluded. The treatment regimen for both TRT and standard-of-care groups were similar to those summarized above. After removal of one subject because of scheduling, 38 participants were randomized to receive TRT (n=19) or standard of care with placebo generator (n=19), with follow-up at six, 12, and 18 months. The primary outcome was impairment measured using the Tinnitus Handicap Inventory (THI), a 25-question standardized survey determining level of impairment with scores between 0 (denoting no handicap; grade 1) to 100 (catastrophic handicap; grade 5). At six, 12, and 18-month follow-up, both arms demonstrated significant improvement in THI scores compared with baseline. After excluding dropout or loss to follow-up participants, a small but significant difference was noted in THI scores between the TRT treatment when compared with standard of care at 12 months (–28 vs –19, $P<.05$). However, no significant difference was noted between groups at 18 months of follow-up (–29 vs –19, $P=.07$). EBP

Jacquelynn Hatch, DO
Vishal Anand, MD
Alice Treen, MD
St. Joseph Mercy Livingston Family Medicine Residency Program, Brighton, MI

The authors declare no conflicts of interest.

References

1. Scherer RW, Formby C. Effect of tinnitus retraining therapy vs standard of care on tinnitus-related quality of life: a randomized clinical trial. *JAMA Otolaryngology–Head Neck Surg.* 2019;145:597-608. [STEP 2]
2. Bauer CA, Berry JL, Brozoski TJ. The effect of tinnitus retraining therapy on chronic tinnitus: a controlled trial. *Laryngoscope Investig Otolaryngol.* 2017;2:166-177. [STEP 2]

Does prenatal pelvic floor muscle training help to prevent stress urinary incontinence in late pregnancy, postpartum, and interpregnancy period?

EVIDENCE-BASED ANSWER

Prenatal pelvic floor muscle training (PFMT) helps to prevent urinary incontinence (UI) in the late pregnancy (62% lower risk) and early postpartum period (29% lower risk; SOR: **A**, meta-analysis of clinical trials). However, PFMT may not help prevent UI in the interpregnancy period (SOR: **C**, small clinical trial). PFMT is recommended as an initial treatment of all types of UI but not specifically as prevention of stress UI in pregnancy (SOR: **C**, expert opinion).

DOI 10.1097/EBP.0000000000001440

A 2020 meta-analysis of 46 quasi-randomized or randomized controlled trails (RCTs) examined pelvic floor muscle training (PFMT) for prevention or treatment of urinary incontinence (UI; N=10,832).[1] Patients were pregnant women or women who had delivered their baby within the last three months and who were continent at recruitment. The intervention group took part in a PFMT program, which typically included one or more sets of exercises per day, performed on at least several days of the week for at least eight weeks during pregnancy. Control groups were asked not to do PFMT, did not receive instruction on PFMT, or received usual antenatal or postnatal care (eg, information on PFMT) or an alternative PFMT intervention. The primary outcome was self-reported UI of any type. Compared with usual care, women performing antenatal PFMT had a 62% lower risk of reporting UI in late pregnancy (6 trials, N=624; risk ratio [RR] 0.38; 95% CI, 0.20–0.72) and 29% lower risk in the mid-postnatal period (>3–6 months postpartum; 5 trials, N=673; RR 0.71; 95% CI, 0.54–0.95). No difference was noted for the late postnatal or interpregnancy period (>6–12 months; 1 trial, n=44; RR 1.2; 95% CI, 0.65–2.2). Limitations included lack of blinding of participants given the nature of the intervention, varied specific exercises of each PFMT routine, and limited details on group allocation.

The 2015 American College of Obstetricians and Gynecologists and the American Urogynecologic Society published practice guidelines by summarizing scientific evidence regarding diagnosis and management of UI in women of all ages.[2] The practice guidelines on UI listed PFMT as a first-line, noninvasive treatment of UI in women of various ages and etiologies. However, the guidelines made no recommendation on use of PFMT for prevention of UI in pregnancy, postpartum, or interpregnancy interval. EBP

Renee Cherian, DO
Erin Morgan, DO
In His Image Family Medicine Residency, Tulsa, OK

The authors declare no conflicts of interest.

References

1. 'Woodley SJ, Lawrenson P, Boyle R, et al. Pelvic floor muscle training for preventing and treating urinary and faecal incontinence in antenatal and postnatal women. *Cochrane Database Syst Rev*. 2020;5(5):CD007471. [STEP 1]
2. ACOG Practice Bulletin No. 155: urinary incontinence in women. *Obstet Gynecol*. 2015;126(5):e66-e81. [STEP 5]

Do CCBs affect male fertility factors?

EVIDENCE-BASED ANSWER

Hypertension is associated with lower semen volume, sperm motility, total sperm count, and total motile sperm count, whereas treatment with calcium channel blockers, such as verapamil, seems to increase semen volume but have little effect on sperm concentration (SOR: **B**, based on a cross-sectional study). Verapamil may cause a dose-related elevation of prolactin and decreased testosterone, and thus have a possible effect on fertility in a small number of patients (SOR: **C**, based on a cross-sectional study of disease-oriented outcomes).

