June 2021

EVIDENCE-BASED PRACTICE

A Peer-Reviewed Journal of the Family Physicians Inquiries Network

June 2021
Volume 24 | Number 6

EVIDENCE-BASED PRACTICE
Volume 24 | Number 6

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 24 | NO. 6 | June 2021

EVIDENCE-BASED PRACTICE

A PEER-REVIEWED JOURNAL OF THE FAMILY PHYSICIANS INQUIRIES NETWORK

INDEPTH

DIVING FOR PURLs

GEMs

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*

SPOTLIGHT ON PHARMACY

EVIDENCE-BASED PRACTICE

A PEER-REVIEWED JOURNAL OF THE FAMILY PHYSICIANS INQUIRIES NETWORK

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 screening for social determinants of health improve patient outcomes?

EVIDENCE-BASED ANSWER

Probably. Screening for social determinants of health likely results in positive impacts and results in referrals for services to help with unmet social needs (SOR: **B**, systematic review of randomized controlled trial and cohort studies). One current practice guideline in this area includes screening for intimate partner violence (IPV) in women of reproductive age, which reduces IPV exposure, reduces depression scores, and improves birth outcomes (SOR: **B**, evidence-based guideline).

DOI 10.1097/EBP.0000000000000869

A 2017 systematic review examined the existing literature on screening for social determinants of health (SDOH) in clinical settings.[1] The study reviewed 67 articles with various study designs (24 randomized controlled trials [RCTs]) covering 37 unique programs that screened patients (total N not provided) for a number of social and/or economic determinants of health and then linked patients to social services. The studies covered a variety of patient populations defined by demographic conditions, disease states, and/or specific social circumstances. They also targeted a variety of SDOH, including housing, employment, education, economic security, personal safety, childcare, food security, and legal issues. Study outcomes included process measures (69%), SDOH outcomes (48%), health outcomes (30%), health care cost impacts (27%), and provider outcomes (13%). The health outcomes measured varied and included disease-specific metrics such as asthma control scores to more generalized outcomes such as mortality, quality of life, and health-related behaviors.

This review revealed screening tools effectively identified unmet social needs and provided referrals, and most trials reported positive impacts on SDOH. However, the association between positively impacting SDOH and improving health outcomes were less clear. Of the 30% reviewed studies that assessed health outcomes, results found more positive than negative impacts of SDOH screening and intervention. Data were not pooled because of different outcomes being measured and heterogeneity of the studies. The authors noted that less than one quarter of the studies in this review met criteria for high quality based on GRADE[2] standards. The study was also limited by the exclusion of articles that included medical interventions in combination with social services.

Current practice guidelines in the area of screening for health determinants include the United States Preventive Services Task Force (USPSTF) recommendation for intimate partner violence (IPV) screening in women of reproductive age (grade B, high certainty that the net benefit is moderate or there is moderate certainty that the net benefit is moderate to substantial). This grade was based on a 2018 systematic review of 30 studies (15 RCTs and 15 cross-sectional trials) investigated the efficacy of screening for IPV.[3] Eleven of the 30 trials (all RCTs, n=6,740) evaluated whether screening women for IPV and performing an intervention improved health outcomes. Interventions included home visits, brief clinic-based counseling, behavioral counseling, and interpersonal psychotherapy. Because of the heterogeneity of outcome reporting, data were not pooled. Screening and intervention reduced IPV exposure in two of 10 studies, reduced depression in three of five studies, and improved birth outcomes in one study. No difference was found in the quality of life (three studies), anxiety (one study), and PTSD (one study). EBP

Riley Smith, MD
Cleveland Piggott, MD
University of Colorado FMR, Denver, CO

The authors declare no conflicts of interest.

References

1. Gottlieb LM, Wing H, Adler NE. A systematic review of interventions on patients' social and economic needs. *Am J Prev Med*. 2017; 53(5):719–729. [STEP 3]
2. Guyatt GH, Oxman AD, Kunz R, Vist GE, Falck-Ytter Y, Schünemann HJ. What is "quality of evidence" and why is it important to clinicians? *BMJ*. 2008; 336(7651):995–998. [STEP 5]
3. Feltner C, Wallace I, Berkman N, et al. Screening for intimate partner violence, elder abuse, and abuse of vulnerable adults. *JAMA*. 2018; 320(16):1688–1701. [STEP 1]

Pain, pain go away. What dose of ibuprofen should I use today?

Motov S, Masoudi A, Drapkin J, et al. Comparison of oral ibuprofen at three single-dose regimens for treating acute pain in the emergency department: a randomized controlled trial. *Ann Emerg Med*. 2019; 74(4): 530–537.

DOI 10.1097/EBP.0000000000001132

This single-center, randomized, double-blind trial of 225 patients compared three doses of oral ibuprofen (400, 600, and 800 mg) and resultant pain scores at 60 minutes. Patients were at least 18 years old and presented to the emergency department (ED) with acute pain. Patients were excluded if NSAIDs were contraindicated for various reasons or if they had received pain medications within four hours before ED arrival. To assist with blinding, ibuprofen was administered orally in a liquid formulation compounded into one of the three doses by a pharmacist onsite. Patients provided pain scores on a standard 0 to 10 numeric scale, adverse effects, and other pain medications administered at baseline and 60 minutes. The primary outcome was difference in mean pain scores across the three groups at 60 minutes; 1.3 points was deemed to represent a clinically meaningful difference. Secondary outcomes included differences in pain within treatment groups from baseline to 60 minutes, rates of adverse effects, and use of other pain medications. Seventy-five patients were randomized to each group, and groups were not different at baseline. The most common indication was musculoskeletal pain (at least 55% in each group) followed by cutaneous pain. At 60 minutes, patients who received 400 mg of ibuprofen had a change in pain score from 6.48 to 4.36 (mean difference, 2.12; 95% CI, 1–4). Patients in the 600-mg group had a change of 6.35 to 4.50 (mean difference, 1.85; 95% CI, 1–3) and patients in the 800-mg group had a change of 6.46 to 4.50 (mean difference, 1.95; 95% CI, 1–4). No clinically meaningful differences were found, and pain reduction was similar between the groups. Four patients in the 400-mg and 800-mg group required an additional pain medication, whereas only one patient in the 600-mg group did. Adverse effects were not different between the groups.

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

Bottom line: Ibuprofen doses of 400, 600, or 800 mg provide similar pain response without increases in adverse events at 60 minutes. In the absence of any meaningful differences and short patient follow-up, providers could use lower ibuprofen doses for analgesia in the acute setting. A longer study is needed to adequately assess a practice change.

Linda Hogan, PhD
Gregory Castelli, PharmD, BCPS, BC-ADM
UPMC St. Margaret FMRP, Pittsburgh, PA

The authors declare no conflicts of interest.

Testosterone and VTE: Worthy of discussion, but not yet practice changing

Walker RF, Zakai NA, MacLehose RF, Cowan LT, Adam TJ, Alonso A, Lutsey PL. Association of testosterone therapy with risk of venous thromboembolism among men with and without hypogonadism. *JAMA Intern Med*. 2019; 180(2):190–7.

DOI 10.1097/EBP.0000000000001172

This retrospective, case-crossover study used large claims database data to evaluate whether short-term testosterone therapy increases short-term venous thromboembolism (VTE) risk in men with and without hypogonadism. The International Business Machine (IBM) MarketScan Commercial Claims and Encounters database was used to identify 39,622 men without cancer who had a VTE between January 1, 2011 and December 31, 2017 and at least 12 months of continuous enrollment before the event. These men in the case period, defined as six months before their VTE, were matched with themselves during the control period, defined as six to 12 months before their VTE. Billed testosterone prescriptions were compared between the

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

periods, controlling for hospitalizations, outpatient visits, and ED visits. Exposure to testosterone was associated with increased VTE risk in both men with (odds ratio [OR] 2.3; 95% CI, 2.0–2.7) and without (OR 2.0; 95% CI, 1.5–2.8) hypogonadism. Subanalyses examining differences by age and testosterone route found no statistically significance difference in VTE. Limitations of this study included the retrospective nature of the trial and lack of control for confounders or control for other VTE risk factors.

Methods

This article was identified as a potential priority updates of the relevant literature (PURL) through the standard systematic methodology that has been described here.[1]

Bottom line: Although this study identifies an important potential risk of testosterone therapy, the case-crossover design and inadequate control for confounding factors does not provide compelling evidence of a causal relationship between short-term testosterone use and VTE. Prescription of testosterone should always involve shared decision making and counseling around risks and benefits. Results of this study may change provider counseling practices to include a discussion of VTE risk; however, ultimately, this study is more hypothesis generating than practice changing.

Haley Ringwood, MD, MPH
Roxanne Radi, MD, MPH
University of Colorado FMR, Denver, CO

The authors declare no conflicts of interest.

More expensive isn't always better: NPH versus insulin analogs as initial long-acting insulin

Lipska KJ, Parker MM, Moffet HH, Huang ES, Karter AJ. Association of initiation of basal insulin analogs versus neutral protamine hagedorn insulin with hypoglycemia-related emergency department visits or hospital admissions and with glycemic control in patients with type 2 diabetes. *JAMA*. 2018; 320(1):53–62.

DOI 10.1097/EBP.0000000000001140

This retrospective observational study examined the time to a hypoglycemia-related emergency department (ED) visit or hospital admission and the change in HbA1c within 1 year of the initiation of insulin analog (glargine or detemir) or Neutral Protamine Hagedorn (NPH) was initiated. The patients all enrolled in the Kaiser Permanente of Northern California. This study included adults over 18 years old with type II diabetes and a full health plan and prescription coverage who were started on basal insulin therapy with either NPH or insulin analog between 2006 and 2014. The primary outcome was time to hypoglycemia-related ED visit or hospital admission. A secondary outcome was the change in HbA1c. Per 1,000 person-years, 11.9 events (95% CI, 8.1–15.6) were noted in the insulin analog group and 8.8 events (95% CI, 7.9–9.8) in the NPH group with a between-group difference of 3.1 events (95% CI, –1.5 to 7.7; $P=.07$). The secondary outcome was change in HbA1c from baseline within 1 year of initiation of therapy. HbA1c decreased 1.26 percentage points (95% CI, 1.16–1.36) for those patients on insulin analogs and 1.48 percentage points (95% CI, 1.39–1.57) for patients on NPH. A difference-in-difference for glycemic control was –0.22% (95% CI,

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

–0.09% to –0.37%), which is statistically, but likely not clinically, significant.

Bottom line: Choosing NPH as the initial long-acting insulin is a reasonable option that is cost saving and will not significantly alter hypoglycemia episodes.

Ben Arthur, MD, FAAFP
Ashley Smith, MD
Nick Bennett, DO, FAAFP
David Bury, DO, FAAFP
FMR at Fort Lewis Madigan Army
Gig Harbor, 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.

Efficacy of anti-inflammatory medications in treating major depression

Bai S, Guo W, Feng Y, et al Efficacy and safety of anti-inflammatory agents for the treatment of major depressive disorder: a systematic review and meta-analysis of randomised controlled trials. J Neurol Neurosurg Psychiatry. 2020; 91(1):21–32.

DOI 10.1097/EBP.0000000000001093

Methods

This article was identified as a potential practice update from the research literature (PURL) through the standard systematic methodology that has been described here. An additional literature search was conducted by searching UpToDate, DynaMed, USPSTF, and PubMed with the terms screening, low dose computed tomography, lung cancer to find additional literature to place this research into the context of current clinical practice.

This systematic review and meta-analysis pooled data from 26 randomized control trials (RCTs) (N=1,610) published through the end of 2018. The study objective was to determine the efficacy of anti-inflammatory medications in the treatment of major depression disorder (MDD) based on high-level evidence. The primary outcome was the efficacy of anti-inflammatory medications in treating depression. This was defined as the mean change in depression scores from the baseline using validated depression scales. Secondary outcomes included response and remission rates as well as the quality of life (QoL) indicators. Safety and adverse events incidence were ascertained. Inclusion criteria specified RCTs, using a placebo arm, that studied anti-inflammatory monotherapy or adjunct therapy and used validated, standardized scales for depression and QoL. Studies were excluded if they were unpublished abstracts, patients had bipolar disease, or the RCTs did not include a placebo arm. Diagnosis was made using Diagnostic and Statistical Manual (DSM) IV or V criteria for MDD. Response to treatment was defined as a 50% reduction in depression score from baseline. The authors used the Cochrane Systematic Review guidelines, focusing on allocation concealment and randomization to assess bias risk. Anti-inflammatory agents moderately reduced depressive symptoms (standardized mean difference [SMD] –0.55, 95% CI –0.75 to –0.35, I^2=71%). The group receiving anti-inflammatory medications also demonstrated higher response (relative risk [RR] 1.52; 95% CI, 1.30–1.79; I^2=29%) and remission rates (RR1.79; 95% CI, 1.29–2.49; I^2=41%). The greatest difference was seen in studies using anti-inflammatory agents as adjunctive therapy (SMD –0.70; 95% CI, –0.97 to –0.43, I^2=74%). Both modafinil and N-acetylcysteine showed nonsignificant effects on MDD. QoL differences between groups were not significant. Gastrointestinal events (higher with anti-inflammatories, especially N-acetylcysteine) constituted the only significant side effect difference between groups. There was moderate to high heterogeneity between studies with I^2 values>40% for most variables (an I^2 value≥50% indicates high heterogeneity).

Bottom line: There is limited evidence that anti-inflammatory medications may be effective in treating major depressive disorder. Due to moderate to high study heterogeneity and inadequate time for adequate adverse effects evaluation, we do not currently recommend a change in practice.

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

Tyler Rogers, MD
Erica Meisenheimer, MD
Samuel Tiglao, DO
Madigan Army Medical Center

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.

A Novel Treatment of COVID-19

Remdesivir for the Treatment of COVID-19—A Preliminary Report

Beigel JH, Tomashek KM, Dodd LE, et al. Remdesivir for the Treatment of COVID-19—Preliminary Report [published online ahead of print, 2020 May 22]. *N Engl J Med*. 2020;NEJMoa2007764. doi:10.1056/NEJMoa2007764

DOI 10.1097/EBP.0000000000001162

KEY TAKEAWAY: Compared with placebo, remdesivir shortened time to recovery in hospitalized patients with coronavirus disease 2019 (COVID-19) with evidence of lower respiratory tract infection requiring supplemental oxygen.

STUDY DESIGN: Double-blind randomized controlled trial, multisite, multinational.

LEVEL OF EVIDENCE: Step 2.

BRIEF BACKGROUND INFO: At the time publication, no known effective treatment for COVID-19 was found. Remdesivir is an antiviral therapy with potential to treat COVID-19 patients.

PATIENTS: Adult hospitalized patients with COVID-19.

INTERVENTION: IV Remdesivir.

CONTROL: Placebo.

OUTCOME: Time to recovery.

METHODS BRIEF DESCRIPTION:

- Hospitalized COVID-19 patients, 18 years old or older, were recruited from 60 sites across 10 countries, including 45 sites in the United States.
- Patients were randomized to an experimental group that received remdesivir at a 200-mg loading dose and a 100-mg daily maintenance dose for total of 10 days, or a control group that received placebo and supportive care per hospital protocol.
- Inclusion criteria:
  - Laboratory-confirmed COVID-19
  - At least one the following: infiltrates on radiography, SpO_2 ≤94% on room air, requiring supplemental oxygen, requiring mechanical ventilation
  - Agreeing not to participate in another COVID-19 treatment clinical trial through day 29
  - Practicing abstinence or study-specified contraception through day 29 for women of childbearing potential.
- Exclusion criteria:
  - ALT or AST greater than five times the upper limit of normal;
  - eGFR less than 30 mL/min
  - Allergy to study product
  - Pregnancy or breastfeeding
  - Anticipated discharge or transfer from the hospital within 72 hours of enrollment.
- Groups were stratified by disease severity into eight categories on an ordinal scale, with higher numbers representing more severe illness.
- Patients were followed for up to 29 days and their clinical status was recorded each day on ordinal scale of 1–8:
  - 1=not hospitalized, no limitations of activities
  - 2=not hospitalized, limitation of activities, home oxygen requirement, or both
  - 3=hospitalized, not requiring supplemental oxygen and no longer requiring ongoing medical care (used if hospitalization was extended for infection-control reasons)
  - 4=hospitalized, not requiring supplemental oxygen but requiring ongoing medical care (COVID-19—related or other medical conditions)
  - 5=hospitalized, requiring any supplemental oxygen
  - 6=hospitalized, requiring noninvasive ventilation or use of high-flow oxygen devices
  - 7=hospitalized, receiving invasive mechanical ventilation or extracorporeal membrane oxygenation
  - 8=death.
- Primary outcome was defined as time to recovery as a 1, 2, or 3 on the ordinal clinical severity scale.
- Primary means of statistical analysis was the stratified log-rank test of recovery time for remdesivir compared with placebo.
- Analysis was by intention to treat.

INTERVENTION (# IN THE GROUP): 538

COMPARISON (# IN THE GROUP): 521

FOLLOW-UP PERIOD: 29 days

RESULTS:

- Primary outcomes:
  - Patients receiving remdesivir had a shorter median recovery time (median, 11 days [95% CI, 9–12 days], as compared with 15 days [95% CI, 13–19 days])
  - Recovery rate ratios (remdesivir vs placebo) for each category of disease severity at baseline, >1 indicates benefit for remdesivir
    - No improvement for hospitalized patients not requiring supplemental oxygen
      - 1.38 (95% CI, 0.94–2.03)

    - Improvement for hospitalized patients requiring supplemental oxygen
        - 1.47 (95% CI, 1.17–1.84)
    - No improvement for hospitalized patients on noninvasive ventilation or high-flow oxygen
        - 1.38 (95% CI, 0.79–1.81)
- Secondary outcomes:
    - Odds of improvement in the ordinal scale were greater at 15 days in the remdesivir group as compared with the placebo group: 1.50 (95% CI, 1.18–1.91; $P=.001$; 844 patients)
    - Mortality was numerically lower in the remdesivir group than in the placebo group, but the difference was not significant (hazard ratio for death 0.70; 95% CI, 0.47–1.04).

LIMITATIONS:

- Median time from symptom onset to randomization was nine days.
- The following limitations likely impacted data collection and monitoring of adverse effects:
    - Restricted travel, hospitals restricted the entrance of nonessential personnel
    - Training initiation visits and monitoring visits performed remotely
    - Research staff dealt with other clinical duties
    - Staff illnesses strained research resources
    - Not enough supplies

EBP

Nathan Razbannia, MD
CMU College of Medicine
Saginaw, MI

Remdesivir—Pushing the Envelope Amidst Covid-19?

Compassionate Use of Remdesivir for Patients With Severe Covid-19

Grein J, Ohmagari N, Shin D, et al. Compassionate use of remdesivir for patients with severe covid-19. *N Engl J Med*. 2020. doi:10.1056/NEJMoa2007016.
DOI 10.1097/EBP.0000000000001192

KEY TAKEAWAY: Hospitalized patients with SARS-CoV-2 infection (covid-19) treated with remdesivir may experience improvement in oxygenation and show decreased mortality. The medication appeared to have more effect in patients younger than 70 years and those not receiving mechanical ventilation.

STUDY DESIGN: Prospective Cohort Study.

LEVEL OF EVIDENCE: Step 3.

BACKGROUND: Supportive care with oxygen and off-label medications is being urgently explored as interventions to decrease poor patient outcomes in covid-19. Remdesivir, a nucleotide analogue, has shown effectiveness in vitro against the virus. The drug manufacturer has provided remdesivir for compassionate use in hospitalized patients with covid-19. This study details initial data from a larger planned study of remdesivir in covid-19 patients.

PATIENTS: Patients—Adult patients hospitalized with COVID-19.

INTERVENTION: Intervention—Treatment with 10-day course of remdesivir.

CONTROL: Comparison—None.

OUTCOME: Outcome—Incidence of key clinical events:

- Changes in oxygen-support requirements
- Hospital discharge
- Reported adverse events (including discontinuation of treatment, serious adverse events)
- Death
- Proportion of patients with clinical improvement per live discharge and decrease of at least two points from baseline on modified ordinal scale recommended by the World Health Organization.

METHODS BRIEF DESCRIPTION: Clinicians enrolled hospitalized patients with positive SARS-CoV-2 reverse transcriptase polymerase chain reaction tests in a compassionate remdesivir program.

INCLUSION/EXCLUSION CRITERIA: Oxygen saturation of less than 94% on room air or need for oxygen support, creatinine clearance of >30 mL/min, alanine aminotransferase and aspartate aminotransferase of less than five times the normal range, and agreement to forego other agents for covid-19.

Patients received a 10-day course of remdesivir with day 1 IV loading dose of 200 mg, followed by 9 days of 100 mg.

Patients were followed for 28 days or until discharge or death.

Regulatory and Institutional review board, ethics committee approval, and patient consents were obtained.

Gilead Sciences sponsored and conducted this study.
Clinical improvement and mortality were evaluated with Kaplan-Meier analysis. Associations with pretreatment characteristics were evaluated with Cox proportional hazards regression.

INTERVENTION (# IN THE GROUP): 61 (53 included in analysis)
COMPARISON (# IN THE GROUP): N/A

FOLLOW-UP PERIOD: 28 days or until discharge or death

RESULTS:
Total of 53 patients enrolled: 22 from United States, nine from Japan, 12 from Italy, one from Austria, one from Canada, four from France, two from Germany, one from the Netherlands, and one from Spain.
Forty patients (75%) were men, and median age was 64 years old. At trial start 34 patients (64%) were receiving invasive ventilation: 30 (57%) receiving mechanical ventilations and 4 (8%) ECMO.
Treatment lengths varied from less than 5 days to 10 days of treatment; 40 received the full 10-day course.
Median follow-up was 18 days.
Outcomes
- Changes in oxygen-support requirements: 36 of 53 patients (68%) showed improvement in oxygen support, eight of 53 (15%) worsened.
- Hospital discharge: 25 of 53 patients (47%) discharged.
- Reported adverse events: 32 of 53 patients (60%) reported increased hepatic enzymes, diarrhea, rash, renal impairment, and hypotension), 12 of 53 patients (23%) reported multiple-organ dysfunction syndrome, septic shock, acute kidney injury, and hypotension.
- Mortality seven of 53 (13%) died. Risk of death of those older than 70 years versus younger age hazard ratio was 11.34 (95% CI 1.36–94.17). Mortality was also higher in patients receiving ventilation as opposed to those receiving noninvasive ventilation with hazard ratio of 2.78 (95% CI, 0.33–23.19)
- Proportion of patients with clinical improvement per live discharge and decrease of at least two points from baseline on modified ordinal scale was 84% (95% CI, 70–99). Clinical improvement occurred less frequently in patients older than 70 years versus those younger than 50 years with hazard ratio of 0.29 (95% CI, 0.11–0.74) and in ventilated versus noninvasive ventilated patients with hazard ratio of 0.33 (95% CI, 0.16–0.68)

LIMITATIONS:
- Small cohort without comparison group.
- No prespecified end points.
- Industry funded.
- Lack of completely uniform therapy.

EBP

Catherine Marzan, MD
Janel Kam-Magruder, MD
Lawrence Smith, MD
Alaska Family Medicine Residency Program
Anchorage, AK

The authors declare no conflicts of interest.

Do botulinum toxin type A injections decrease the frequency and severity of chronic migraines?

EVIDENCE-BASED ANSWER

Yes. In patients with chronic migraine, botulinum toxin type A injections decrease headache frequency by 1.6 to 3.1 days per month (SOR: **A**, meta-analyses of randomized controlled trials [RCTs]) and severity by more than two points on 0 to 10 visual analog pain scale (SOR: **B**, meta-analyses of small RCTs). However, it also leads to more adverse events and withdrawals from protocol than placebo (SOR: **A**, meta-analyses of RCTs). Botulinum should be offered to patients with chronic migraine to increase headache-free days and improve health-related quality of life (SOR: **C**, expert opinion).

DOI 10.1097/EBP.0000000000000960

A 2018 systematic review and meta-analysis of 28 randomized controlled trials (RCTs) (N=4,190) evaluated the effectiveness of botulinum toxin type A in the treatment of either chronic or episodic migraines in adults with the primary outcome of number of migraine days per month. Six RCTs in the review (N=1,572) compared botulinum with placebo in patients with chronic migraines for at least one of the outcomes of interest (see **TABLE 1**).[1] Chronic migraine was defined as >15 days of headache per month with attacks lasting four hours or more. Eighty-eight percent of participants were in two large RCTs and were treated with botulinum toxin type A with a minimum of 155 units at fixed injection sites and could also get 40 additional units at up to eight sites based on the location of their pain (Food and Drug Administration–recommended dosage is 155 units). The remainder of participants were treated with a wide range of dosing between 25 and 155 units. Abortive therapies were allowed as needed in all study participants. The primary outcome was number of migraine days per month with secondary outcomes including number of any headache days per month, headache severity, and adverse events. Consistent with standard dosing intervals of three months, pooled analysis at 90 days found that, compared with placebo, botulinum toxin significantly reduced the number of migraine days per month and number of headache days per month (see **TABLE 1**). Botulinum also decreased the severity of chronic migraine measured on a 10-cm visual analog scale. Adverse events were high in both groups with more in the botulinum group resulting in a high withdrawal rate.

A 2019 systematic review and meta-analysis of 17 RCTs (N=3,646) evaluated botulinum toxin versus placebo in adults for the primary outcome of monthly headache episodes at 90 days, six of which (N=1,546) specifically reviewed chronic migraine.[2] The slight difference in primary outcomes of this meta-analysis compared with the previous resulted in this meta-analysis including the two largest RCTs from the preceding meta-analysis and four unique, smaller RCTs that used different amounts of botulinum toxin type A (96–205 units) compared with placebo, each with fixed injection sites. Participants were allowed to use abortive therapies as needed. A significant

TABLE 1. Meta-analysis of randomized controlled trials comparing botulinum toxin type A versus placebo in the prevention of chronic migraine headache[1]

Outcome	No. of studies	No. of participants	Analysis of difference (95% CI)	Favored treatment
Migraine days per month at 90 d	4	1,497	**MD –3.1 (–4.7 to –1.4)**	**Botulinum toxin**
Headache days per month at 90 d	2	1,384	**MD –1.9 (–2.7 to –1.0)**	**Botulinum toxin**
Severity of migraine (0–10 visual analogue scale)	2	75	**MD –2.7 (–3.3 to –2.1)**	**Botulinum toxin**
Total adverse events	5	1,494	**RR 1.2 (1.1 to 1.4)**	**Placebo, NNH=9**
Adverse event—muscle weakness	2	1,379	**RR 13 (3.5 to 46)**	**Placebo, NNH=19**
Adverse event—neck pain	3	1,432	**RR 2.5 (1.5 to 4.1)**	**Placebo, NNH=26**
Withdrawals because of adverse events	2	1,384	**RR 3.7 (1.4 to 10.0)**	**Placebo, NNH=49**

Statistically significant differences in bold. MD=mean difference; NNH=number needed to harm; RR=relative risk.

TABLE 2. Meta-analysis of randomized controlled trials comparing botulinum toxin type A versus placebo in prevention of chronic migraine headache[2]

Outcome	No. of studies	No. of participants	Analysis of difference (95% CI)	Favored treatment
Headache episodes per month at 90 d	5	1,546	**MD –1.6 (–3.1 to –0.07)**	**Botulinum toxin**
Headache episodes per month at 60 d	5	1,546	**MD –1.6 (–2.7 to –0.47)**	**Botulinum toxin**
Quality of life at 90 d	4	1,520	**SMD –0.39 (–0.51 to –0.28)**	**Botulinum toxin**
Total adverse events at 90 d	5	1,509	**RR 1.2 (1.1 to 1.3)**	**Placebo**

Statistically significant differences in bold. MD=mean difference; RR=relative risk; SMD=standardized mean difference.

reduction in migraine frequency per month was noted with botulinum compared with placebo at both 60 and 90 days (see **TABLE 2**). Quality of life was assessed on different scales such as the Headache Disability Inventory, Beck Depression Inventory-II, Migraine Disability Assessment, and the Headache Impact Test at 90 days posttreatment, so pooled analysis was reported using standardized mean differences. Botulinum led to a small to moderate improvement in quality of life. Total adverse events were more common with botulinum (see **TABLE 2**).

In 2016, the American Academy of Neurology published an evidence-based guideline on treating chronic migraine with botulinum toxin type A based on the two large RCTs included in the above systemic reviews.[3] They concluded that botulinum was effective and should be offered to patients with chronic migraine to increase headache-free days and improve health-related quality of life in chronic migraine. EBP

Heidi Beery, MD
Harry Taylor, MD, MPH
Jeff MacEwen, OMS-IV
Roseburg Family Medicine Residency, Roseburg, OR

The authors declare no conflicts of interest.

References

1. Herd CP, Tomlinson CL, Rick C, et al. Botulinum toxins for the prevention of migraine in adults. *Cochrane Database Syst Rev*. 2018; 6:CD011616. [STEP 1]
2. Bruloy E, Sinna R, Grolleau J-L, Bout-Roumazeilles A, Berard E, Chaput B. Botulinum toxin versus placebo: a meta-analysis of prophylactic treatment for migraine. *Plast Reconstr Surg*. 2019; 143(1):239–250. [STEP 1]
3. Simpson DM, Hallett M, Ashman EJ, et al. Practice guideline update summary: botulinum neurotoxin for the treatment of blepharospasm, cervical dystonia, adult spasticity, and headache: report of the Guideline Development Subcommittee of the American Academy of Neurology. *Neurology*. 2016; 86(19):1818–1826. [STEP 5]

In women with prelabor rupture of membranes at term, is oxytocin more effective than misoprostol in preventing chorioamnionitis?

EVIDENCE-BASED ANSWER

For women with prelabor rupture of membranes (PROM), induction with oxytocin does not seem to be superior to misoprostol for preventing chorioamnionitis (SOR: **B**, systematic review of low-quality randomized control trials [RCTs] and single RCTs). Experts recommend immediate induction of labor over expectant management for pregnant patients with PROM at term to lessen the risk of chorioamnionitis, noting that either option is a valid method for labor induction (SOR: **C**, practice bulletin).

DOI 10.1097/EBP.0000000000000962

A 2017 systematic review and meta-analysis evaluated 23 randomized control trials (RCTs) (N=8,615) to determine the effectiveness of labor induction within 24 hours of ruptured membranes versus expectant management for pregnant women with prelabor rupture of membranes (PROM).[1] The review included trials with women with singleton pregnancies and PROM at 37 weeks or greater gestation. Interventions in 21 RCTs were IV oxytocin (10 trials), vaginal prostaglandin E2 (six trials), and oral, sublingual, or vaginal misoprostol (six trials). Expectant management was typically observation and monitoring for up to 24 hours. The primary outcome was maternal infectious morbidity (chorioamnionitis or endometritis). Overall, labor induction within 24 hours reduced the risk of maternal infections

by 50% (eight trials, N=6,864; relative risk 0.5; 95% CI, 0.3–0.7). In subgroup analyses, IV oxytocin and sublingual misoprostol were more effective than expectant management; however, the benefit from oral misoprostol or vaginal prostaglandin E2 was not significantly different from expectant management (see **TABLE**). The authors deemed the evidence to be low quality, mainly because of unclear risk of bias.

The largest RCT in the above review[1] evaluated 5,041 pregnant women at term with PROM to determine if labor induction was superior to expectant management in reducing the risk of maternal or neonatal infection.[2] The patients were on average 28.4 years old at 38.9 weeks' gestation with a singleton pregnancy, recruited from 72 hospitals in Australia, Canada, Denmark, Israel, Sweden, and the United Kingdom; 59.6% were primigravidas. Women were assigned to one of four groups: a) immediate induction with IV oxytocin (per location protocol); b) immediate induction with 1 or 2 mg of prostaglandin E2 gel inserted into the posterior vaginal fornix, repeated (if necessary) six hours later, and ultimately IV oxytocin (if needed); c) expectant management, supplemented with IV oxytocin induction if complications arose or labor had not started after four days; and d) expectant management, supplemented with prostaglandin gel induction if complications or no labor after four days, and ultimately IV oxytocin (if needed). Overall, 57.2% received oxytocin, ranging from 43.1% of the prostaglandin immediate induction to 91.9% of the oxytocin immediate induction groups, and 26.8% received prostaglandin, ranging from 0.6% of the oxytocin immediate induction to 88.3% of the prostaglandin immediate induction groups. In the immediate induction with oxytocin group, 4% developed chorioamnionitis compared with 8.6% of women in the expectant management (with oxytocin if needed) group ($P<.001$). Chorioamnionitis rates were also higher in the prostaglandin immediate induction (6.2%) and expectant management (with prostaglandin if needed) (7.8%) groups, but these were not significantly different from the oxytocin immediate induction group ($P>.045$). The study was limited by the inability to blind the patients and clinicians to the intervention; however, an adjudication committee that was not aware of the group assignments assessed the outcomes.

A 2018 RCT of 184 women assessed the safety and efficacy of vaginal prostaglandin compared with oxytocin for induction of labor in women with term PROM.[3] On average, the patients were 26.1 years old (8.7% >35 years old) at 37 to 42 weeks' gestation; 56.5% were nulliparous. They were given either IV oxytocin or prostaglandin E2 gel, dosed as 1 mg (for multigravida) or 2 mg (for primigravida) inserted into the posterior vaginal fornix every six hours (for up to three doses), and ultimately IV oxytocin (if needed). Overall, 67.9% received oxytocin, ranging from 46.7% of the prostaglandin induction to 88.3% of the oxytocin induction groups. No difference was found in rates of maternal infection (maternal fever, chorioamnionitis, or other maternal infection) in the women given oxytocin (11.7%) or prostaglandin (11.1%), relative risk 1.1; 95% CI, 0.5–2.4. Limitations included the lack of blinding of patients and study staff, and small number of patients.