DOI 10.1097/EBP.0000000000001428

A 2017 cross-sectional study 1 (n=1,076) evaluated the effects of calcium channel blockers (CCBs) on male infertility. The study included male participants of infertile couples who were diagnosed with hypertension (per International Classification of Diseases 9 codes) before or within one year after semen analysis and taking CCBs (formulation and dosing undefined, n=42). These male participants were compared with those without hypertension (n=1,034). Compared with the control group, men with hypertension, when evaluated by absolute semen parameters, were found to have lower semen volume (2.1 vs 3 mL, $P<.001$), sperm motility (41% vs 47%, $P=.008$), total sperm count in ejaculate (104 vs 147 million, $P=.005$), and total motile sperm count (43 vs 69 million, $P=.03$). When stratified by the World Health Organization fifth criteria, hypertensive men had a higher incidence of subfertile semen volume 18% vs 10%, $P=.03$), sperm concentration (19% vs 12%, $P=.02$), and total motile sperm count (26% vs 16%, $P=.01$). Compared with the control group, the CCB group had a significantly higher semen volume (median 3.0 mL, $P<.05$) and statistically nonsignificant decreased sperm concentration. No other semen parameters differed. Study limitations included treatment with multiple medications and no clear exclusion criteria, method of sampling, or clear age range of study group.

A 1996 cross-sectional study 2 (n=615) evaluated the prevalence and degree of hyperprolactinemia associated with verapamil in the clinical setting. The study included adult male participants (mean age 66 years old) with hypertension taking verapamil for at least two months (n=449). Patients with a history of renal failure, hypothyroidism, or use of other drugs raising prolactin levels were excluded. The control group consisted of 166 patients not treated with verapamil. Serum prolactin levels, frequency of elevated prolactin, and total testosterone levels were measured in all patients. A normal mean value of 135 mU/L (range of 31–462 mU/L) was established for serum prolactin levels. Mean prolactin levels for the verapamil group were 207 mU/L compared with 203 mU/L in the control group ($P<.001$). Hyperprolactinemia occurred in 8.5% (38/449) of patients in the verapamil group and 3% in the control group (5/166) ($P=.018$). A subanalysis was performed of those patients with elevated prolactin (n=38) to evaluate the effects of various verapamil doses. The percentage of patients with abnormal prolactin was 3.1%, 3.1%, 9.21%, 6.0%, and 9.4% corresponding to verapamil doses of 0, 1 to 120, 121 to 240, 241 to 360, and 361 to 480 mg/d ($P=.03$) Patient prolactin levels were then obtained at least 30 days from the initial measurement. Of the 24 patients available for follow-up, 15 continued to take verapamil (group 1) and nine discontinued verapamil between days 6 and 120 (group 2). Fourteen patients (93.3%) taking verapamil continued to be hyperprolactinemic (odds ratio>120, $P<.00001$). Prolactin levels of all patients who discontinued verapamil returned to normal from initial measurement (586–277 mU/L, $P=.001$). Mean serum testosterone levels at follow-up were lower in group 1 (6.2 nmol/L) than in group 2 (9.4 nmol/L, $P=.029$). Two patients in each group were excluded because of primary testicular failure as was one patient with a microadenoma. This study is limited by lack of blinding and possibility of observation bias. EBP

Quratulanne H. Jan, MD
Angela Abouassi, MD MPH
Katherine Serrano, MD
Kenneth C. Stupka, MD MPH
Jamal Islam, MD MS
University of Texas Medical Branch-Family Medicine Residency Program, Galveston, TX

The authors declare no conflicts of interest.

References

1. Guo D, Li S, Behr B, Eisenberg ML. Hypertension and male fertility. *World J Mens Health*. 2017;35(2):59-64. [STEP 2]
2. Romeo JH, Dombrowski R, Kwak YS, Fuehrer S, Aron DC. Hyperprolactinemia and verapamil: prevalence and potential association with hypogonadism in men. *Clin Endocrinol (Oxf).* 1996;45(5):571-575. [STEP 2]

Does the consumption of a high flavonoid diet improve cognitive function in adults?

EVIDENCE-BASED ANSWER

Perhaps. Flavonoid consumption may improve cognitive function in children, adults, and older adults (SOR: **C**, low-quality systematic review of small randomized controlled trials [RCTs] and crossover trials). In older adults with mild, self-perceived cognitive decline, it may decrease perceived cognitive inefficiency, but these effects are not maintained if supplementation is discontinued (SOR: **C**, small RCT). Supplementation with cocoa flavanols in elderly adults may dose-dependently improve cognitive function. Duration of benefit is unclear (SOR: **C**, small RCT).

DOI 10.1097/EBP.0000000000001427

A 2017 systematic review of three long-term (12–16 weeks) randomized controlled trials (RCTs; with 72 patients) and four acute (1–6 hours) crossover trials (with 89 patients) evaluated the impact of diet-based anthocyanin consumption (a subclass of dietary flavonoids) on cognitive outcomes in patients with mild-to-moderate cognitive impairment.[1] Long-term interventions included adults 68 years old and older, whereas the acute interventions included children and young adults. In each study, varying anthocyanin-containing beverages (with 55–253 mg) were compared against little to no anthocyanin consumption (0–7.7 mg). In the acute and long-term trials, cognitive outcomes were assessed with commonly used validated tests aimed at measuring short-term learning, episodic memory, long-term memory, working memory, spatial memory, semantic memory attention, higher executive functioning, speed of processing, response inhibition, response interference, verbal learning, and verbal fluency. Numerical results (ie, magnitude of treatment effect) and confidence intervals were not provided. Additionally, given the heterogeneity between studies, a pooled analysis was not feasible. Authors concluded that subjects within the acute trials of anthocyanin supplementation improved their attention and working memory, speed of processing, verbal learning, and response inhibition. Subjects consuming anthocyanin in the long-term trials improved response interference from baseline, verbal learning, verbal fluency, short-term memory, and long-term memory. The studies included in the review were largely heterogeneous and lacked comprehensive randomization, limiting both applicability and quality of the review.

A 2018 double-blind RCT of 94 patients in the U.S. between 62 and 80 years old with mild, self-perceived cognitive decline investigated eicosapentaenoic acid (EPA), docosapentaepenoic acid (DPA), and anthocyanin consumption (in the form of fish oil and blueberries) for improved cognitive performance.[2] Patients had normal scores in several ratings and tests used to screen across multiple cognitive domains for dysfunction in memory, mood, cognition, and verbal learning. Patients were randomly assigned to receive blueberry powder (269 mg anthocyanin daily) and placebo oil (BB), fish oil (400 mg EPA and 200 mg DPA daily) and blueberry powder (BB+FO), fish oil and placebo powder (FO), or placebo oil and placebo powder (PL) for 24 weeks along with a diet restriction of other fish oil and anthocyanin-containing foods. Assessments were taken at the start of the trial, after 24 weeks of treatment, and at 48 weeks, using the Dysexecutive Questionnaire to characterize self-assessed change in cognitive effectiveness in everyday activities. Seventy-six test subjects completed the 24-week treatment period and 65 completed the 48-week postassessment. There were no observed effects in any cognitive domain in the BB+FO group. But, at 48-week follow-up, a reduction in cognitive symptoms was seen in both the FO group (Cohen's f=0.46, P=.02) and the BB group (Cohen's f=0.41, P=.03). (Cohen's f is an effect size measure, where 0.1 is considered small, 0.25 moderate, and 0.4 large.) During the study, 16 individuals dropped out with no causes related to treatment. The study was limited by the absence of a run-in period to provide a period of abstinence before supplementation and by measurement limitations with respect to subjectivity in assessing for modest cognitive enhancement.

A 2015 double-blind RCT of 90 individuals in Italy, who were 61 to 85 years old without cognitive dysfunction, evaluated the effect of flavanol consumption on cognitive performance.[3] Inclusion criteria consisted of body mass index ≤30 kg/m^2, Mini-Mental State Examination (MMSE) >27, and Geriatric Depression Scale <11. Those with clinically significant comorbidities, current smokers, habitual users of antioxidant supplements or chocolate, or use of medications interfering with cognitive function were excluded. Patients were randomly assigned to consume daily drinks containing high (993 mg), intermediate (520 mg), or low (48 mg) cocoa flavanols for eight weeks. Cognitive function of each group was assessed before and after eight weeks with the MMSE, Trail Making Test (TMT) parts A and B, and the Verbal Fluency Test (VFT). No differences in MMSE scores were noted between groups. The high

flavonoid group completed the TMT A faster than the intermediate and low flavonoid groups (–8.6 vs –6.7 and –0.8 seconds, respectively; $P<.001$ for both groups). The high flavonoid group also completed the TMT B faster than the intermediate and low flavonoid groups (–16.5 vs –14.2 and –1.1 seconds, respectively; $P<.001$ for both groups). Likewise, VFT scores improved across each group in a dose-dependent fashion by 7.70 words/min in the high flavonoid group ($P<.0001$), 3.57 words/min in the intermediate flavonoid group ($P<.007$), and 1.33 words/min in the low flavonoid group ($P<.01$). One patient discontinued treatment because of gastric discomfort but otherwise no adverse effects or outcomes were reported. The main limitation of this study was its generalizability to the typical elderly population because this cohort had few comorbidities. Supplement consumption was also lower than typical in the general population because habitual consumers of dietary supplements with antioxidant properties were excluded. EBP

William Leber, MD
David Evans, MD
Keith Petersen, DO
Tyler Raymond, DO, MPH, FAAFP
Madigan Army Medical Center, Tacoma, WA

The authors declare no conflicts of interest.
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 Army Medical Department, the Army at large, or the Department of Defense.

References

1. Kent K, Charlton KE, Netzel M, Fanning K. Food-based anthocyanin intake and cognitive outcomes in human intervention trials: a systematic review. *J Hum Nutr Diet*. 2017; 30(3):260-274. [STEP 1]
2. McNamara RK, Kalt W, Shidler MD, et al. Cognitive response to fish oil, blueberry, and combined supplementation in older adults with subjective cognitive impairment. *Neurobiol Aging*. 2018;64:147-156. [STEP 2]
3. Mastroiacovo D, Kwik-Uribe C, Grassi D, et al. Cocoa flavanol consumption improves cognitive function, blood pressure control, and metabolic profile in elderly subjects: the Cocoa, Cognition, and Aging (CoCoA) Study—a randomized controlled trial. *Am J Clin Nutr*. 2015;101(3): 538-548. [STEP 2]

What are the risks and benefits to breastfeeding infants when mothers are on medication-assisted therapy with methadone and buprenorphine?

EVIDENCE-BASED ANSWER

Breastfed infants of mothers receiving medication-assisted treatment (MAT) were 26% to 32% less likely to need treatment of neonatal abstinence syndrome and stayed an average of 5 to 6 fewer days in the hospital. Infants breastfed by mothers on methadone (but not buprenorphine) had a decreased average duration of treatment when treated for neonatal abstinence syndrome. No risks of breastfeeding on MAT were identified in any of these studies (SOR: **B**; systematic review and retrospective cohort studies).