A 2018 American College of Obstetrics and Gynecology practice bulletin recommended labor induction for women with PROM at 37 weeks or greater, generally using oxytocin, if labor did not occur near the time of presentation, and assuming no contraindications existed (level B: based on limited and inconsistent scientific evidence).[4] The authors noted that labor induction in these patients has been shown to reduce rates of chorioamnionitis, endometritis, or both, and infant admission to the neonatal intensive care unit, without increasing the risk of cesarean section or operative vaginal delivery. Furthermore, the bulletin stated that prostaglandin induction had been shown to be equally effective as oxytocin, but the former had been associated with higher rates of chorioamnionitis, based on the previously mentioned RCT[2] (no level of evidence given). EBP

TABLE. Relative risk of maternal infections (chorioamnionitis or endometritis) for pregnant women with prelabor rupture of membranes with labor induction within 24 hours of ruptured membranes (intervention) versus expectant management

No. of trials	No. of patients	Intervention	Relative risk (95% CI)
5	3,625	IV oxytocin	**0.6 (0.4–0.9)**
1	84	Oral misoprostol	0.01 (0.01–1.6)
1	560	Sublingual misoprostol	**0.2 (0.1–0.4)**
2	2,595	Vaginal prostaglandin E2	0.7 (0.4–1.2)

Statistically significant results in bold font. Data from a systematic review and meta-analysis of randomized control trials.[1]

Erin Longley, MD
Joseph Angel, MD
Danielle Davies, MD
University of Washington, Seattle, WA

The authors declare no conflicts of interest.

References

1. Middleton P, Shepherd E, Flenady V, McBain RD, Crowther CA. Planned early birth versus expectant management (waiting) for prelabour rupture of membranes at term (37 weeks or more). *Cochrane Database Syst Rev*. 2017; 1: CD005302. [STEP 1]
2. Hannah ME, Ohlsson A, Farine D, et al. Induction of labor compared with expectant management for prelabor rupture of the membranes at term. TERMPROM Study Group. *N Engl J Med*. 1996; 334(16):1005–1010. [STEP 2]
3. Mahomed K, Wild K, Weekes CR. Prostaglandin gel versus oxytocin—prelabour rupture of membranes at term—a randomised controlled trial. *Aust N Z J Obstet Gynaecol*. 2018; 58(6):654–659. [STEP 2]
4. Committee on Practice Bulletins-Obstetrics. ACOG practice bulletin no. 188: prelabor rupture of membranes. *Obstet Gynecol*. 2018; 131(1):e1–e14. [STEP 5]

What is the most accurate noninvasive test for diagnosis of *H pylori* infection?

EVIDENCE-BASED ANSWER

Urea breath tests appear to have higher diagnostic accuracy for *Helicobacter pylori* infection than serology or stool antigen tests in symptomatic patients without gastrectomy or recent proton pump inhibitor or antibiotic use (SOR: **B**, systematic review and meta-analysis of indirect test comparisons with methodologic limitations). Urea breath testing and stool *H pylori* antigen are preferred over serology for the noninvasive diagnosis of *H pylori* infection (SOR: **C**, expert opinion).

DOI 10.1097/EBP.0000000000001025

A 2018 systematic review and meta-analysis evaluated the diagnostic accuracy of four commonly used noninvasive tests for the diagnosis of *H pylori* infection.[1] The review included 99 diagnostic accuracy studies with a total of 10,799 patients; 5,694 (53%) of whom had *H pylori* infection diagnosed by the reference standard; 14 studies enrolled only children. Most of the studies (57%) evaluated patients with dyspepsia or abdominal pain, and the authors excluded case-control studies and those enrolling patients with acute upper gastrointestinal bleeding. The median *H pylori* prevalence was 53.7% (range, 15.2–94.7%); half of the studies had an *H pylori* prevalence between 42.0% and 66.5% (interquartile range). The four index tests were the urea breath test labeled with either carbon-13 or carbon-14 radioisotopes, *H pylori* stool antigen, and serology. The reference standard was endoscopic biopsy with histology confirmation. The authors calculated the diagnostic odds ratio (DOR)—the likelihood ratio (LR) of a positive test divided by the LR of a negative test (LR+/LR–). Using pooled data from 99 studies (N=10,799; 5,694 diagnosed with *H pylori* by the reference standard), the authors indirectly compared the four index tests with the reference standard. In these meta-analyses, both urea breath tests had significantly higher DORs than either stool antigen or serology (see **TABLE**). However, in seven studies (N=495) directly comparing the carbon-13 urea breath test against serology, there was no significant difference between the DORs. Similarly, in seven studies (N=608) directly comparing carbon-13 urea breath test against *H pylori* stool antigen assay, there was also no significant difference between the DORs. The review was limited by the paucity of studies directly comparing the index tests, as well as the variable reference standard used in the indirect comparisons. Additionally, many studies excluded participants with prior gastrectomy and recent antibiotic or proton pump inhibitor use, limiting generalizability to these groups.

A 2016 clinical practice guideline of Alberta, Canada, for the diagnosis and treatment of *H pylori* infection in adults recommended the urea breath test for noninvasive testing of *H pylori* infection in patients without alarm symptoms, describing its sensitivity and specificity as "superior to any other diagnostic test" (no strength of recommendation or evidence grade provided).[2] The guideline recommended against using IgG serology or stool antigen tests for the diagnosis of *H pylori.* The urea breath test was also recommended for the confirmation of eradication of *H pylori*, with consideration for the stool antigen test only if urea breath test was not available. The guideline emphasized the importance of a 28-day washout period after completing antibiotics and three days after discontinuing proton pump inhibitor therapy before repeating urea breath testing. Guideline authors also noted that stool antigen test results might be affected by stool temperature, consistency, and

TABLE. Diagnostic accuracy of noninvasive tests for detecting *Helicobacter pylori* infection compared against a reference standard of endoscopic biopsy with histology

Index test	No. of studies	No. of patients	Sensitivity (95% CI)	DOR[a] (95% CI)
Carbon-13 urea breath test	34	3,139	94% (89%–97%)	153 (73.7–316)
Carbon-14 urea breath test	21	1,810	92% (89%–94%)	105 (74.0–150)
Serology	34	4,242	84% (74%–91%)	47.4 (25.5–88.1)
Stool antigen	29	2,988	83% (73%–90%)	45.1 (24.2–84.1)

A higher DOR correlates with a more accurate test.[1] [a] The DOR is defined as the LR of a positive test divided by the likelihood ratio of a negative test (LR+/LR–). DOR=diagnostic odds ratio; LR=likelihood ratio.

time interval between collection and measurement. The guideline recommendations were based on a committee consensus after reviewing evidence from a systematic review of the literature.

In 2019 the American Society for Clinical Pathology (ASCP) released a "Choosing Wisely" statement that recommended against serologic evaluation of patients to determine the presence of *H pylori* infection (no strength of recommendation or evidence grade provided).[3] The ASCP noted that other noninvasive methods, such as urea breath and stool antigen tests, have greater accuracy in detecting the presence of *H pylori* and have demonstrated higher clinical utility, sensitivity, and specificity. Choosing Wisely is a 2012 initiative of the American Board of Internal Medicine Foundation, and member organizations provide evidence-based statements to support and engage physicians in making better decisions in utilization of limited health care resources. EBP

Chelsea Carlson, MD
Family Medicine Residency of Idaho
Boise, ID

Adam Voelckers, MD
Northwest Washington Family Medicine Residency
Bremerton, WA

Jeremy Ginoza, DO
Skagit Regional Health Family Medicine Residency
Mount Vernon, WA

The authors declare no conflicts of interest.

References

1. Best LMJ, Takwoingi Y, Siddique S, et al. Non-invasive diagnostic tests for Helicobacter pylori infection. *Cochrane Database Syst Rev.* 2018; (3):CD012080. [STEP 1]
2. Toward Optimized Practice (TOP) Working Group for H. pylori. *Diagnosis and Treatment of Helicobacter pylori: Clinical Practice Guideline*. Edmonton, AB: Toward Optimized Practice; 2016. [STEP 5]
3. American Society for Clinical Pathology. *Choosing Wisely: Thirty Things Physicians and Patients Should Question*. Released February 21, 2014 (1–5), February 3, 2015 (6–10), September 14, 2016 (11–15), September 19, 2017 (16–20), September 25, 2018 (21-25) and September 4, 2019 (26-30). https://www.ascp.org/content/docs/default-source/get-involved-pdfs/istp_choosingwisely/2019_ascp-30-things-list.pdf. Accessed January 17, 2020. [STEP 5]

How effective is caffeine plus an analgesic compared with analgesic alone for the treatment of acute headache pain in adults?

EVIDENCE-BASED ANSWER

When compared with analgesic such as acetaminophen or ibuprofen alone, the combination of caffeine plus an analgesic appears superior in the treatment of acute headache in both pain relief and time to meaningful pain relief (SOR: **A**, consistent randomized controlled trials). The evidence is conflicting on the relative efficacy of caffeine-analgesic combinations and sumatriptan (no SOR given).

DOI 10.1097/EBP.0000000000000895

A 2014 meta-analysis of four randomized controlled, double-blind, crossover trials compared the efficacy of a combination of 250 mg acetylsalicylic acid, 250 mg

acetaminophen, and 65 mg caffeine (AAC) versus 1,000 mg acetaminophen (APAP) in treating tension-type headaches.[1] The study included 1,376 patients with 2,737 occurrences of tension headaches. Eligible patients were between 18 and 65 years old and averaged 4 to 10 headaches per month. Patients were given a single dose of one of the above interventions and then observed for four hours. More patients in the AAC group had complete resolution of headaches two hours after medication administration compared with APAP (29% AAC vs 21% APAP, $P<.0001$). A higher percentage of patients in the AAC group also reported any reduction in headache intensity, one hour after administration (8.6% AAC vs 6.1% APAP, $P=.004$) and two hours after administration (67% AAC vs 58% APAP, $P<.001$).

A 2014 randomized placebo-controlled controlled trial evaluated the efficacy of severe headache treatment with acetaminophen 500 mg, acetylsalicylic acid 500 mg, and caffeine 130 mg (AAC) versus ibuprofen 400 mg or placebo in 660 healthy adults with migraine headache and two to six attacks per month.[2] Patients were given a single dose of one of the above interventions and then monitored for six hours. AAC showed greater clinical improvement in total pain relief when compared with ibuprofen ($P=.037$, results presented in graphic form) and time to meaningful pain relief (132 minutes vs 148 minutes, $P=.026$).

A 2005 multicenter, double-blind, randomized, parallel-group, single-dose study evaluated headache pain relief at two hours and time to meaningful headache pain relief with the combination of acetaminophen 500 mg, acetylsalicylic acid 500 mg, and caffeine 130 mg (AAC) to ibuprofen 400 mg.[3] This trial included 1,335 adults, 18 years old and older who experience two to six migraine headaches per month. Patients received a single dose of one of the above interventions and were monitored for 6 hours. AAC was found to have significantly better pain relief at two hours (0 to 4 pain scale; 2.7 vs 2.4, $P<.03$) and a shorter time to meaningful pain relief (128 minutes vs 148 minutes, $P<.036$) than ibuprofen.

A 2012 randomized controlled trial evaluated the effectiveness of migraine treatment with indomethacin 25 mg, prochlorperazine 2 mg, and caffeine 75 mg (Indoprocaf®) compared with sumatriptan 50 mg.[4] The trial included 297 adults 18 to 65 years old with one to six migraines per month. Patients were treated with an initial dose of one of the above interventions for two headaches separated by at least 48 hours. Patients were also given a second dose of the same intervention for each headache to be used if needed. Initial dosing of Indoprocaf and sumatriptan were similarly effective. However, pain-free rates at two hours after a second dose were significantly higher with Indoprocaf than with sumatriptan (60% vs 50%, $P<.05$). Headache relief was also significantly better after a second dose with Indoprocaf than sumatriptan at two hours (65% vs 45%, $P<.05$).

A 2007 double-blind, double-dummy, crossover randomized controlled trial compared the efficacy of paracetamol with caffeine (PCF) versus sumatriptan for the treatment of migraine headaches in 108 adults.[5] Patients were 18 to 62 years old with two to eight migraines per month. Patients treated three consecutive migraine headaches with a single dose of an intervention medication as determined by a randomized crossover sequence generator. Both treatments were similar in efficacy for achieving total pain relief (74% with PCF vs 72% with sumatriptan; $P=.98$). Potential confounders were that inclusion criteria included consumption of at least two cups of coffee a day for all patients and that patients were able to take a rescue medication three hours after the initial administration of the study medicine. Medications were self-administered at home, making it impossible to verify which medication was taken. EBP

Greg Maeder, MD, MBA
Amanda Vester, DO
Dominick Ruybal, DO
Amanda Torczynski, MD
Mollie Ebner, DO
Gabriela Kaldan, DO
HealthONE Family Medicine Residency, Denver, CO

The authors declare no conflicts of interest.

References

1. Diener HC, Gold M, Hagen M. Use of a fixed combination of acetylsalicylic acid, acetaminophen and caffeine compared with acetaminophen alone in episodic tension-type headache: a meta-analysis of four randomized, double-blind, placebo-controlled, crossover studies. *J Headache Pain*. 2014; 15:76. [STEP 1]
2. Goldstein J, Hagen M, Gold M. Results of a multicenter, double-blind, randomized, parallel-group, placebo-controlled, single-dose study comparing the fixed combination of acetaminophen, acetylsalicylic acid, and caffeine with ibuprofen for acute treatment of patients with severe migraine. *Cephalalgia*. 2014; 34(13):1070–1078. [STEP 2]
3. Goldstein J, Siberstein S, Saper J. Acetaminophen, aspirin, and caffeine in combination versus ibuprofen for acute migraine: results from a multicenter, double-blind, randomized, parallel-group, single-dose, placebo-controlled study. *J Headache*. 2006; 46:444–453. [STEP 2]

4. Sandrini G, Cerbo R, Del Bene E, et al. Efficacy of dosing and re-dosing of two oral fixed combinations of indomethacin, prochlorperazine and caffeine compared with oral sumatriptan in the acute treatment of multiple migraine attacks: a double-blind, double-dummy, randomised, parallel group, multicentre study. *Int J Clin Pract.*. 2007; 61(8):1256–1269. [STEP 2]
5. Pini LA, Guerzoni S, Cainazzo M, Ciccarese M, Prudenzano MP, Livrea P. Comparison of tolerability and efficacy of a combination of paracetamol + caffeine and sumatriptan in the treatment of migraine attack: a randomized, double-blind, double-dummy, cross-over study. *J Headache Pain*. 2012; 13(8):669–675. [STEP 3]

In patients with simple lacerations, are adhesive interventions as effective as standard wound closure methods?

EVIDENCE-BASED ANSWER

Yes. When compared with sutures, staples, or adhesive strips, adhesive interventions provide no difference in wound cosmesis while decreasing overall procedure time, rate of wound erythema, and pain levels, but carry a small increased risk of dehiscence (SOR: **A**, systematic review of randomized controlled trials). Length of emergency department stay was 26 minutes shorter in patients whose wounds were repaired using tissue adhesive interventions (SOR: **C**, cohort study).

DOI 10.1097/EBP.0000000000000824

A 2009 meta-analysis of 13 randomized controlled trials (N=1,328) compared tissue adhesives with standard wound closure (sutures, staples, or adhesive strips) or another tissue adhesive in adults and children in an emergency or primary care setting.[1] Tissue adhesive included octylcyanoacrylate and butylcyanoacrylate. The primary outcome measured was cosmetic quality, using the Cosmetic Visual Analogue Scale (CVAS) or Wound Evaluation Score (WES). The CVAS is a validated system designed to evaluate laceration repairs using a visual analog scale of 0 to 100 mm (100=best possible scar). A measurement of 12 to 15 mm was considered the minimum clinically significant difference between optimal and suboptimal scars. The WES assesses six clinical variables of each scar, including edge inversion, absence of step-off, contour irregularities, wound margin separation (>2 mm), excessive distortion, and overall cosmetic appearance. A WES score of six (out of a possible six) is considered "optimal," whereas a score of five or less is considered "suboptimal." Studies reported the number of optimal results per group. Secondary outcomes included pain scores, procedure completion time, and the occurrence of complications. Pain was evaluated using a visual analogue scale (VAS) by parents of pediatric patients (higher score indicating worse pain). At five to 14 days, one to three months, and nine to 12 months, the cosmetic outcomes using CVAS showed no difference between tissue adhesives and standard wound closure (one study, n=52; weighted mean difference [WMD] 0.0 mm; 95% CI, –4.8 to 4.8 mm; seven studies, n=549; WMD 1.6 mm; 95% CI, –3.2 to 6.4; and four studies, n=364; WMD 1.5 mm; 95% CI, –3.1 to 6.1). Additionally, the number of optimal results using WES scoring at five to 14 days, one to three months, and nine to 12 months, no difference was found between tissue adhesives and standard wound closure (two studies, n=195; relative risk [RR] 0.98; 95% CI, 0.87–1.1; four studies, n=364; RR 0.99; 95% CI, 0.89–1.1; and two studies, n=140; RR 1.1; 95% CI, 0.89–1.3). Pain scores and procedure time significantly favored tissue adhesives (five studies, n=434; WMD –13 mm on 100 mm parent-reported VAS; 95% CI, –20.0 to –6.9 mm and six studies, n=584; WMD –4.7 minutes; 95% CI, –7.2 to –2.1 min). Dehiscence was higher with tissues adhesives than standard wound closure (nine studies, n=834; risk difference 2.4%; 95% CI, 0.1–4.9%). Studies were limited by lack of randomization (selection bias), inability to double-blind because of the nature of the interventions, heterogeneity, lack of gold standard for pain criterion, and inadequate allocation concealment.