DOI 10.1097/EBP.0000000000001439

A 2019 systematic review of eight studies (N=779) examined newborn feeding methods (breastfed or formula fed) and outcomes related to neonatal abstinence syndrome in infants with intrautero exposure to medication-assisted treatment (MAT) with methadone, buprenorphine, or both.[1] The eight studies included five retrospective cohort studies, one prospective cohort study, one matched-subject design, and one mixed-methods, randomized, control trial. All studies utilized the Finnegan scoring tool, which is a 21-item scale to qualitatively measure the severity of neonatal withdrawal from opioids, with a maximum score of 37, and a score of ≥8 indicated that the infant needs additional pharmacologic intervention. One retrospective cohort study (n=190) of neonates with in -tero exposure to methadone found the development of neonatal abstinence syndrome to occur significantly later in breastfed infants

than in formula-fed infants (10 vs 3 days; $P=.001$). They also found that breastfed newborns were less likely to require pharmacologic treatment (53% vs 79%; $P<.001$) and had a significantly shorter hospital lengths of stay (15 vs 19 days; $P=.049$). One prospective cohort study (n=124) found that breastfed neonates exposed to methadone in utero were less likely to require pharmacologic treatment (53% vs 80%; $P<.05$) and had a shorter duration of treatment of neonatal abstinence syndrome (31 vs 49 days; $P\leq.05$), but no significant difference in the pharmacologic need for neonatal abstinence syndrome (NAS) treatment or duration of treatment in infants exposed to buprenorphine in utero was found. A retrospective cohort study (n=128) of breastfed infants exposed to methadone in utero found that they had a nonsignificant shorter median duration of methadone pharmacotherapy, but they had a significantly shorter hospital stay compared with formula-fed infants (13 vs 19 days; $P=.01$). Another retrospective cohort review (n=28) looking at newborns exposed to methadone in utero found that mean Finnegan scores were lower in breastfed infants than in formula-fed infants (4.9 vs 6.9; $P=.0001$). A final retrospective cohort study(n=194) showed no statistical difference in mean NAS scores of breastfed versus formula-fed infants but showed that breastfeeding during the first two days of life was associated with a delayed onset of NAS ($P=.04$). None of the eight studies showed any risk of breastfeeding while on MAT. Limitations of the review included the use of primarily chart reviews, and lack of data on the dosing of methadone or buprenorphine.

A 2019 retrospective cohort study (n=228), not included in the systematic review above, studied mothers on MAT (methadone or buprenorphine) and compared prenatal breastfeeding intention with breastfeeding rates at hospital discharge and at two months postpartum.[2] Researchers also looked at whether breastfeeding at hospital discharge reduced the need for pharmacological treatment of neonatal opioid withdrawal syndrome. The study found a decreased need for pharmacological treatment of neonatal abstinence syndrome among neonates who were exclusively breastfed at discharge compared with infants who did not breastfeed at discharge (16% vs 48%; $P<.05$). No risks of breastfeeding on MAT were identified in this study. This study was limited by the inability to randomize patients and the lack of blinding.

A 2018 retrospective cohort study (n=89), not included in the systematic review above, looked at whether maternal buprenorphine dose affected the severity or incidence of neonatal abstinence syndrome, as well as confounding factors such as maternal breastfeeding.[3] Electronic medical record audit included women 18 to 53 years old (and their infants) with a Diagnostic and Statistical Manual of Mental Disorders IV diagnosis of opioid dependency. Neonatal abstinence syndrome was assessed every three hours using the Finnegan scale, with oral morphine administered for neonates with three consecutive Finnegan scores totaling >24. The study found no significant associations between the maternal buprenorphine dose and the infant's peak NAS score, NAS severity requiring morphine, time to morphine start, peak morphine dose, or days on morphine. Exclusively breastfed infants had a decreased need for any morphine compared with nonexclusively breastfed infants (23% vs 55%; $P=.003$). Exclusively breastfed infants also had a shorter hospital stay (8 vs 13 days; $P=.01$). No risks of breastfeeding on MAT were identified in this study. Limitations of this study included questionable reliability of the medical record and a small sample size. EBP

Lisa Gieseke, DO
Sterling Adams, DO
Jennifer Speirs, DO
Matthew Orobona, DO
Katherine Stuart, DO
HealthONE Family Medicine Residency Program, Aurora, CO

The authors declare no conflicts of interest.

References

1. McQueen K, Taylor C, Murphy-Oikonen J. Systematic review of newborn feeding method and outcomes related to neonatal abstinence syndrome. *JOGN Nurs.* 2019;48(4):398-407. [STEP 1]
2. Yonke N, Maston R, Weitzen S. Breastfeeding intention compared with breastfeeding postpartum among women receiving medication-assisted treatment. *J Hum Lact.* 2019; 35(1):71-79. [STEP 3]
3. Wong J, Saver B, Scanlan J, et al. Does maternal buprenorphine dose affect severity or incidence of neonatal abstinence syndrome? *J Addict Med.* 2018;12(6):435-441. [STEP 3]

Does coenzyme Q_{10} enzyme relieve symptoms in patients with heart failure?

EVIDENCE-BASED ANSWER

In patients with heart failure, coenzyme Q_{10} (CoQ10) supplementation reduces mortality and increases exercise capacity but has no benefit on New York Heart Association classification (SOR: **A**, subgroup analysis of 4 and 3 randomized controlled trials [RCTs], respectively). Furthermore, CoQ10 reduces hospitalizations by 51% at two years (SOR: **B**; single RCT).

DOI 10.1097/EBP.0000000000001433

A 2017 meta-analysis of 14 randomized control trials (N=2,149) assessed the efficacy of CoQ10 supplementation in patients with heart failure as compared with placebo for endpoints of death, left heart ejection fraction, exercise capacity, and New York Heart Association (NYHA) classification.[1] The meta-analysis included 2,149 patients with heart failure; left and right heart failure were not differentiated. The clinical endpoints of exercise capacity and NYHA classification (class I no symptoms, class II mild symptoms with activity, class III marked limitations because of symptoms with activity, and class IV symptoms at rest) were pooled using standardized mean difference (SMD). Four RCTs were included in the analysis of exercise capacity with a total of 264 patients, half of whom received CoQ10 and half received placebo. Exercise capacity was measured as walking distance, exercise duration, or both and was not defined further. Exercise capacity was improved with CoQ10 compared with placebo (SMD=0.62; 95% CI, 0.02–1.12; P=.04). Three RCTs were included in the analysis of NYHA classification with a total of 126 patients, 66 of whom received CoQ10 and 60 received placebo. No significant difference was noted in NYHA classification or left heart ejection fraction. However, the CoQ10 group had a decreased rate of all-cause mortality (risk ratio [RR] 0.69; 95% CI, 0.50–0.95; P=.02), with no difference in left heart ejection fraction. Although trials included in the meta-analysis were middle or high-quality studies, limitations included heterogeneity of the design characteristics (different doses and duration of treatment) and some trials lacked detailed descriptions of allocations, concealment, and blinding.

A 2014 two-year prospective randomized control trial (n=420) compared supplementation with 100 mg of CoQ10 three times a day versus placebo, in addition to standard heart failure therapy, and assessed short- and long-term improvements in heart failure outcomes.[2] Although this study was referenced in the meta-analysis above, it was not used in the subgroup analyses considered in the aforementioned paragraph and is discussed separately. Patients were NYHA class II through IV, with 88% in class III. No specified cutoff for ejection fraction was noted. The treatment and placebo groups were similar for standard heart failure treatment, baseline duration of disease, and initial six-minute walk test. The primary short-term endpoints analyzed at 16 weeks were changes in the NYHA classification, six-minute walk test, and N-terminal pro-B type natriuretic peptide levels. The primary long-term endpoint analyzed at 2 years was time to first major adverse cardiovascular event (MACE). No differences in short-term endpoints were noted during the study. However, at two years, 15% of patients taking CoQ10 suffered their first MACE compared with 26% of patients in the placebo group (hazard ratio [HR] 0.50; 95% CI, 0.32–0.80; P=.003). CoQ10 supplementation also improved cardiovascular mortality (9% vs 16%, P=.026), all-cause mortality (10% vs 18%, P=.018), and the incidence of hospital stays for heart failure (HR 0.51; 95% CI, 0.27–0.95; P=.033) at two years. Limitations included not reaching the target number of 550, increasing risk for type 2 (false negative) error. Additionally, NYHA classification improved during the trial on standard therapy alone, which may have confounded the long-term mortality benefit endpoint.

A recent 2021 Cochrane systematic review of 11 RCTs (N=1,573; 7 studies overlapped with the meta-analysis above) assessed the efficacy of CoQ10 compared with placebo or standard heart failure therapy in patients with chronic heart failure for primary endpoints of all-cause mortality, cardiovascular mortality, MI, stroke, PCI and CABG, hospitalization due to heart failure, all-cause hospitalization, and NYHA classification.[3] Many of

the outcomes were only reported by a single study, so meta-analysis could not be completed. All-cause mortality, cardiovascular mortality, MI, and PCI/CABG data were only reported by the previously discussed 2014 RCT. Two studies evaluated risk of hospitalization for heart failure; however, these studies were conducted more than 20 years apart. All-cause hospitalization was not reported by any included study. Seven studies reported data regarding NYHA classification; however, they could not be pooled because of differences in reporting methods. The authors concluded the presence of moderate-quality evidence that CoQ10 reduced all-cause mortality (1 study, n=420; relative risk [RR] 0.58; 95% CI, 0.35–0.95) and hospitalization for heart failure (2 studies, N=1,061; RR 0.62; 95% CI, 0.49–0.78). Prespecified secondary endpoints, including left ventricular ejection fraction and exercise capacity, had very low-quality evidence raising uncertainty regarding the effect of CoQ10. Limitations of this review included relatively small sample sizes, varied dosing of CoQ10, varied periods of follow-up, unclear methods of randomization among the studies, and differing endpoints or reporting methods resulting in inability to perform meta-analysis. EBP

Molly Chandler, MD
Marco Jose Dela Cruz, MD
Gary Clark, MD
Christopher Taylor, MD
Madigan Army Medical Center, Tacoma, WA

The authors declare no conflicts of interest.
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 Army Medical Department, the Army at large, or the Department of Defense.

References

1. Lei L, Liu Y. Efficacy of coenzyme Q10 in patients with cardiac failure: a meta-analysis of clinical trials. *BMC Cardiovasc Disord.*. 2017;17(1):1-7. [STEP 1]
2. Mortensen SA, Rosenfeldt F, Kumar A, et al. The effect of coenzyme Q10 on morbidity and mortality in chronic heart failure: results from Q-SYMBIO: a randomized double-blind trial. *JACC Heart Fail.* 2014;2(6):641-649. [STEP 2]
3. Al Saadi T, Assaf Y, Farwati M, et al. Coenzyme Q10 for heart failure. *Cochrane Database Syst Rev.* 2021(2):Art. No.: CD008684. [STEP 1]

In the initial management of adults with acute respiratory failure, does high-flow nasal oxygen lead to better outcomes than noninvasive ventilation?

EVIDENCE-BASED ANSWER

In hospitalized adults with either hypoxic or hypercapnic acute respiratory failure (ARF), it is unclear if high-flow nasal oxygen (HFNO) changes mortality rates or the need for intubation compared with noninvasive ventilation (NIV) (no SOR given, conflicting randomized controlled trials [RCTs]). However, in hypercapnic ARF alone, intubation rates and mortality appear similar with either approach (SOR: **B**, meta-analysis of RCTs and cohort studies). HFNO may be more comfortable for patients than NIV (SOR: **C**, inconsistent evidence from meta-analyses of RCTs).