A 2019 retrospective cohort study of 8.7 million patients who received either standard wound closure methods or tissue adhesives evaluated emergency department length of stay.[2] Of included patients, 63% were female, 42% were pediatric patients, and the mean age was 30 years old. The population consisted of adult and pediatric patients who presented to the emergency department with an isolated traumatic laceration identified using International Classification of Diseases codes from the National Hospital Ambulatory Medical Care Survey database. Of patients with a single laceration closed with either tissue adhesives or mechanical means (sutures and staples), emergency department length of stay was significantly shorter in patients whose wounds were closed with tissue adhesives (101 vs 136 minutes; $P = .001$). After adjusting for potential confounding variables, the use of

tissue adhesives was still associated with a shorter emergency department length of stay (26 minutes; 95% CI, 9–44 minutes). This study was limited by the presence of confounding variables, such as laceration length, depth, complexity, and patient preferences. EBP

Tyson Schwab, MD, MS
Sarah Daly, DO, FAAFP
Utah Valley FMR, Provo, UT

The authors declare no conflicts of interest.

References

1. Farion KJ, Russell KF, Osmond MH, et al. Tissue adhesives for traumatic lacerations in children and adults. *Cochrane Database Syst Rev.* 2002; (3):CD003326. [STEP 1]
2. Otterness K, Thode HC Jr, Singer AJ. Methods of laceration closure in the ED: a national perspective. *Am J Emerg Med*. 2019. In press. [STEP 3]

How effective are mass media interventions for smoking cessation in adults?

EVIDENCE-BASED ANSWER

Mass media (including billboards, television, and radio ads) regarding smoking cessation appear to have a modest effect on smoking cessation and intention to quit. The use of interactive social media applications increases cessation, but a Quit Line was more effective than an interactive app (SOR: **B**, based on systematic review of randomized controlled trials [RCTs]). "Why to quit" adds with personal stories or graphic images may also be effective with smoking cessation (SOR: **B**, RCT).

DOI 10.1097/EBP.0000000000000866

A systematic review of 11 controlled trials reviewed effect of mass media campaigns in reducing levels of smoking among adults (N=1,965,478).[1] The demographics of the population included adults of 25 years and older with additional identifiers not defined. Mass media campaigns included television, radio, newspaper, billboards, posters, leaflets, or booklets alone or in conjunction with tobacco control programs. The primary outcome was tobacco cessation, including quit and prevalence rates. The trials compared the effects of mass media campaigns in the exposed areas and to control areas not exposed to the campaign. Two trials were state wide (California and Massachusetts) compared with the rest of the United States. Smoking outcomes were examined in the whole population, mainly by prevalence. For smoking prevalence with follow-up six to 18 months, the state-wide campaigns resulted in greater declines in smoking compared with the rest of the United States (results summarized, no data presented). Two other trials focused on Vietnamese-American men, where one trial detected a difference in smoking prevalence at two years (one trial; odds ratio [OR]: 1.7; 95% CI, 1.3–2.2), but the other trial did not. Of the seven community trials, four resulted in a significant difference in quit rates in the campaign area, whereas two did not report a significant difference. The data were not combined because of significant heterogeneity of the trials. The authors concluded that mass media campaigns may affect smoking behavior, but the available trials are of variable quality and are conducted in settings where there are other influences on tobacco behavior, making it difficult to assess the mass media intervention.

A systematic review of four randomized controlled trials and three feasibility studies (N=9,755) examined the use of social media interventions for smoking cessation.[2] Patient demographics included people who use social media and smoke tobacco. Additionally, patients could be from any population group. All seven studies used social media interventions either on its own or in conjunction with a phone app to promote smoking cessation via interactive modalities (private group discussions, tweets, etc.). Smokers using a Smoker's Helpline, versus those who used an interactive app were more likely to have higher quit rates at three months (one trial; N=238; 14% vs 32%, $P<.001$). In another study, abstinence was greater among participants using the "Tweet2Quit" program versus current standard resources at 60 days (one trial; N=160; 55% vs 41%; $P=.021$). Another study demonstrated that posting more comments in a specific Facebook group related to cessation was associated with biochemically verified abstinence at three months (one trial; N=79; $P=.036$; no data provided).

A randomized field trial compared high-dose media markets (HDM) to standard media markets (SDM) to assess the role of media in smoking cessation among randomly selected populations (N=8,576).[3] Both smokers and nonsmokers were included to the study. The primary outcome for smokers was defined as at least one quit attempt lasting one day or longer (within 30 days to six months), and for nonsmokers, key outcome measures were communication with friends and family about the dangers of smoking. A three-month national media campaign was supplemented

within 67 (of 190) randomly selected media markets, with high dose being defined as three times exposure of that of standard dose markets. Based on the collected survey data after the termination of the campaign, the quit attempt rate among smokers was found to be higher in HDM versus SDM (39% vs 35%; $P<.04$), this impact being the most meaningful among African Americans (51% vs 32%; $P<.01$). Nonsmokers in HDM were also found to be much more likely to talk to family and friends about smoking cessation (43% vs 36%; $P<.01$).

A randomized control trial evaluated smoking-related beliefs and attitudes, quit intentions, and smoking behaviors over four weeks in adult smokers (N=3,002) who viewed 30-second why-to-quit ads.[4] These ads consisted of why-to-quit ads with emotion-evoking personal testimonies (WTQ-T) or graphic images (WTQ-G), how-to-quit ads (HTQ), or a combination of those. These groups were compared with a control group that viewed no ads. In three groups who viewed ads, smokers were more likely to have quit smoking at 4 weeks than the control group: the WTQ-T group (OR, 10; 95% CI, 3.5–30), the WTQ-G group (OR, 6.8; 95% CI, 2.7–17), and the WTQ-T + HTQ (OR, 5.9; 95% CI, 1.5–23). The group who viewed HTQ alone had no decrease in smoking cessation after four weeks. EBP

Marc Pascuzzo, PsyD
Andrew Schmutz, DO
Brandon Hoy, DO
Ciprian Viziteu, DO
Sarah Patzner, DO
HealthONE Family Medicine Residency at Sky Ridge
Lone Tree, CO

The authors declare no conflicts of interest.

References

1. Bala MM, Strzeszynski L, Topor-Madry R. Mass media interventions for cessation in adults. *Cochrane Database Syst Rev*. 2017; 11:CD004704. **[STEP 2]**
2. Naslund JA, Kim SJ, Aschbrenner KA, et al. Systematic review of social media interventions for smoking cessation. *Addict Behav*. 2017; 73:81–93. **[STEP1]**
3. McAfee T, Davis KC, Shafer P, Patel D, Alexander R, Bunnell R. Increasing the dose of television advertising in a national antismoking media campaign: results from a randomised field trial. *Tob Control*. 2017; 26(1):19–28. **[STEP 2]**
4. Duke JC, Nonnemaker JM, Davis KC, Watson KA, Farrelly MC. The impact of cessation media messages on cessation-related outcomes: results from a national experiment of smokers. *Am J Health Promot*. 2014; 28(4):242–250. **[STEP 2]**

In patients who had appropriate initial weight loss after bariatric surgery, what patient behaviors are risk factors for regaining weight?

EVIDENCE-BASED ANSWER

When taken as a group, the eating behaviors of patients, such as grazing, picking and nibling, loss-of-control eating, and binge eating, significantly increase the risk for weight regain after bariatric surgery (SOR: **B**, meta-analysis of cross-sectional and cohort studies). When analyzed separately, grazing behavior is inconsistently associated with weight regain (SOR: **C**, systematic review of inconsistent cross-sectional studies). Eating fast food, eating when full, eating continuously throughout the day, and greater sedentary time are also significantly associated with weight regain after bariatric surgery (SOR: **B**, prospective cohort study).

DOI 10.1097/EBP.0000000000000803

A 2019 systematic review and meta-analysis analyzed 11 cross-sectional and two prospective cohort studies (N=1,766) to evaluate whether weight regain after bariatric surgery is associated with psychopathology.[1] Of these studies, five studies (four cross-sectional studies and one prospective cohort study, N=361) specifically analyzed eating behaviors such as grazing, picking and nibling, loss-of-control eating, and binge eating and provided data that could be pooled for meta-analysis. In these five studies, behaviors were diagnosed using validated instruments such as the Diagnostic and Statistics Manual IV or the Eating Disorder Examination. Patients were adults who underwent any type of weight loss surgical procedure (Roux-en-Y gastric bypass, laparoscopic-assisted gastric banding, or sleeve gastrectomy) 18 to 24 months previously and after initial weight loss had weight regain ranging from 3% to greater than 15% of nadir. All five studies showed significant positive associations between weight regain and these eating behaviors, and pooled results showed the presence of these eating behaviors more than doubled the risk of weight regain (relative risk [RR] 2.2; 95% CI, 1.6–3.2; $I^2=70\%$) compared with patients without these behaviors. The heterogeneity was moderately high, but this was caused by one study

according to a sensitivity analysis. When that study was excluded, the I^2 statistic fell to 7%, and the RR remained significant at 1.9 (95% CI, 1.5–2.2). Limitations of this study included lack of uniformity in weight regain reporting. Additionally, in the meta-analysis, three studies were rated as good quality, one rated as fair, and one rated as poor; the latter two studies rated lower based on methods of patient selection and lack of comparison with nonrespondents.

A 2017 systematic review analyzed five cross-sectional studies (N=994) to investigate the effect of grazing behavior on weight regain after bariatric surgery.[2] Two of these studies were unique and not included in the meta-analysis above (N=547). The first of these studies aimed to assess the relationship of eating behaviors, weight outcomes, and quality of life and analyzed the responses of 497 adults who had Roux-en-Y gastric bypass surgery three to 10 years previously. It defined grazing behavior as "a pattern of eating or nibbling continuously at least 2 days a week for a 6-month period over an extended period of time in addition to an inability to stop or control their eating while nibbling." Weight regain was defined as any amount of increased weight after bariatric surgery nadir weight, and the prevalence of grazing behavior was 18%. This study showed modest correlation between grazing behavior and weight regain (correlation coefficient [*r*] 0.39, $P<.001$). The second study aimed to evaluate whether eating behavior could predict short- and long-term success postsurgery and analyzed the responses of 50 patients who had Roux-en-Y gastric bypass, laparoscopic-assisted gastric banding, or sleeve gastrectomy surgery 12 months or more previously. It defined grazing behavior as "unplanned, continuous, and repetitive eating of small amounts of food through extended time period, associated with loss of control overeating." The definition of weight regain was not reported, but the prevalence of grazing behavior was 44%. This study did not show a significant correlation between grazing behavior and weight regain (r –0.15, $P=.33$). The studies were rated as good quality, though limited insofar as they did not have a standardized grazing definition and were cross-sectional and thus causal relationships could not be determined.

A 2019 prospective cohort study aimed to identify patient behaviors and characteristics related to weight regain after Roux-en-Y gastric bypass surgery (N=1,278).[3] The patients were adults who underwent Roux-en-Y gastric bypass surgery in six U.S. cities between 2006 and 2009 and who had adequate follow-up (undefined). The age range was 19 to 75 years old with a median age of 46 years old; 80% of the patients were female, and 86% were white. Weight regain was calculated as the percentage of maximum weight lost. Weight regain occurring during or within six months after pregnancy was excluded from analysis. The study used a multivariate analysis controlling for maximum weight loss as a static fixed effect and for weight-loss medication as a time-dependent fixed effect. Results were reported as beta coefficients (β), measuring the mean difference in weight regain (as percentage of weight loss) for every 1 unit of change in patient behavior. Each behavior was compared with its reference, defined as the absence of that behavior—although sedentary time was divided into quartiles with the lowest quartile (<2 h/d) being the reference. The following patient behaviors were independently associated with postoperative weight regain compared with absence of the behavior: eating fast food meals (per each additional meal per week, β 0.45; 95% CI, 0.19–0.71), eating when feeling full (β 2.9, 95% CI 1.2–4.6), eating continuously throughout the day (β 1.6; 95% CI, 0.09–3.1), and binge eating or loss-of-control eating (β 8.0; 95% CI, 5.1–11.0). Compared with the reference of less than two hours of sedentary time per day, only the highest quartile of sedentary time (>4.5 hours per day) was significantly associated with weight regain (β 3.0; 95% CI, 1.2–4.8). One limitation was that the associations in the postsurgery models may be bidirectional; that is, just as behaviors may influence weight regain, weight regain may influence behaviors, and a cohort study cannot adjudicate directionality. EBP

Eric Christensen, MD
Thomas Satre, MD
University of Minnesota/St. Cloud Hospital Family Medicine Residency
St. Cloud, MN

The authors declare no conflicts of interest.

References

1. Mauro M, Papelbaum M, Brasil M, et al. Is weight regain after bariatric surgery associated with psychiatric comorbidity? A systematic review and meta-analysis. *Obes Rev*. 2019; 20: 1413–1425. [STEP 3]
2. Pizato N, Botelho P, Gonçalves V, Dutra E, de Carvalho KMB. Effect of grazing behavior on weight regain post-bariatric surgery: a systematic review. *Nutrients* 2017; 9: 1322. [STEP 3]
3. King W, Belle S, Hinerman A, Mitchell J, Steffen K, Courcoulas A. Patient behaviors and characteristics related to weight regain after Roux-en-Y gastric bypass: a multicenter prospective cohort study. *Ann Surg*. In press. [STEP 3]

What method of screening for anal cancer is most accurate in adults at increased risk?

EVIDENCE-BASED ANSWER

Anal cytology and human papillomavirus testing are both moderately effective at detecting anal dysplasia (SOR: **B**, three systematic reviews of cohort and cross-sectional studies with limitations).

DOI 10.1097/EBP.0000000000001013

A 2019 systematic review and meta-analysis included 12 cohort and cross-sectional studies of screening for anal cancer using cytology and human papillomavirus (HPV) testing in adult men and women with known HIV status (N=2,541) who were at increased risk for anal cancer.[1] The gold standard for diagnosis of cancer was histopathology. Based on meta-analysis of 12 studies (n=2,541), pooled sensitivity and specificity values were determined (see **TABLE**). Likelihood ratios were also calculated. The likelihood ratios for anal cytology were in the range of two to range, suggesting an association of positive cytology with a slight to moderate increased chance of disease. The area under the receiver operating curve (AUC) for cytology was 0.75, which is considered moderate test accuracy (AUC of 1 indicates a perfect test). In a subset of eight studies (n=2,079), HPV testing was evaluated (see **TABLE**) and showed similar sensitivity and specificity as cytology for detection of anal intraepithelial neoplasia (AIN). The AUC for HPV testing was 0.74. This systematic review was limited by significant heterogeneity because of differences in study design, patient population, and specimen preparation.

A 2018 review and meta-analysis also evaluated the performance of anal cytology and HPV testing.[2] All were compared with histopathology for detection of AIN-2+ samples. The review included 18 cohort and cross-sectional studies (n=6,018). Three studies (n=1,010) were included in the 2019 review. The studies used cytology and HPV testing in men with known HIV status. Anal cytology was evaluated in 14 studies (n=6,018) and HPV testing was performed in a subset of 10 studies (n=4,789). Additional sensitivity analyses were performed looking at the subset of HIV+ patients. The pooled sensitivity and specificity were obtained, and the positive and negative likelihood ratios were calculated for each test and subset (see **TABLE**). Overall, HPV testing had better sensitivity but lower specificity. Overall, this review was limited by significant heterogeneity between studies.