DOI 10.1097/EBP.0000000000001442

A 2021 meta-analysis of 29 randomized controlled trials (RCTs; N=4,968) examined the effectiveness of high-flow nasal oxygen (HFNO) versus noninvasive ventilation (NIV) or conventional oxygen therapy in hospitalized adults with acute respiratory failure (ARF), of which eight studies (N=1,372) compared HFNO with NIV in the setting of initial management of ARF as opposed to the setting of postextubation management of ARF.[1] Patients had an average age of 63 years old and were treated in the emergency department for hypercapnic respiratory failure (one study, n=204), intensive care unit (ICU) for hypoxic respiratory failure (four studies, N=1,169), and hospital/ward/step-down unit for hypoxic or hypercapnic respiratory failure (three studies, N=234). Underlying diagnoses included pneumonia, heart

failure, chronic obstructive pulmonary disease, and cystic fibrosis. ARF was defined as oxygen saturations (SpO_2) less than 90%, PaO_2-FiO_2 ratio less than 300, PaO_2 <60 mmHg, or $PaCO_2$ >45 mmHg. The intervention groups received HFNO at a rate greater than 20 L/min, whereas the comparison group received NIV with inspiratory pressures 5 to 10 cm H_2O and expiratory pressures 5 to 10 cm H_2O titrated to a goal SpO_2 of 92% to 95%. HFNO compared with NIV in the initial management of ARF reduced all-cause mortality (one study, n=216; relative risk [RR] 0.46; 95% CI, 0.24–0.79) and reduced the need for intubation (two studies, N=420; RR 0.71; 95% CI, 0.53–0.95), No difference in ICU admissions, length of stay, and incidence of hospital acquired pneumonia was observed. Three studies reported improved patient comfort with HFNO compared with NIV, whereas four studies reported no differences, but results were not pooled due to heterogeneity in outcome measures. Adverse events, such as facial pressure, sores, or nasal ulcerations, were not reported. Two of the studies were rated as low risk of bias, whereas six had moderate bias (deficiencies not reported).

A 2021 meta-analysis of 31 RCTs (N=5,136) compared HFNO with conventional oxygen therapy or NIV for the treatment of ARF in adult patients treated exclusively in the ICU.[2] Seven studies compared HFNO with NIV in the initial management of ARF; however, only two of these studies (N=383) evaluated the primary outcomes of treatment failure (need for intubation or progression of respiratory support) and in-hospital mortality. One of these studies was included in the review above; it is unclear why the other study was not. Patients were on average 61 years old and treated in the ICU for ARF (PaO_2/FiO_2 <300 mmHg). Intervention and comparison treatments were similar to the studies in the review above. Subgroup analysis only on studies of the initial management of ARF showed no difference in treatment failure between HFNO and NIV (2 studies, N=286; RR 0.77; 95% CI, 0.58–1.03). Subgroup analysis was not done for mortality, but as summarized in the review above, one of these studies showed decreased in-hospital mortality with HFNO; however, the study unique to this review showed no difference between HFNO and NIV (RR 0.87; 95% CI, 0.41–1.8). When pooled with three studies in the setting of postextubation ARF, no difference was found between HFNO and NIV for in-hospital mortality (5 studies, N=1,758; RR 0.92; 95% CI, 0.64–1.3). Adverse events were not evaluated in the initial management studies. The two RCTs were determined to be low risk of bias but had unclear allocation concealment and one had possible selective reporting of outcomes.

A 2020 meta-analysis of six RCT and two cohort studies (N=621) examined whether HFNO is more effective than NIV at reducing the mortality and intubation rate in patients with hypercapnic ARF.[3] Three RCTs (N=200) focused on the initial management, whereas three other RCTs (n=251) included postextubation patients. Only one study (n=168) was included in the first review above, and no studies were there in common with the second review above. Patients were on average 72 years old and had acute hypercapnic respiratory failure defined by a $PaCO_2$ more than 45 mmHg. The intervention group received HFNO (starting at 35–50 L/min and titrated via SpO_2 or blood gas analysis) and the control group received NIV (initial inspiratory pressures of 10 cm H_2O and expiratory pressures of 4–5 cm H_2O titrated via SpO_2 or blood gas measurements). Pooled analysis of only the RCTs found no statistical difference between the HFNO and NIV groups in mortality at one month (5 studies, N=317; RR 1.3; 95% CI, 0.68–2.6) or in need for intubation (5 studies, N=283; RR 0.92; 95% CI, 0.45–1.9). Subgroup analysis of the RCTs evaluating the initial management of ARF was not done. Comfort and adverse event data was not pooled and the scales used were not reported, but all RCTs evaluating each outcome were reported to be consistent in showing that HFNO was more comfortable (3 studies, N=296), caused less flatulence (2 studies, N=115), and resulted in less nasofacial skin breakdown (two studies, N=129). The reviewers noted that three of the RCTs had high risk of bias due to the lack of description of randomization and allocation concealment; the other three RCTs had low risk of bias. EBP

Christopher Mullet, DO
Thomas Satre, MD
University of Minnesota/CentraCare St Cloud Family Medicine Residency, St Cloud, MN

The authors declare no conflicts of interest.

References

1. Baldomero AK, Melzer AC, Greer N, et al. Effectiveness and harms of high-flow nasal oxygen for acute respiratory failure: an evidence report for a clinical guideline from the American college of Physicians. *Ann Intern Med*. 2021;174(7):952-966. [STEP 1]
2. Lewis SR, Baker PE, Parker R, Smith AF. High-flow nasal cannulae for respiratory support in adult intensive care patients. *Cochrane Database Syst Rev*. 2021(3):CD010172. [STEP 1]
3. Huang Y, Lei W, Zhang W, Huang JA. High-flow nasal cannula in hypercapnic respiratory failure: a systematic review and meta-analysis. *Can Respir J*. 2020;2020:7406457.[STEP 2]

EVIDENCE-BASED ANSWER

Multiple risk factors are associated with post-traumatic stress disorder (PTSD) for both adult and child survivors of earthquakes. Higher education is the only protective factor that has been identified (SOR: **B**, meta-analysis of observational studies). For adult hurricane survivors losing a loved one, being present for the storm and seeking health care during the storm are associated with development of PTSD (SOR: **C**, cross-sectional study).

DOI 10.1097/EBP.0000000000001466

What risk factors predispose individuals to post-traumatic stress disorder after an environmental disaster?

A meta-analysis of 52 observational studies (N=56,722 adults and 22,931 children) evaluated risk factors for post-traumatic stress disorder (PTSD) after earthquakes.[1] Fifty-one cross-sectional studies and one cohort study were included. Total sample size ranged from 91 to 14,207; time after disaster ranged from three months to 60 months. Fifteen studies reviewed children and the remaining 37 examined adults. Various methods of evaluating PTSD were used, with the most common being the PTSD checklist (14 studies), which is a 14-item questionnaire. Studies came from nine different countries with 31 from China. The meta-

TABLE 1. Adult risk factors of PTSD[1]

Characteristics	No. of studies	Odds ratio (95% CI)	Heterogeneity I^2 (P^a)
Experience of fear	4	3.0 (1.8–5.0)	92.9% (<.001)
Bereavement	23	2.5 (2.0–3.0)	70.8% (<.001)
Loss of employment	14	2.1 (1.5–2.9)	85.5% (<.001)
Personal injury	14	2.1 (1.3–3.2)	90.0% (<.001)
House damage	21	1.9 (1.5–2.3)	Not cited
Female gender	42	1.9 (1.7–2.0)	51.1% (<.001)
Being trapped	6	1.8 (1.5–2.2)	Not cited
Low socioeconomic status	16	1.7 (1.2–2.5)	89.0% (<.001)
Loss of property	11	1.7 (1.3–2.2)	Not cited
Prior trauma	2	1.6 (1.1–2.4)	55.3% (=.135)
Older age	32	1.2 (1.1–1.3)	73.2% (<.001)
Higher education	30	0.8 (0.8–0.9)	33.3% (=.041)

[a] Low *P* scores (<.05) indicate that heterogeneity is statistically significant.

TABLE 2. Pediatric risk factors of PTSD[1]

Characteristics	No. of studies	Odds ratio (95% CI)	Heterogeneity I^2 (P^a)
Fear	2	2.2 (1.5–3.3)	Not cited
Bereavement	18	2.2 (2.0–2.6)	Not cited
Personal injury	11	2.1 (1.7–2.5)	Not cited
Witnessing injury	6	2.0 (1.4–2.8)	64.9% (=.014)
Being trapped	9	1.9 (1.5–2.5)	Not cited
Loss of property	11	1.8 (1.5–2.0)	Not cited
Higher education	13	1.6 (1.1–2.2)	77.7% (<.001)
Female gender	20	1.5 (1.3–1.6)	Not cited
Age >7 yr old	7	1.3 (1.1–1.6)	Not cited

[a] Low *P* scores (<.05) indicate that heterogeneity is statistically significant.

analysis reviewed pre-trauma characteristics, trauma characteristics, and post-trauma characteristics to determine risk factors for PTSD. Risk factors of PTSD for adults included experience of fear, bereavement, loss of employment, personal injury, house damage, female gender, being trapped, low socioeconomic status, loss of property, prior trauma, and older age (not defined by authors) (see **TABLE 1**). A high-level of education (not defined) was negatively associated, whereas marital status, ethnicity, religion, disease history, witnessing injury, and being involved in rescue were not correlated. For children, risk factors included fear, bereavement, personal injury, witnessing injury, being trapped, loss of property, higher education, female gender, and age >7 years old (see **TABLE 2**). Religion, ethnicity, prior trauma, presence of social support, and house damage were not associated. Limitations of this study included its reliance on solely observational studies; data based largely on self-reporting measures; significant heterogeneity among studies in sampling, design, measurement, and statistical analysis; publication bias; publications limited to English language; and many variables examined by only a small proportion of studies.

A cross-sectional study in 2011 (n=747) evaluated risk factors for PTSD one year after Hurricane Katrina by surveying patients 18 years old or older present in the only operating New Orleans ER 11 to 12 months after the hurricane.[2] Inclusion criteria consisted of those who gave verbal consent to answer questions and were residents of greater New Orleans at the time of the storm. Interviews consisted of a structured questionnaire and included a four-question Primary Care PTSD screen, where two or more positive responses indicated a positive screen. Risk factors associated with positive PTSD screener were the death of a loved one due to the disaster (odds ratio [OR] 1.9; 95% CI, 1.3–2.8), being in New Orleans during the storm (OR 1.7; 95% CI, 1.2–2.3), and seeking health care during the storm (OR 1.7; 95% CI, 1.2–2.4). No significant correlation was noted for female gender, material losses, or seeking health care after returning to the city after evacuation. Limitations of this study were that surveys were conducted in an Emergency Department waiting room and required participation from otherwise sick patients requiring emergency room (ER) care, and concurrent psychiatric disorders were not examined as contributing factors to PTSD. EBP

Jared Baird, DO
Jennifer Svarverud, DO
Nellis Family Medicine Residency, Las Vegas, NV

The authors declare no conflict of interest.
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 Uniformed Services University of the Health Sciences, the U.S. Air Force Medical Department, the Air Force at large or the Department of Defense.