A 2007 Canadian health technology assessment systematic review evaluated screening test characteristics for anal cancer using cytology followed by anoscopic examination compared with histopathology.[3] The review included nine cohort studies (n=2,221) set in hospital-based specialty HIV/AIDS care clinics including mainly HIV+ men. One study from this systematic review was also included in the 2019 meta-analysis (n=401). Sensitivity and specificity values for

TABLE. Accuracy of cytology and HPV testing for detection of anal cancer[a]

	Cytology						HPV testing			
year	Total n	Percent sensitivity (95% CI)	Percent specificity (95% CI)	Positive LR	Negative LR	Subset n	Percent sensitivity (95% CI)	Percent specificity (95% CI)	Positive LR	Negative LR
2019 Meta-analysis[1]										
	2,541	0.79 (0.77–0.82)	0.66 (0.64–0.69)	2.3	3.1	2,079	0.85 (0.82–0.87)	0.46 (0.43–0.49)	1.6	0.33
2018 Meta-analysis[2]										
	6.018	0.77 (0.65–0.86)	0.56 (0.47–0.65)	1.75	0.41	4,789	0.91 (0.79–0.97)	0.33 (0.22–0.46)	1.4	0.15
	2,670 HIV+	0.81 (0.69–0.89)	0.54 (0.42–0.66)	1.8	0.35	1,518	0.95 (0.85–0.99)	0.24 (0.16–0.33)	1.3	0.21

[a] Values represent pooled estimates. HPV=human papillomavirus; LR=likelihood ratio.

anal cytology to detect AIN ranged from 46% to 69%. Specificity of anal cytology to detect AIN ranged between 59% and 81%. No CIs were given. The results of the individual studies were not pooled because of significant heterogeneity among the studies in varying prevalence of abnormal cytology, length of follow-up, and different thresholds used by the scoring pathologists. EBP

Melanie Bird, PhD
Laura Morris, MD, MSAM
University of Missouri School of Medicine
Columbia, MO

The authors declare no conflicts of interest.

References

1. Chen CC, Chou YY. Predictive value of the anal cytology for detecting anal intraepithelial neoplasia or worse: a systematic review and meta-analysis. *Diagn Cytopathology*. 2019; 47:307–314. [STEP 1]
2. Clarke MA, Wentzensen N. Strategies for screening and early detection of anal cancers: a narrative and systematic review and meta-analysis of cytology, HPV testing, and other biomarkers. *Cancer Cytopathol.* 2018; 126: 447–460. [STEP 1]
3. Medical Advisory Secretariat. Anal dysplasia screening: an evidence-based analysis. *Ont Health Technol Assess Ser.* 2007; 7(4):1–43. [STEP 1]

In adult patients with acute nonradicular low back pain, does early increased activity, as compared with bed rest, result in pain relief and increased functionality?

EVIDENCE-BASED ANSWER

Yes. Patients with acute nonradicular lower back pain experience improvements in pain and function if they receive advice to stay active compared with advice to rest in bed. (SOR: **A**, meta-analysis of randomized controlled trials). Clinicians should educate patients with acute lower back pain to remain active, provide information about effective self-care options, and avoid bed rest (SOR: **C**, guideline and expert opinion).

DOI 10.1097/EBP.0000000000000868

A 2010 meta-analysis of 10 randomized controlled trialss (N=1,923) examined the effects of advice to rest in bed or stay active on individuals with acute low back pain (pain lasting for less than six weeks) with or without sciatica.[1] Three randomized control trials specifically compared acute low back pain and functional status in patients with nonradicular pain who received advice to rest in bed versus advice to stay active (n=481). The population consisted of 59% women with an average age of 46 years. The review excluded patients with compressive disease, fracture of lumbar spine, history of cancer, infections, inflammatory disease, posttraumatic injuries, pregnancy, radiating pain below the buttocks, sciatica, spinal tumors, or urinary tract disease. The primary outcome was pain level, functional status, recovery, and return to work. Outcomes were assessed at intervals ranging from six days to three months. Interventions included either instruction for bed rest or to stay active. Best rest included instructions to take two to four days of complete bed rest, then resume activities as tolerated after. Advice to stay active included instructions to avoid bed rest and continue normal routines as actively as possible within the limits permitted by their back pain. Patients with acute lower back pain who received advice to stay active, opposed to bed rest, did not experience significantly better improvement in pain at two to four weeks (three studies, n=480; standardized mean difference [SMD], 0.02; 95% CI, –0.16 to 0.2) but did experience a small incremental improvement in pain level at 12 weeks (two studies, n=393; SMD, 0.25; 95% CI, 0.05–0.45). Additionally, patients with acute lower back pain who receive advice to stay active reported a small improvement in functional status at three to four weeks (two studies, n=400; SMD, 0.29; 95% CI, 0.09–0.49) and at 12 weeks (two studies, n=393; SMD, 0.24; 95% CI, 0.04–0.44). Included studies were limited by the lack of blinding and incomplete outcomes data (attrition bias).

In 2007, the American College of Physicians and the American Pain Society recommended clinicians educate patients with acute lower back pain to remain active and provide information about effective self-care options (Recommendation rating A, moderate-quality evidence, based on systematic reviews).[2] In 2019, the North American Spine Society joined the American Board of Internal Medicine Foundation's "Choosing Wisely" campaign and provided recommendations to specifically avoid bed rest for acute low back pain. Additionally, the Choosing Wisely program recommended patients remain as active as possible, seek positions of comfort, and participate in activities that avoid provoking symptoms (no level of evidence provided).[3] EBP

Tyson Schwab, MD, MS
Sarah Daly, DO, FAAFP
Utah Valley FMR, Provo, UT

The authors declare no conflicts of interest.

References

1. Dahm KT, Brurberg KG, Jamtvedt G, Hagen KB. Advice to rest in bed versus advice to stay active for acute low back pain and sciatica. *Cochrane Database Syst Rev*. 2010; (6): CD007612. [STEP 1]
2. Chou R, Qaseem A, Snow V, et al; Clinical Efficacy Assessment Subcommittee of the American College of Physicians; American College of Physicians; American Pain Society Low Back Pain Guidelines Panel. Diagnosis and treatment of low back pain: a joint clinical practice guideline from the American College of Physicians and the American Pain Society. *Ann Intern Med*. 2007; 147(7):478–491. [STEP 3]
3. North American Spine Society. Five things physicians and patients should question. Choosing wisely. http://www.choosingwisely.org/societies/north-american-spine-society/. Published April 2019. Accessed October 14, 2019. [STEP 3]

Are children who receive antibiotics in early childhood more likely to develop allergic disease?

EVIDENCE-BASED ANSWER

Probably so. Exposure to antibiotics before a patient is six months old is associated with increased risk of asthma, atopic dermatitis, rhinitis, anaphylaxis, and allergic conjunctivitis compared with those without history of antibiotic use; findings are similar in patients exposed to antibiotics before two years old. (SOR: **B**, large retrospective cohort and small, single-center, prospective, cohort study). Receiving at least one prescription for antibiotics before seven years is associated with increased risk of milk allergy, nonmilk allergy, and allergic rhinitis (SOR: **B**, large case-control study).

DOI 10.1097/EBP.0000000000000961

A 2018 Department of Defense Tricare beneficiaries retrospective cohort study (N=792,130) examined the association of antibiotics and treatment for acid reflux with allergic diseases.[1] Included patients were born between 2001 and 2013 with Tricare enrollment until at least 1 year of age and followed for 12 years. Infants requiring more than seven days hospitalization at birth or given an allergy diagnosis before age six months were excluded. Exposure was defined as a histamine-2 receptor antagonist, proton pump inhibitor, or antibiotic dispensed before six months old. The primary outcome measure was diagnosis of allergic disease, including food allergy, anaphylaxis, asthma, atopic dermatitis, allergic rhinitis, allergic conjunctivitis, urticaria, contact dermatitis, medication allergy, or other allergy at age six months or older based on International Classification of Diseases, 9th revision, Clinical Modification (*ICD-9*-CM) diagnostic codes derived from insurance claim files. Hazard ratios were adjusted for prematurity, cesarean delivery, sex, and other drug classes. The risk of all allergic diagnoses increased with the use of antibiotics before six months of age: asthma (adjusted hazard ratio [aHR], 2.1; 95% CI, 2.0–2.1), allergic rhinitis (aHR, 1.8; 95% CI, 1.7–1.8), anaphylaxis (aHR, 1.5; 95% CI, 1.4–1.7), and allergic conjunctivitis (aHR, 1.4; 95% CI, 1.3–1.5). This observational study is large but can only support association between exposure and outcome. Additionally, it is unclear if patients receiving antibiotics were only prescribed antibiotics or also received acid reflux treatment.

A 2017 single-center, Japanese, prospective birth cohort study (N=1,550) investigated the association between antibiotic exposure before two years old and development of allergic disease compared with those without antibiotic exposure.[2] Pregnant women were recruited during antenatal clinic visits. Patients were excluded if there were missing data or multiple births. Parents reported any antibiotic exposure history on a questionnaire when their child was two years old. Primary outcomes included a history of wheezing, asthma, rhinitis, or eczema in the past 12 months on parental questionnaire when the child was five years old. Exposure to antibiotics before two years old increased the risk of later development of asthma (adjusted odds ratio [aOR], 1.7; 95% CI, 1.1–2.7), atopic dermatitis (aOR, 1.4; 95% CI, 1.1–1.9), and allergic rhinitis (aOR, 1.6; 95% CI, 1.1–2.6) in children at the age of five years. This study included a single Japanese hospital, limiting its generalizability. Outcomes are

based on parent report without confirmatory chart documentation.

A 2017 case-control study (N=30,060) explored the association between early antibiotic use and development of food allergy and allergic disease among children up to seven years old from a health system in Pennsylvania.[3] Cases included a diagnosis of milk allergy after two months old (mean age of diagnosis 10 months) or nonmilk food or other allergies after three months old (mean age of diagnosis, two to three years). Cases were each matched to five patients without an allergy diagnosis by sex and age. To identify exposure, researchers counted the total number of outpatient, inpatient, and emergency department antibiotic orders. Patients with antibiotics ordered within 30 to 60 days of the date of allergy diagnosis were excluded. One or two antibiotic orders were associated with increased rate of milk allergy (odds ratio [OR], 1.5; 95% CI, 1.2–2.0), nonmilk food allergy (OR, 1.4; 95% CI, 1.1–1.8), and nonfood allergies (OR, 1.7; 95% CI, 1.5–1.9) compared with no antibiotic use. Three antibiotic orders were associated with a greater increased rate of milk allergy (OR, 2.4; 95% CI, 1.3–4.6), nonmilk food allergy (OR, 1.7; 95% CI, 1.3–2.1), and nonfood allergies (OR, 3.1; 95% CI 2.7–3.5). Limitations of this study included no confirmatory testing of allergy diagnosis, possibility of antibiotic orders going unfilled, and possibility of exposure to antibiotics from outside the system. EBP

Alex Katubig, MD
Carmen Boessen, MD
Nikole Cronk, PhD
Department of Family and Community Medicine
University of Missouri, Columbia, MO

The authors declare no conflicts of interest.

References

1. Mitre E, Susi A, Kropp L, Schwartz D, Gorman GH, Nylund CM. Association between use of acid-suppressive medications and antibiotics during infancy and allergic diseases in early childhood. *JAMA Pediatr*. 2018; 172(6):e180315. [STEP 3]
2. Yamamoto-Hanada K, Yang L, Narita M, Saito H, Ohya Y. Influence of antibiotic use in early childhood on asthma and allergic diseases at age 5. *Ann Allergy Asthma Immunol*. 2017; 119(1):54–58. [STEP 3]
3. Hirsch A, Pollak J, Glass T, et al. Early-life antibiotic use and subsequent diagnosis of food allergy and allergic diseases. *Clin Exp Allergy*. 2017; 47(2):236–244. [STEP 4]

Does a restrictive intravenous fluid strategy lead to improved outcomes in patients undergoing major abdominal surgery as compared with liberal intravenous fluids?

EVIDENCE-BASED ANSWER

In older or sicker patients, perioperative fluid restriction (resulting in zero net fluid balance) has no effect on disability-free survival but increases risk of acute kidney injury (SOR: **B**, large randomized controlled trial [RCT]). In a wider group of patients, restrictive fluid regimens may be associated with shorter time to flatus and hospital stay (SOR: **B**, meta-analysis of heterogeneous RCTs).

DOI 10.1097/EBP.0000000000001053

An international assessor-blinded 2018 randomized controlled trial (RCT) included 3,000 adults undergoing major abdominal surgery with increased risk of complications (defined as age ≥70 years old or having heart disease, diabetes, renal impairment, or morbid obesity).[1] Urgent or minor laparoscopic procedures were excluded. Patients received either a restrictive (net zero fluid balance with median of 1.7 L intraoperatively and additional 1.9 L in 24-hour postoperative period) or liberal (3.0 L during surgery plus an additional 3.0 L in 24-hour postoperative period) intravenous fluid regimen. The primary outcome was disability-free survival at one year, whereas secondary outcomes included acute kidney injury at 30 days, renal replacement therapy at 90 days, and a composite of septic complications, surgical site infection, or death. No differences were noted in disability-free survival at one year (82% vs 82%) or rates of septic complications or death (22% vs 20%). No differences were noted in renal replacement therapy at 90 days (0.9% vs 0.3%) or surgical site infections (17% vs 14%) after adjusting for multiple comparisons. The rate of acute kidney injury at

30 days was higher in the restrictive group compared with the liberal group (8.6% vs 5.0%; hazard risk 1.7; 95% CI, 1.3–2.3; number needed to harm 28).

A 2017 meta-analysis compared postoperative morbidity, recovery, and length of hospital stay for restrictive versus liberal perioperative fluid therapies for adults undergoing major abdominal surgery (13 RCTs; N=1,052).[2] Surgical procedures varied and included intestinal or colorectal procedures, pancreatobiliary procedures, and abdominal aortic repair or bypass. The definition of "restrictive" varied by the study and included total perioperative fluid input of 1.4 to 6.0 L, while "liberal" or "conventional" regimens included input of 1.6 to 6.6 L. Primary outcome was rate of total postoperative complications (specific complications and follow-up were not defined). Secondary outcomes included time to flatus and length of hospital stay. No difference was noted in rate of total postoperative complications (11 trials, n=932; pooled odds ratio 0.59; 95% CI, 0.34–1.04). Patients who received the restricted regimen did have shorter time to flatus (six trials, n=345; pooled difference in the mean=–0.67; 95% CI, –1.3 to –0.06) and shorter hospital stay (eight trials, n=566; pooled difference in the mean=–1.5; 95% CI, –2.9 to –0.1) than patients who received the liberal regimen. Significant heterogeneity was observed in the type of surgery performed, fluid protocol used, and total perioperative fluid input. All studies were judged of good quality, but several studies were limited by lack of clear blinding of investigators and patients. EBP

Kane Laks, MD
Seth Freeman, MD
University of Missouri School of Medicine
Columbia, MO

The authors declare no conflicts of interest.

References

1. Myles PS, Bellomo R, Corcoran T, et al. Restrictive versus liberal fluid therapy for major abdominal surgery. *N Engl J Med*. 2018; 378(24):2263–2274. [STEP 2]
2. Jia FJ, Yan QY, Sun Q, et al. Liberal versus restrictive fluid management in abdominal surgery: a meta-analysis. *Surg Today*. 2017; 47(3):344–356. [STEP 1]

Are cyclical progestogens effective in treating heavy menstrual bleeding?

EVIDENCE-BASED ANSWER

Cyclical progestogens, taken as a short or long course, reduce menstrual blood loss compared to baseline but are less effective than other medical treatments for heavy menstrual bleeding (SOR: **A**, meta-analysis of randomized controlled trials [RCTs]). Long course progesterone therapy may be more effective than short course (SOR: **B**, systematic review of RCTs and one nonrandomized study with significant limitations).

DOI 10.1097/EBP.0000000000001074

A 2019 Cochrane review of 15 randomized controlled trials (RCTs) (N=1,071) compared treatments for heavy menstrual bleeding, defined as excessive blood loss interfering with quality of life.[1] Bleeding was assessed by the alkaline hematin method (five studies, n=210), pictorial blood assessment chart (PBAC) (eight studies, n=786), or women's subjective assessment of blood loss (two studies, n=70). A subset of six trials (n=145) compared short course norethisterone (5 mg, two or three times daily, for eight or 11 days) to tranexamic acid, danazol, or a levonorgestrel intrauterine device. The trial duration ranged from two cycles to six months. Short course progesterone was less effective than other medical treatment options for reduction of blood loss (mean difference [MD] 37.3 mL per cycle; 95% CI, 17.7–56.9). A subset of four trials (n=355) pooled data comparing long course medroxyprogesterone acetate (5 mg or 10 mg, twice daily on days 5–26 or 5–25 of the menstrual cycle) or long course norethisterone (5 mg three times daily, on days 5–26 of the menstrual cycle) to tranexamic acid 500 mg, four times per day on first five days of menses, ormeloxifene 60 mg twice per week, or a levonorgestrel intrauterine device. Trial duration ranged from three to nine months. Long course progesterone was less effective than other medical treatment options for reduction of blood loss (MD 17 points on the PBAC;

95% CI, 10.9–22.8). No RCTs compared progestogen treatment with placebo. Limitations included heterogeneity of comparison treatments.