References

1. Tang B, Deng Q, Glik D, et al. A meta-analysis of risk factors for post-traumatic stress disorder (PTSD) in adults and children after earthquakes. *Int J Environ Res Public Health.* 2017;14(12):1537. [STEP 1]
2. Mills LD, Mills TJ, Macht M, et al. Post-traumatic stress disorder in an emergency department population one year after Hurricane Katrina. *J Emerg Med.* 2012;43(1):76-82. [STEP 3]

Does melatonin prevent procedure-related anxiety in adults and children?

EVIDENCE-BASED ANSWER

In both adults and pediatric patients, melatonin reduces anxiety scores in the perioperative setting compared with placebo (SOR: **B**, meta-analysis with low-quality randomized controlled trials [RCTs] and small RCTs).

DOI 10.1097/EBP.0000000000001443

A 2014 systematic review of 24 randomized controlled trials (RCTs; N=1,794) analyzed melatonin's effect on various perioperative outcomes, including anxiety.[1] The RCTs had patient populations ranging from children to adults, although specific age ranges and gender were not specified. All studies used between 3 and 10 mg of melatonin, either oral or sublingual, and medication administration ranged between the night before surgery to 30 minutes before surgery. Most control groups were given an inactive placebo, but a few trials used midazolam or clonidine, and one trial used dexmedetomidine. Fourteen RCTs (N=1,146) investigated perioperative anxiety. These studies used four different validated scales for measurement of anxiety, so results were pooled using standardized mean difference (SMD). Anxiety was measured before, during, and after surgery. In every individual RCT, melatonin either reduced intraoperative and postoperative anxiety or was no different from control. With pooled results, melatonin reduced the preoperative anxiety score compared with control (SMD –0.88; 95% CI, –1.33 to –0.44). Melatonin was not associated with any serious side effects, but several trials registered sedation as an adverse effect. The authors noted "extreme heterogeneity" among studies within the meta-analysis.

A 2018 RCT (N=90) examined the effectiveness of a preoperative one-time dose of melatonin, gabapentin, or placebo for the treatment of postoperative anxiety related to lumbar spine surgery in adult Iranian patients 18 to 60 years old.[2] Patients were excluded if they had an anxiety disorder, their body mass index was greater than 35 kg/m^2, if they were allergic to melatonin or gabapentin, or if it was an emergency surgery. Patients received 6 mg melatonin, 600 mg gabapentin, or placebo 100 minutes before surgery. Patients were evaluated 15 minutes before surgery and at 1, 2, 6, 12, and 24 hours postoperatively using verbal anxiety score (from 0: no anxiety to 10: maximum anxiety). Study authors noted a decrease in verbal anxiety score 15 minutes before surgery in patients given melatonin or gabapentin compared with control (3.8 for melatonin vs 4.0 for gabapentin and 5.5 for control, $P=.04$). Patients given melatonin or gabapentin had significantly lower anxiety scores at all periods after surgery compared with control, including 24 hours postoperatively (1.4 for melatonin, 1.2 for gabapentin, and 2.6 for control, $P=.03$). No adverse effects were noted.

A 2016 RCT (n=100) evaluated the preoperative anxiolytic effects of melatonin versus midazolam or placebo for Indian children 5 to 15 years old.[3] Each child received one of the following in syrup form: melatonin 0.5 mg/kg, melatonin 0.75 mg/kg, midazolam 0.5 mg/kg, or placebo. Patient anxiety was assessed via a modified Yale Preoperative Anxiety Scale (range 0–60, with 60 being maximum anxiety). Researchers assessed anxiety before medication administration and at 30 and 60 minutes after administration of medication or placebo. Patients receiving placebo had higher anxiety scores than those receiving melatonin 0.5 mg/kg, melatonin 0.75 mg/kg, or midazolam at 30 minutes (50 vs 37 vs 37 vs 35, $P=.00001$) and at 60 minutes (41 vs 28 vs 24 vs 29, $P=.0001$). When compared with one another, patients receiving melatonin 0.5 mg/kg and melatonin and 0.75 mg/kg had comparable anxiety scores ($P=.17$). EBP

Adrienne Ide, DO
Kindra Galloway, DO
Michael Jones, DO
Biana Dorfman, DO
Kyle Adkins, DO
HealthONE Family Medicine Residency Program, Aurora, CO

The authors declare no conflicts of interest.

References

1. Andersen LP, Werner MU, Rosenberg J, Gögenur I. A systematic review of peri-operative melatonin. *Anaesthesia*. 2014;69(10):1163-1171. [STEP 2]
2. Javaherforooshzadeh F, Amirpour I, Janatmakan F, Soltanzadeh M. Comparison of effects of melatonin and gabapentin on post operative anxiety and pain in lumbar spine surgery: a randomized clinical trial. *Anesth Pain Med.* 2018; 8(3):e68763. [STEP 2]
3. Kurdi MS, Muthukalai SP. A comparison of the effect of two doses of oral melatonin with oral midazolam and placebo on pre-operative anxiety, cognition and psychomotor function in children: a randomised double-blind study. *Indian J Anaesth.* 2016;60(10):744-750. [STEP 2]