A 2015 systematic review (91 clinical trials and observational studies; N=5,929) included a subset of 11 RCTs (n=361) and one nonrandomized trial (n=6) assessing the impact of short course (≤2 weeks or ≤14 days per cycle) and long course (≥3 weeks or ≥21 days per cycle) oral progestogens on heavy menstrual bleeding.[2] Study drugs included norethisterone 5 mg, twice or three times per day, and medroxyprogesterone acetate 5 or 10 mg daily or twice daily. Comparison drugs included danazol (200 mg per day), progesterone intrauterine system (60 μg progesterone daily), mefenamic acid (500 mg or 1 g, four times per day), ormeloxifene (60 mg, three times per day, two days per week), vaginal ring, and levonorgestrel-releasing intrauterine system. Heavy menstrual bleeding was defined as interfering with quality of life by self-report or blood loss of at least 80 mL per cycle. No *P* values nor CIs were reported.

In a subset of four trials (n=157), short course oral progestogens showed a median reduction in blood loss of 2% to 30% relative to baseline for up to six months of study. One study (n=21) of short course norethisterone 5 mg twice daily showed a mean reduction in blood loss of 20% over two cycles of treatment. One study (n=11) of short course 10 mg medroxyprogesterone daily showed a blood loss reduction of 25% after one month of treatment, 41% reduction after two months, but 12% mean increase after three months of treatment assessed using the alkaline hematin method and pictorial blood loss assessment chart.

Data were not pooled for long course progesterone. In a single study of long course norethisterone acetate 5 mg three times daily compared to the levonorgestrel-releasing intrauterine system (n=44), long course progesterone reduced median blood loss from baseline by 63% during cycle one and 78% during cycle three of treatment. In another small study of long course progesterones comparing medroxyprogesterone acetate 10 mg twice daily to norethisterone acetate 5 mg three times daily (n=5), long course methods resulted in mean blood loss reductions between 32% and 37% over baseline. No studies compared progestogens to placebo. Limitations include small studies, lack of reported measures of effect, heterogeneity of bleeding definitions, treatment regimens, and study duration. EBP

Hillary Lee, MD
Nikole Cronk, PhD
MO—University of Missouri FP—Columbia
Columbia, MO

The authors declare no conflicts of interest.

References

1. Bofill Rodriguez M, Lethaby A, Low C, Cameron IT. Cyclical progestogens for heavy menstrual bleeding. *Cochrane Database Syst Rev.* 2019; 8:CD001016. [STEP 1]
2. Bitzer J, Heikinheimo O, Nelson A, Calaf-Alsina J, Fraser I. Medical management of heavy menstrual bleeding: a comprehensive review of the literature. *Obstet Gynecol Surv*. 2015; 70(2):115–130. [STEP 1]

What are the harms associated with decreased sleep duration in adults?

EVIDENCE-BASED ANSWER

Those who routinely slept ≤6 hours per day were 6% to 18% more likely to develop type-2 diabetes mellitus or impaired fasting glucose, whereas those who slept <5 hours per day were 48% more likely (SOR: **A**, meta-analysis of cohort studies and a cohort study).

DOI 10.1097/EBP.0000000000000908

A 2016 meta-analysis of 36 cohort studies (N=1,061,555) determined the risk of developing type-2 diabetes in adults with different types of sleep disturbances.[1] Of the 36 studies, 14 studies (N=583,263) looked at sleep duration. Patients included were approximately 42% male, body mass index (BMI) when reported was 24 to 33 kg/m^2 with the majority less than 30 kg/m^2, and mean ages ranged from 36 to 70 years (overall mean, 62 years). Sleep duration was determined through self-report and was categorized into short (≤5 h/d or 6 h/d), normal (seven to eight h/d), and long (≥9 h/d) durations. Follow-up was from two to 32 years. The primary outcome was the incidence of diabetes. The diagnosis was through self-report, medical chart review, or using blood tests (fasting plasma glucose [FPG], oral glucose tolerance test,

or HbA1c). The risk of developing diabetes was increased for those with sleep durations of <5 h/d and six h/d (14 studies, n=not provided; relative risk [RR], 1.5; 95% CI, 1.3–1.8; I^2=81% and 10 studies, n=not provided; RR, 1.2; 95% CI, 1.1–1.3; I^2=55%). Two of the 14 studies used slightly different sleep parameters (normal duration was set at seven h/d and six to eight h/day), although the pooled RR were similar when these two studies were excluded.

A 2017 prospective cohort study including 162,121 adults of 20 to 80 years evaluated the effect of sleep duration on healthy adults.[2] Patients were 47% male with a BMI of 22 to 23 kg/m^2. Patients had no major diseases as determined by a baseline medical screening (obesity, impaired fasting glucose/diabetes, hypertension, thyroid disease, among others). All patients visited the MJ Health Management Institute in Taiwan periodically for medical screening throughout the study. The number of visits ranged from two to 19 with 99% going annually. Using self-reported sleep duration, patients were separated into three groups (19% with <6 h/d, 73% with six to eight h/d, and 8.6% with >8 h/d). At baseline, no significant differences were observed in demographic or cardiovascular risk factors among the three groups. The primary outcome was metabolic syndrome, including impaired fasting glucose and diabetes. Impaired fasting glucose was defined as FPG of 100 to 125 mg/dL, and diabetes defined as FPG of ≥125 mg/dL. At the end of the 18-year long study, it was found that those in the <6 h/d group were more likely to develop impaired fasting glucose and diabetes (adjusted hazard ratio, 1.1; 95% CI, 1.03–1.1). There were multiple limitations to this study, including sleep duration obtained through self-report rather than through actigraphy or polysomnography. Insomnia was assessed through a single question that was beneficial for a large screening but did not quantify the severity of the insomnia. Potential confounders were not addressed such as variability in weekday-weekend sleep, contraceptive use, hormonal replacement therapy, household income, or number of children in the household. EBP

Matt McKee, DO
Sarah Daly, DO FAAFP
Utah Valley FMR
Provo, UT

The authors declare no conflicts of interest.

References

1. Thunyarat A, Sirimon R, Eve Van C, et al. Sleep disturbances compared to traditional risk factors for diabetes development: systematic review and meta-analysis. *Sleep Med Rev.* 2016; 30, 11–24. [STEP 2]
2. Han-Bing D, Tony T, Benny Chung Y, et al. Short sleep duration increases metabolic impact in healthy adults: a population-based cohort study. *Sleep*. 2017: 40(10). [STEP 3]

Does *Pelargonium sidoides* root extract shorten duration and lessen severity of the common cold?

EVIDENCE-BASED ANSWER

Adults with the common cold treated with *Pelargonium sidoides* as compared with placebo are 59% less likely to fail to recover by day 10 as compared with placebo. No difference was found in failure to recover rates at day 5 (SOR: **B**, meta-analysis of randomized controlled trials [RCTs]). Clinical cure and major improvement may also more likely at 10 days with *P sidoides* than with placebo (SOR: **B**, RCT).

DOI 10.1097/EBP.0000000000000909

A 2013 meta-analysis of eight randomized controlled trials (RCTs) (N=1,771) examined the efficacy of *Pelargonium sidoides* (an African geranium) in acute respiratory infections (acute bronchitis, acute sinusitis, or the common cold).[1] Patients had one or more acute respiratory symptoms (<48 hours to symptom onset) and did not require antibiotic therapy. Only one study evaluated patients with cold symptoms. This study included 103 adult patients, 18 to 55 years old, with cold symptoms for 24 to 48 hours. The patients were randomized to receive either 30 drops (1.5 mL) of the liquid herbal drug preparation of *P sidoides* EP or placebo three times a day for 10 days. The primary outcome was failure to recover by day 5 and day 10. *Pelargonium sidoides* compared with placebo showed no difference in rates of failure to recover by day five, but there was less failure to recover by day 10 (one RCT, N=103; risk ratio [RR], 0.96; 95% CI, 0.9–1.03 and RR, 0.41; 95% CI, 0.29–0.6). There was a slightly increased risk for minimal adverse events (gastrointestinal complaints such as nausea, vomiting, diarrhea or heartburn, allergic skin reactions with pruritus and urticaria) in the treatment group compared with placebo (eight RCT, N=1,771; RR, 1.3; 95% CI, 1.04–1.7). None of the reported events were

considered serious. There were some concerns about the effectiveness of the blinding, the use of unvalidated scores for outcome assessment, minor attrition problems, and, in one instance, selective reporting.

A 2018 RCT (N=207) examined the efficacy of *P sidoides* in adults of 18 to 55 years old (average 35 years and 78% female) with cold symptoms (cough, headache, nasal discharge or congestion, sneezing, sore throat, scratchy throat, hoarseness, muscle aches, or fever) for 24 to 48 hours.[2] Patients in whom antibiotic or other specific therapy was recommended, such as streptococcal infections, pneumonia, diphtheria, tuberculosis, infections in immunocompromised or elderly persons, any life-threatening or chronic condition (asthma, COPD, etc), were excluded. Treatments included standard dose (3 × 30 drops) or high dose (3 × 60 drops) of *P sidoides* extract per day for 10 days as compared with placebo. In the case of a fever>39°C, paracetamol tablets were allowed. Clinical cure was defined as a complete resolution of all cold symptoms with cold intensity score of zero points or complete resolution of all but one cold symptom. The cold intensity score is a 40-point verbal rating scale of cold symptom severity where higher scores indicate more severe symptoms. After 10 days, more patients receiving active treatment were clinically cured compared with placebo (90% vs 21%; $P<.0001$). Complete recovery or major improvement was significantly better at day 5 for the active treatment group compared with the control group (71% vs 9.6%; $P<.0001$). Mild-to-moderate adverse events—all nonserious—occurred in 15% of those receiving active treatment versus in 5.8% for the control group (*P* value not reported). These adverse events included epistaxis, mild epigastric discomfort, and abdominal pain. Study authors supported by pharmaceutical industry. EBP

Benz Gulati Pimsakul, MD
Sarah Daly, DO
Utah Valley Family Medicine Residency, Provo, UT

The authors declare no conflicts of interest.

References

1. Timmer A, Günther J, Motschall E, Rücker G, Antes G, Kern WV. Pelargonium sidoides extract for treating acute respiratory tract infections. *Cochrane Database Syst Rev.* 2013; 10:CD006323. [STEP 1]
2. Riley DS, Lizogub VG, Zimmermann A, Funk P, Lehmacher W. Efficacy and tolerability of high-dose Pelargonium extract in patients with the common cold. *Altern Therapies Health Med*. 2018; 24(2):16–26. [STEP 2]

Are stimulant medications effective for treating adults with major depressive disorder?

EVIDENCE-BASED ANSWER

Yes, in certain situations. Psychostimulants used as adjunct therapy, but not monotherapy, relieve clinical symptoms by 50% in patients with major depressive disorder or bipolar depression (SOR: **A**, meta-analysis of randomized controlled trials [RCTs]). In subgroup analyses, patients with major depressive disorder treated with armodafinil/modafinil or dextroamphetamine showed clinical improvement but not with lisdexamfetamine or methylphenidate (SOR: **A**, meta-analysis of RCTs). Low-dose stimulants may be considered as adjunct therapy for depressed patients with suboptimal response to treatment (SOR: **C**, consensus guideline). The Food and Drug Administration does not approve armodafinil, modafinil, amphetamine, dextroamphetamine, lisdexamfetamine, or methylphenidate for the treatment of depression either as monotherapy or as adjunct treatment.

DOI 10.1097/EBP.0000000000001045

A 2017 meta-analysis of 21 randomized controlled trials evaluated the efficacy of psychostimulants as adjunct or monotherapy in adults with unipolar or bipolar depression (N=3,713).[1] Patients, adults ≥18 years old (61% female, average age 43 and 44 years old), were clinically diagnosed with major depressive disorder or bipolar depression, as defined by the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition. The treatment group received psychostimulants as adjunct or monotherapy with armodafinil, modafinil, amphetamine, dextroamphetamine, lisdexamfetamine, or methylphenidate (doses not provided) compared with placebo for two to 16 weeks. Response to treatment was defined as a >50% improvement on the Hamilton Depression Rating Scale, the Inventory of Depressive Symptomatology Clinician-Rated 30-Item scale, or the Montgomery-Asberg Depression Rating Scale. Results of the meta-analysis showed greater response to treatment with psychostimulants compared with placebo in patients with major depressive disorder and bipolar depression

(15 studies, N=2,047; odds ratio [OR] 1.4; 95% CI, 1.1–1.8; I^2=19%; and six studies, N=1,628; OR 1.4; 95% CI, 1.1–1.8; I^2=11%). Depression symptoms clinically improved with psychostimulant agents as adjunct therapy compared with placebo but not as monotherapy (17 studies, N=3,550; OR 1.4; 95% CI, 1.2–1.6; and four studies, N=125; OR 2.3; 95% CI, 0.67–7.5). Based on subgroup analysis, clinical improvement was noted in depression with armodafinil/modafinil or dextroamphetamine over placebo (10 studies, N=2,190; OR 1.5; 95% CI, 1.2–1.8; I^2=15 and one study, N=22; OR 7.1; 95% CI, 1.1–46). However, no difference was found between lisdexamfetamine or methylphenidate and placebo (four studies, N=1,020; OR 1.2; 95% CI, 0.94–1.6; I^2=0; and seven studies, N=443; OR 1.5; 95% CI, 0.88–2.5; I^2=13). Limitations included possible overestimation of effects because of small sample sizes and confounding data from patients with concurrent substance use disorders and medical disorders (brain injury, cancer, ADHD). Authors of the meta-analysis were supported by multiple medical pharmaceutical companies.

The 2010 American Psychiatric Association practice guidelines for the treatment of major depression discussed the use of stimulants as adjunct therapy for major depression.[2] Based on a few clinical trials and case reports, the American Psychiatric Association stated that stimulants may help ameliorate otherwise suboptimal response to therapy (strength of recommendation not provided). Several practice guideline panel members reported consulting, research, or speaking for multiple pharmaceutical companies. This guideline was reviewed by an independent review panel without conflicts of interest. EBP

Blake Taylor, DO
Sarah Daly, DO
Utah Valley FMR, Provo, UT

The authors declare no conflicts of interest.

References

1. McIntyre RS, Lee Y, Zhou AJ, et al. The efficacy of psychostimulants in major depressive episodes: a systematic review and meta-analysis. *J Clin Psychopharmacol.* 2017; 37(4):412–418. [STEP 1]
2. Practice guideline for the treatment of patients with major depressive disorder (revision). American Psychiatric Association. *Am J Psychiatry.* 2000; 157(4 suppl):1–45. [STEP 5]

In athletes with a history of low-grade ankle sprains, does long-term bracing prevent future ankle sprains?

EVIDENCE-BASED ANSWER

Yes. External bracing reduces ankle sprain recurrence by up to 63% (SOR: **A**, meta-analysis of randomized controlled trials). External bracing can reduce ankle sprain recurrence by up to 47% compared with neuromuscular proprioceptive exercises (SOR: **B**, single randomized control trial). Ankle braces are effective for secondary prevention in athletes who play football, volleyball, soccer, and basketball (SOR: **A**, meta-analysis of randomized control trials).

DOI 10.1097/EBP.0000000000001004

A 2018 systematic review and meta-analysis of five randomized controlled trials (RCTs) (N=6,725) studied the efficacy of ankle bracing for primary and secondary prevention of ankle injury in athletes.[1] Patients in the study were male and female athletes, 16 to 26 years old who played volleyball, basketball, football, and soccer. All RCTs compared bracing during sports to no interventions. Ankle injuries were defined as ankle sprains, syndesmotic sprains, or fractures. Braces studied included nonrigid, semi-rigid, rigid, and lace-up ankle braces. The most common braces used were semi-rigid braces and lace-up braces. Patients were asked to wear the brace for all practices and games during the follow-up period (ranging from one season to two years). The study collected data on secondary injury (ankle sprain or fracture) from patient report, physician report, and athletic trainer report. Braces were effective for secondary prevention of acute ankle injuries (risk ratio [RR] 0.37; 95% CI, 0.24–0.58) with a number needed to treat of 12 (95% CI, 10–18). No differences between brace types were observed, but the data was limited on this subject. No significant adverse events in the bracing group were reported. Study limitations included inability to blind the treatment group.

A 2014 three-armed RCT (N=384) not included in the above meta-analysis compared different modalities

for preventing ankle sprain recurrence for 12 months after lateral ankle sprains.[2] Study participants were self-identified athletes 18 to 70 years old (52% male), who suffered a lateral ankle sprain within two months of study inclusion. Participants were randomly assigned to bracing, neuromuscular training, or combined groups. The study controlled for age, education, high-risk sports, previous ankle injury, ankle sprain grade, and chronic ankle instability. Patients in the bracing group received a semi-rigid Aircast® brace to be worn during all sports activities for 12 months. The neuromuscular training group underwent an eight-week at-home program with a balance board and increasingly challenging balance exercises. The combination group received both the exercise program and the brace. The primary outcome was incidence of inversion ankle sprains, sensations of giving way, or severe injuries requiring time off from activity evaluated via patient-reported monthly questionnaires for 12 months. Bracing was found superior to neuromuscular training (15% vs 27% ankle sprain recurrence; RR 0.53; 95% CI, 0.29–0.97). Bracing and neuromuscular training in combination was not statistically significant compared with training alone (19% vs 27% ankle sprain recurrence; RR 0.71; 95% CI, 0.41–1.23). Severity of the recurrent sprains was the same in all treatment groups. No adverse outcomes were reported in this study. Study limitations included lack of a control group, low patient compliance in all three groups, and reliance on patient-reported symptoms without physician validation. EBP

Keyur Desai, MD
Anne Marie Zeller, DO, MSc
University of Chicago (NorthShore) FMRP
Glenview, IL

The authors declare no conflicts of interest.

References

1. Barelds I, van der Broek A, Huisstede B. Ankle bracing is effective for primary and secondary prevention of acute ankle injuries in athletes: a systematic review and meta-analyses. *Sports Med*. 2018; 48(12):2775–2784. [STEP 1]
2. Janssen KW, Mechelen WV, Verhagen EALM. Bracing superior to neuromuscular training for the prevention of self-reported recurrent ankle sprains: a three-arm randomised controlled trial. *Br J Sports Med*. 2014; 48(16):1235–1239. [STEP 2]

For patients with EPL, is the addition of mifepristone to vaginal misoprostol more effective than medication management with misoprostol alone?

EVIDENCE-BASED ANSWER

For patients with first trimester early pregnancy loss, pretreatment with mifepristone before misoprostol increases treatment success more than vaginal misoprostol alone, with a number needed to treat between 3.5 and 6 (SOR: **A**, randomized controlled trials [RCTs]). Patients who receive pretreatment with mifepristone are more likely to have gestational sac expulsion on follow-up ultrasonography (SOR: **A**, RCTs) and may be less likely to need surgical intervention (SOR: **B**, RCTs with inconsistent results).

DOI 10.1097/EBP.0000000000001016

A 2018 USA randomized controlled trial (RCT) assessed women with early pregnancy loss (EPL) to determine if pretreatment with oral mifepristone before vaginal misoprostol improved the rate of successful gestational sac expulsion over misoprostol alone.[1] The study included 300 women (mean age 30 years old) diagnosed with EPL between five and 12 weeks' gestation. It excluded women with a viable or ectopic pregnancy, an open cervical os, an absent gestational sac on ultrasonography, a hemoglobin less than 9.5 g/dL, any contraindication to study medications, or an intrauterine device in place. The pretreatment group (N=149) received 200 mg oral mifepristone (directly observed), whereas the misoprostol-alone group (N=151) received no medication. Women in both groups self-administered 800 μg of vaginal misoprostol at home 24 hours later. At follow-up a median of two to three days (range 0.5–9.6 days) after treatment, an investigator blinded to the therapy performed transvaginal ultrasonography on each woman to look for a gestational sac. Women with gestational sacs were offered a second dose of misoprostol, uterine aspiration, or expectant management, with subsequent follow-up and ultrasonography one week later. Treatment success (absence of a gestational sac at the initial follow-up visit

and no additional interventions needed within 30 days) was significantly higher for the mifepristone-pretreatment group (83.8%) compared with the misoprostol-alone group (67.1%), with a relative risk (RR) of 1.3 (95% CI, 1.1–1.4; number needed to treat [NNT]=6). Women in the mifepristone-pretreatment group were also less likely than those in the misoprostol-alone group to undergo uterine aspiration within 30 days of treatment (8.8% vs 23.5%, respectively; RR 0.37; 95% CI, 0.2–0.7; NNT=6.8). Study limitations included the lack of blinding of the participants, variable time between treatment and the initial follow-up visit, and lack of supervision of the misoprostol administration.

A 2018 double-blind RCT in India assessed the efficacy of pretreatment with mifepristone before misoprostol compared with misoprostol alone for the management of EPL.[2] Participants included 92 women (mean age 25 years old; median parity 1) diagnosed with EPL at or before 12 weeks' gestation. Exclusion criteria included incomplete or inevitable abortion, hemodynamic instability, bleeding disorder, infection, hemoglobin less than 8 g/dL, or a contraindication to the study medications. Patients in the pretreatment group received 200 mg of oral mifepristone, while those in the misoprostol-only group received placebo. Both groups were admitted to the hospital 48 hours later and received 800 μg of vaginal misoprostol. If no expulsion of the gestational sac occurred within four hours, the investigators gave patients oral misoprostol 400 mcg every three hours (maximum two to four doses, depending on gestational age). The primary outcome was completed abortion (absence of vaginal bleeding and a well-defined endometrial line with maximum thickness of <15 mm on transvaginal ultrasonography on day 14); patients who did not achieve this underwent surgical evacuation. Patients who received pretreatment with mifepristone had a success rate of 86.7% compared with 57.8% in the control group (RR 1.5; 95% CI, 1.1–2, NNT=3.5). Findings were not affected by patient age, parity, previous abortions, gestational age, clinical symptoms, or gestational sac size. The study was limited by the need for hospital admission, which could be impractical for many women with EPL.

A 2019 meta-analysis[3] included the two studies detailed above as well as a 2009 RCT[4] from China of limited quality. In the 2009 study, 50% of women in the misoprostol-alone group required emergent dilation and curettage for hemorrhage and were subsequently excluded from analysis. The study was limited by unclear allocation concealment, lack of information on blinding of investigators and participants, and absence of details on outcomes of the 15 patients who were excluded from the analysis. In contrast to the two 2018 studies detailed above, the 2009 study did not find increased efficacy with the addition of mifepristone to misoprostol for management of EPL. The 2019 meta-analysis[3] of the three studies also did not find increased efficacy with the addition of mifepristone to misoprostol, with significant heterogeneity (I^2=70.55%).

Mifepristone is not readily available in the United States because of a Risk Evaluation and Mitigation Strategy required by the Food and Drug Administration. The American College of Obstetricians and Gynecologists and American Academy of Family Physicians[5] both state that such status is no longer necessary. EBP

Jen Flament, MD
Claire Thomson, MD
Kristin Nierenberg, MD
Swedish Cherry Hill Family Medicine Residency
Seattle, WA

The authors declare no conflicts of interest.

References

1. Schreiber CA, Sonalkar S, Barnhart KT, Ratcliffe SJ, Creinin MD, Atrio J. Mifepristone pretreatment for the medical management of early pregnancy loss. *New Engl J Med.* 2018; 378(23):2161–2170. [STEP 2]
2. Sinha P, Suneja A, Guleria K, Aggarwal R, Vaid NB. Comparison of mifepristone followed by misoprostol with misoprostol alone for treatment of early pregnancy failure: a randomized double-blind placebo-controlled trial. *J Obstet Gynecol India.* 2018; 68(1): 39–44. [STEP 2]
3. Lemmers M, Verschoor MAC, Kim BV, et al. Medical treatment for early fetal death (less than 24 weeks). *Cochrane Database Syst Rev.* 2019; 6:CD002253. [STEP 1]
4. Fang AH, Chen QF, Zheng W, Li YH, Chen RY. Termination of missed abortion in a combined procedure: a randomized controlled trial. *J Reprod Contraception.* 2009; 20(1): 45–49. [STEP 3]
5. American Academy of Family Physicians. *Removing Risk Evaluation and Mitigation Strategy (REMS) Categorization on Mifepristone: AAFP Resolution No. 506 (Co-Sponsored C)*. 2018. https://www.aafp.org/about/governance/congress-delegates/2018/resolutions2/cosponsored-c.mem.html. Approved April 5-9, 2019. Accessed on December 6, 2019.

In adults with depression, does medication therapy directed by pharmacogenetic testing lead to better outcomes than nondirected medication therapy?

EVIDENCE-BASED ANSWER

Yes. Pharmacogenetic-guided treatment of major depressive disorder in adults results in 36% better response rates and more than 70% better remission rates compared with unguided treatment (SOR: **B**, meta-analyses using heterogenous randomized control trials and cohort studies).

DOI 10.1097/EBP.0000000000000925

A 2019 meta-analysis that included five randomized control trials (RCTs) (n=1,737) evaluated the effectiveness of pharmacogenetic-guided treatment of major depressive disorder compared with usual treatment.[1] All studies enrolled patients with a diagnosis of major depressive disorder by *Diagnostic and Statistical Manual of Mental Disorders*, Fifth Edition (DSM-5) criteria; two trials reported baseline Hamilton Depression Rating Scale-17 (HDRS-17) score of more than 14 and 18 (range, 0–50). One RCT included only patients who had inadequate response or intolerance to at least one psychotropic treatment; the other trials included a mixture of patients new to treatment or who had inadequate response to previous treatment. All studies excluded patients with other active psychiatric disorders. The patients were blinded in all included trials. Used pharmacogenetic tools included Genesight®, CNSDose®, Neuropharmagen®, and NeurolDgenetix®. The number of genes used to guide treatment varied from five to 30, with all tests including CYP2D6 and CYP2C19 as there are dosing guidelines for antidepressant therapy based on these genes. Results of pharmacogenetic testing was reported by categorizing medications into bins as green (use as directed), yellow (use with caution), or red (use with increased caution and more monitoring). The method for determining the specific pharmacotherapy agent was not reported in this meta-analysis. Treatment as usual was generally defined as the use of a trial and error approach until effective treatment that provides full remission is achieved. Remission of depression was defined as HDRS-17 score of seven or less. After eight to 12 weeks, participants who received pharmacogenesis-guided therapy were 71% more likely to achieve remission compared with unguided therapy (relative risk [RR], 1.71; 95% CI, 1.2–2.5). Limitations included the small number of RCTs that fit inclusion criteria and moderate heterogeneity (I^2=71). Recruitment and industry biases were noted based on physician referral of participants and industry sponsorship of RCTs, respectively.

A 2018 meta-analysis that included four blinded RCTs and two unblinded cohort studies (N=1,329) also evaluated the effectiveness of pharmacogenesis-guided treatment of major depressive disorder.[2] The four RCTs were also included in the above meta-analysis; however, this meta-analysis included unblinded studies that allows for the evaluation of the role of expectancy bias in patients who knowingly receive pharmacogenesis-guided treatment. Furthermore, the inclusion of response rates in addition to remission rates provided additional clinically relevant data. The cohort studies included adults diagnosed with major depressive disorder by DSM-IV criteria and HDRS-17 score greater than 14 who were nonrandomly allocated into Genesight-guided or unguided therapy groups. Four studies (n=799) evaluated response to treatment defined as decrease in HDRS-17 score by 50%, and five studies (n=735) evaluated remission defined as HDRS-17 score less than eight. Pooled analysis showed that participants who received pharmacogenesis-guided therapy were 36% more likely to respond (RR, 1.36; 95% CI, 1.1–1.6; number needed to treat [NNT]=7; I^2=9%) and 74% more likely to achieve remission (RR, 1.74; 95% CI, 1.1–2.8; NNT=7; I^2=72%) compared with unguided therapy. Subgroup analysis based on blinded or unblinded status revealed that the response rate in the unblinded cohort studies was 40% in guided therapy and 23% in unguided therapy, whereas the response rate in the blinded RCTs was 53% in guided therapy and 41% in unguided therapy. Although statistical analysis of differences was not reported, the response rate in unguided therapy was notably lower in the unblinded cohort studies compared with the blinded RCTs. This suggests expectancy bias, as patients who were aware that their therapy was unguided may have been more likely to report poor response. Limitations included a small number of studies that fit inclusion criteria, incorporation of nonrandomized studies, and industry bias. Moderate heterogeneity was found in the

evaluation of remission rates likely due to differing pharmacogenetic testing modalities between studies. EBP

John S. Anderson, DO
Thomas J. Satre, MD
University of Minnesota/St. Cloud Family Medicine Residency, St. Cloud, MN

The authors declare no conflicts of interest.

References

1. Bousman CA, Arandjelovic K, Mancusa SG, Eyre HA, Dunlop BW. Pharmacogenetic tests and depressive symptom remission: a meta-analysis of randomized controlled trials. *Pharmacogenomics*. 2019; 20(01):37–47. [STEP 1]
2. Rosenblat JD, Yena L, Mcintyre RS. The effect of pharmacogenomic testing on response and remission rates in the acute treatment of major depressive disorder: a meta-analysis. *J Affective Disord.* 2018; 241:484–491. [STEP 2]

Do vaccine reminders increase vaccination rates in children and adolescents?

EVIDENCE-BASED ANSWER

Vaccination reminders increase vaccination rates in children (number needed to treat [NNT]=14) and adolescents (NNT=14) (SOR **A**: systematic reviews of randomized controlled trials [RCTs]). Telephone or autodialed phone messages, postcards, and text messages increased the vaccination rates by 16% to 127% (SOR **A**: systematic reviews of RCTs). E-mail reminders increase influenza vaccination rates in adolescents (NNT=11) (SOR **B**: single-blinded, parallel-group study). Both the American Association of Pediatrics and the Task Force on Community Preventive Services strongly recommend the use of recall interventions to improve vaccination rates in children and adolescents (SOR **C**: expert opinion).

DOI 10.1097/EBP.0000000000000850

A 2018 meta-analysis of 38 studies (37 randomized controlled trials [RCTs] and one observational study, N=71,232) examined the effectiveness of reminder and recall interventions to increase vaccination rates in children.[1] The trials included children from birth to 18 years who received immunizations in any setting including academic or nonacademic, developed or developing countries. The intervention groups received reminder and recall messages via telephone, letter, postcard, text, electronic telephone calls, online portals, or face-to-face reminders outside of a clinic visit. The frequency of these interventions ranged from only one reminder or recall notification to messages delivered intermittently over approximately one year. Controls received usual care. The primary outcome studied was the receipt of immunizations as reported in individual studies (either by a certain date or age) as measured over four weeks to 24 months. Patient reminders were associated with an increased rate in childhood immunizations (23 RCTs, N=31,099; risk ratio [RR], 1.2; 95% CI, 1.2–1.3; number needed to treat [NNT]=14), childhood influenza immunizations (five RCTs, N=9,265; RR, 1.5; 95% CI, 1.1–2.0; NNT=5), and adolescent immunizations (nine RCTs and one observational study, N=30,868; RR, 1.3; 95% CI, 1.2–1.4; NNT=14). Subanalyses found that telephone interventions increased vaccination rates in children (two RCTs, N=234; RR, 2.3; 95% CI, 1.1–4.4; NNT=2) and adolescents (one RCT, N=418; RR, 2.0; 95% CI, 1.1–4.0; NNT=4). In children, postcards resulted in higher rates of vaccinations (four RCTs, N=2,806; RR, 1.2; 95% CI, 1.1–1.5; NNT=13) as did autodialer phone messages (three RCTs, N=8,583; RR, 1.3; 95% CI, 1.2–1.4; NNT=11) and letters (nine RCTs, N=13,009; RR, 1.2; 95% CI, 1.1–1.3; NNT=19). Text messages also increased vaccination rates in children (one RCT, N=304; RR, 1.2; 95% CI, 1.1–1.3, NNT=14) and adolescents (three RCTs and one observational study, N=7,264; RR, 1.4; 95% CI, 1.2–1.6; NNT=11). Limitations included heterogeneity in study populations, interventions, settings, as well as type and number of vaccinations being targeted.

A 2017 single-blinded, parallel-group study (N=3,545) explored the effectiveness of e-mail reminders on influenza vaccine rates.[2] The study included adolescents (11–17 years old) in four US clinics (21.7% enrolled in Medicaid) and excluded those who had a record of having received the influenza vaccine that season or a waiver refusing the vaccine.

Patients either received two to three monthly or bimonthly e-mail reminders regarding all due immunizations during each of two influenza seasons (N=1,976) or received usual care (N=1,569). Adolescents who received e-mail reminders were 28% more likely to receive the flu vaccine compared with those who did not (adjusted odds ratio, 1.3; 95% CI, 1.1–1.5; NNT=11). Limitations include siblings randomized to different study arms, the study being limited to patients with e-mail addresses, and the reminder e-mails listed all immunizations that were due, not just influenza.

A 2010 evidence-based practice guideline from the American Association of Pediatrics strongly recommended vaccination reminder-recall systems based on a systematic review–level evidence.[3] Earlier (2000) recommendations by the Task Force on Community Preventive Services cited "strong scientific evidence" in recommending the use of reminder-recall interventions to improve vaccination coverage in children and adolescents.[4] EBP

Eve Barker, MD
Janelle Guirguis-Blake, MD
Tacoma Family Medicine Residency, Tacoma, WA

The authors declare no conflicts of interest.

References

1. Jacobson Vann JC, Jacobson RM, Coyne-Beasley T, Asafu-Adjei JK, Szilagyi PG. Patient reminder and recall interventions to improve immunization rates. *Cochrane Database Syst Rev.* 2018; 1:CD003941. [STEP 1]
2. Dombkowski KJ, Cowan AE, Reeves SL, Foley MR, Dempsey AF. The impacts of email reminder/recall on adolescent influenza vaccination. *Vaccine.* 2017; 35(23): 3089–3095. [STEP 2]
3. Hammer LD, Curry ES, Harlor AD, et al. Increasing immunization coverage. *Pediatrics*. 2010; 124(8):1295–1304. [STEP 1]
4. Task Force on Community Preventive Services. Recommendations regarding interventions to improve vaccination coverage in children, adolescents, and adults. *Am J Prev Med.* 2000; 18(1):92–96. [STEP 1]

In patients needing a postdates induction of labor, do outpatient mechanical inductions with single-balloon or double-balloon catheters reduce overall induction time?

EVIDENCE-BASED ANSWER

In patient with various indications for the induction of labor, including postdates, outpatient-based mechanical cervical ripening does not shorten the overall induction-to-delivery time (SOR: **A**, consistent randomized controlled trials [RCTs]). However, outpatient-based mechanical cervical ripening may shorten the overall hospitalization duration for the induction of labor (SOR: **B**, conflicting RCTs).

DOI 10.1097/EBP.0000000000001026

A 2018 randomized controlled trial (RCT) with 129 pregnant patients examined whether inductions started with outpatient cervical ripening shortened time from hospital admission to delivery and the total duration of hospitalization compared with inpatient cervical ripening.[1] Patients were parous women undergoing elective induction at ≥39 weeks of gestation with an unfavorable cervix and reassuring fetal heart tones. In both groups, mechanical ripening was performed with a Foley catheter filled with 30 mL of saline. The outpatient group was instructed to return at a predetermined admission time the following day, even if the catheter was expelled at home or present earlier at the start of labor symptoms. Patients were started on oxytocin at the time of hospital admission, and the catheter was allowed to remain in place up to 24 hours if not already expelled. The inpatient group had oxytocin started concurrently with Foley catheter placement at the time of admission. Nursing staff applied traction to the catheter every one to two hours, and the catheter was allowed to remain in place up to 24 hours. Once admitted, labor was managed in the same fashion for both groups, according to hospital protocol. Outpatient mechanical cervical ripening did not shorten time from hospital admission

to delivery, and there was no difference in total hospitalization duration in days (outpatient 2.6 vs inpatient 2.7; $P=.29$).

A 2017 RCT investigated the effectiveness of time to delivery with outpatient compared with inpatient cervical ripening with a Foley catheter in 130 pregnant patients.[2] Patients were term singleton pregnancies, cephalic presentation, Bishop score of <6, and with a gestational age of ≥41 weeks or a medical indication for induction of labor. This trial did not differentiate based on parity. Mechanical ripening was performed with a Foley catheter filled with 40 mL saline. The outpatient group was instructed to apply manual traction to the catheter every six hours and to return to the hospital for catheter expulsion, rupture of membranes, signs of labor, signs of fetal distress, or after 24 hours. The inpatient group also had manual traction applied to the catheter every six hours. The concurrent use of pharmaceutical induction agents with catheter placement was not discussed; however both groups received either prostaglandins or oxytocin for induction after catheter expulsion, per hospital policy. Initial analysis revealed that the outpatient group had a shorter catheter-to-delivery time (38 vs 45 hours; $P=.01$); however, this difference was not statistically significant after controlling for maternal age, parity, body mass index, gestational age, and indication for induction of labor. The outpatient group had a shorter hospitalization duration compared with the inpatient group (23 vs 36 hours; $P<.01$).

A 2001 RCT examined the effectiveness of outpatient versus inpatient cervical ripening with a Foley catheter with induction time and hospital duration in 111 pregnant patients.[3]. Patients were women with any non–high-risk medical indication for induction at ≥37 weeks of gestational age, term singleton pregnancy, vertex presentation, and reactive nonstress test and amniotic fluid index >5, with Bishop score of ≤5. Patients had a Foley catheter placed and filled to 30 mL with sterile water and were subsequently randomized to the outpatient or inpatient group. The outpatient group was instructed to return at 0600 the following morning, even if the catheter was expelled at home or present earlier for signs of labor or fetal distress. Upon presentation to the hospital, the outpatient group was started on oxytocin regardless of Foley catheter status. The inpatient group had a catheter placed at time of admission and was allowed to ambulate with catheter checks every two to four hours. Oxytocin was initiated upon Foley catheter expulsion. For both groups, the induction was managed at the discretion of the attending physician. No difference was observed in induction time between the groups (outpatient 1,473 minutes vs inpatient 1,472 minutes; $P=.90$). However, the authors stated that the overall induction time for the outpatient group may have been increased due to delay in oxytocin initiation for those patients who remained at home despite catheter explusion. The outpatient group did spent less time in the hospital (9.6 fewer hours of hospitalization overall; no P value provided). EBP

Katharine Kelly, MD
Emily Lines, MD
University of Colorado Family Medicine Residency
Denver, CO

The authors declare no conflicts of interest.

References

1. Kuper SG, Jauk VC, George DM, et al. Outpatient Foley catheter for induction of labor in parous women: a randomized controlled trial. *Obstet Gynecol.* 2018; 132(01):94–101. [STEP 2]
2. Policiano C, Pimenta M, Martins D, Clode N. Outpatient versus inpatient cervix priming with Foley catheter: a randomized trial. *Eur J Obstet Gynecol Reprod Biol.* 2017; 210:1–6. [STEP 2]
3. Sciscione AC, Muench M, Pollock M, et al. Transcervical Foley catheter for preinduction cervical ripening in an outpatient versus inpatient setting. *Obstet Gynecol.* 2001; 98:751–756. [STEP 2]

In adults with multiple chronic medical problems, does addressing one problem per office visit compared with addressing multiple problems per visit affect the time to address each specific medical problem?

EVIDENCE-BASED ANSWER

Probably. In office visits with more than one problem, physicians spent approximately 5 minutes on the primary problem but only 1 minute on each additional issue (SOR: **B**, observational study). Addressing multiple complaints in a single visit is twice as likely to result in longer visits and four times more likely to generate a sense of burden in the physician (SOR: **B**, cohort study).

DOI 10.1097/EBP.0000000000000839

A 2007 observational study (N=392) analyzed the time spent in primary care patient encounters and time allotment for specific topics.[1] Patients were ≥65 years old and met with their regular provider. Video tapes were

analyzed by coders to identify topics, symptoms raised by the patients, and total length of time spent by patient and physician on that topic. The average age of the patients was 74 years with 67% women. The primary outcome was the length of time was spent on any one topic. The average visit duration was 17 minutes. The average number of topics per visit was 6.5 with 72% of topics being biomedical in nature. Physicians spent 5.3 minutes on average for the longest topic and 1.1 minutes on average for each additional topic. The total visit length did not change with the number of topics. Longer time on major topics resulted in less time spent on each subsequent minor topic. Physicians observed in the study were 23% female with no difference in outcome related to physician gender. This observational study was limited by patient participation and prescreening by clinic office managers.

A 2016 cohort study (N=1,505) of patients suffering from bodily distress syndrome (multiple somatic complaints) compared with those without bodily distress syndrome assessed for provider management and time consumption.[2] Patients (average age, 46 years; 65% female) were selected randomly from primary care clinics. Patients who met screening criteria were 18% more likely than controls to be older, female, with lower education, and more chronic conditions. Physicians filled out information on time spent, medical complexity, follow-up, and burden level. Burden level was assessed on a 1 to 10 scale based on physician perception of the importance of the problems assessed at the visit and the number of problems addressed in addition to the main problem. The primary outcome showed that patients with bodily distress syndrome as compared with controls had longer visits (odds ratio [OR]; 1.8; 95% CI, 1.3–2.5) and higher perceived provider burden (OR, 2.5; 95 % CI, 1.8–3.6). No significant difference was noted for time consumption (OR, 1.06; 95% CI, 0.72–1.57) and burden (OR, 1.53; 95% CI, 0.99–2.37) between groups when adjusted for biomedical and psychosocial content. EBP

Matthew Balls, DO
Caroline Kirsch, DO
University of Wyoming Family Medicine Residency-Casper
Casper, WY

The authors declare no conflicts of interest.

References

1. Tai-Seale M, Mcguire TG, Zhang W. Time allocation in primary care office visits. *Health Serv Res*. 2007; 42(5): 1871–1894. [STEP 2]
2. Rask MT, Carlsen AH, Budtz-Lilly A, Rosendal M. Multiple somatic symptoms in primary care patients: a cross-sectional study of consultation content, clinical management strategy and burden of encounter. *BMC Fam Pract.* 2016; 17(1). [STEP 2]

Is lidocaine with epinephrine safe to use in digital blocks?

EVIDENCE-BASED ANSWER

Yes. The use of lidocaine with epinephrine (concentrations 1:100,000–200,000, or 5–10 ug/mL) is safe to use in digital nerve blocks in patients with normal digital circulation and does not cause tissue necrosis, infarction, or gangrene (SOR: **A**, systematic review of randomized controlled trials and cohort studies and a systematic literature review).

DOI 10.1097/EBP.0000000000001018

A 2015 systematic review of 15 controlled trials, 10 cohort studies, nine reviews, and five other articles examined the safety of epinephrine in digital nerve blocks.[1] Of the 15 controlled trials, 12 were randomized and three were nonrandomized (N=494). Patients were either volunteers or from emergency department, hand surgery, or plastic surgery settings. Patients with poor digital circulation, peripheral vascular disease, uncontrolled hypertension, pheochromocytoma, hyperthyroidism, and diabetes were excluded. Epinephrine concentrations mixed with local anesthesia ranged from 1:100,000 to 1:200,000 (5–10 ug/ml) in most trials. One randomized controlled trial (N=22) performed local extremity injection, which was not a digital nerve block technique. All other controls received local anesthesia without epinephrine via digital nerve block. Only two studies (N=103) directly examined the safety of epinephrine in digital nerve block as the primary study outcome. These two studies reported no ischemia-related complications and concluded that epinephrine could safely be used in digital nerve block. In the 13 other controlled trials, no ischemia-related tissue damage was reported, although safety was not a primary outcome in these studies. The only reported complications were in one randomized controlled trial (N=43) that reported three cases of superficial skin infection and one case of hypertensive

crisis in both the intervention and control groups. No other complications were reported from the other 14 controlled trials. Several of the controlled trials reported other benefits of using epinephrine including prolonging and accelerating onset of anesthesia, decreasing need for tourniquets, and decreasing need of additional injections of local anesthesia. Limitations of these controlled trials include small numbers, lack of clarity on whether providers were blinded, exclusion of patients at high risk for ischemia, and primary outcomes that varied widely and were mostly not safety related.

Eight retrospective and two nonblinded prospective cohort studies (N=270,488) were also included in the above 2015 systematic review. Only three cohort studies (N=2,474) directly examined the safety of epinephrine in digital nerve blocks as the primary study outcome. In these three studies, there were no reported complications, including no infarction or necrosis reported. No ischemia-related tissue damage was reported in any of the other seven cohort studies. Four cases of temporary ischemic symptoms relieved by vasodilator therapy were reported in one study examining accidental epinephrine autoinjector injections. Patients with peripheral vascular compromise were excluded in many studies.

A 2001 systematic literature review evaluated all available case reports of digital necrotic and ischemic complications after the use of epinephrine in digits.[2] A total of 20 reported cases of digital gangrene after anesthetic blocks with epinephrine were identified. Sixteen of these cases used an unknown concentration of epinephrine due to manual dilution. Of the four cases where epinephrine concentrations were known, reported concentrations ranged from 1:160,000 to 1:400,000. None of these cases involved commercially prepared lidocaine with epinephrine. There were multiple confounding variables that made it difficult to determine the exact cause of the tissue insult, including concurrent infection, use of hot soaks, and use of tight tourniquets. Most cases occurred more than 50 years ago without current commercially available forms of lidocaine with epinephrine. The procedure of digital nerve block itself and preparation of the anesthetic solution with epinephrine varied widely across case reports. EBP

Matt Ludemann, MD
Drew Blumberg, DO
St. Anthony North Family Medicine Residency
Westminster, CO

The authors declare no conflicts of interest.

References

1. Ilicki J. Safety of epinephrine in digital nerve blocks: a literature review. *J Emerg Med.* 2015; 49(5):799–809. [STEP 1]
2. Denkler KA. Comprehensive review of epinephrine in the finger: to do or not to do. *Plast Reconstr Surg.* 2001; 108(1): 114–124. [STEP 4]

Does oral zinc supplementation decrease the severity or duration of diarrheal disease in children?

EVIDENCE-BASED ANSWER

Oral zinc supplementation decreases the duration of acute diarrheal illness in children under five years old by up to 33 hours in both hospital and community settings in developing countries. The effects of oral zinc supplementation on diarrheal illness are greater in malnourished children (SOR: **B**, extrapolated from meta-analyses of randomized controlled trials [RCTs] in primarily developing countries). Treatment with zinc reduces the duration of hospitalization by 37% and stool frequency by 6% in south Asian countries (SOR: **B**, extrapolated from meta-analyses of RCTs in primarily developing countries).

DOI 10.1097/EBP.0000000000001063

A 2013 systematic review and meta-analysis of 18 randomized controlled trials (N=7,314 children) assessed children with acute diarrhea who received zinc supplementation when compared with placebo or oral rehydration therapy.[1] Studies were performed in both hospital and community-based settings in developing countries, and included children up to five years old with acute diarrhea. The children were administered oral zinc sulfate or gluconate (dose range for elemental zinc was between 2.145 and 45 mg), placebo or oral rehydration therapy for up to 14 days. Primary outcomes included diarrhea duration, stool frequency, and adverse events. Diarrhea was defined as three or more episodes of liquid stool within a 24-hour period. Oral zinc reduced the duration of diarrhea (14 trials, n=4,206, mean difference [MD] −20 hours; 95% CI, −29 to −11). This effect was noted to be greater in malnourished children (five trials, n=1,506, MD −33 hours; 95% CI, −34 to −28), which

was defined as a child having low weight-for-height Z score (<−2) or zinc level below 14 μmol/L. Zinc therapy reduced stool frequency by an average of 2.4 episodes per day (four trials, n=1,350; 95% CI, 3.8–1.0). The only relevant adverse reaction reported was vomiting, which had the highest incidence in the zinc-treated group (0.2% in the treated group vs 0.1% in the placebo group). Oral zinc reduced the prevalence of diarrhea on day three (four trials, n=3,126; 95% CI, 0.60–0.92) and day seven (eight trials, n=5,100; 95% CI, 0.47–0.83) but not on day five. Significant heterogeneity was observed between studies for the outcomes of diarrhea duration and prevalence. The authors believed this to be related to the varying nutritional status of the children across the studies. Additionally, none of the studies were performed in western countries.

A 2013 systematic review and meta-analysis of 104 studies (n=18,822) evaluated the effect of oral zinc supplementation of any zinc salt in comparison with oral placebo or supportive care (fluid infusion, probiotics, and antivirals) in children with acute diarrhea in south Asian countries.[2] Data were pooled from studies conducted in China (89 studies) and regions outside of China (15 studies) in both hospital and community settings. Studies assessed outcomes of diarrhea duration, the proportion of diarrhea episodes lasting greater than three and greater than seven days, duration of hospitalization, duration of fever, duration of vomiting, proportion of cases with vomiting, stool frequency (number), stool output (volume), and death from diarrhea or any cause. Children were under five years old with acute diarrhea, including dysentery, where diarrhea was defined as the passage of at least three loose or watery stools in a 24-hour period. The review excluded studies that exclusively enrolled a particular subgroup of children such as HIV-infected children or preterm infants and studies of persistent diarrhea. Documented treatment doses varied from 2.5 to 280 mg daily and course lengths varied from 3 to 14 days; however, not all studies listed doses or treatment length. Of the non-Chinese studies, acute episodes of diarrhea were 4% (95% CI, 1–8) shorter in duration among children treated with zinc compared with those receiving placebo. Among children hospitalized for diarrhea, treatment with zinc reduced the duration of hospitalization by 37% (95% CI, 21–53) and stool frequency by 6% (95% CI, 2–10). Of the Chinese studies, the reduction in the duration of diarrhea was 37% (95% CI, 35–39) among nonspecific episodes and 31% (95% CI, 29–34) among rotavirus episodes. Among zinc-treated patients, diarrhea lasting beyond three days was reduced with nonspecific (relative risk [RR]=0.73; 95% CI, 0.66–0.79) and rotavirus (RR=0.70; 95% CI, 0.63–0.78) diarrhea. An increased risk of vomiting (RR 1.8; 95% CI, 1.4–2.4) was noted in the non-Chinese children. In China, placebo supplements may not have been readily available and blinding may have been insufficient resulting in a difference between the mean episode duration of zinc-treated and control group children when compared with non-Chinese studies. EBP

David A. Moss, MD
Heather N. Foster, MD
Nellis AFB 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. Galvao TF, Thees MF, Pontes RF, Silva MT, Pereira MG. Zinc supplementation for treating diarrhea in children: a systematic review and meta-analysis. *Rev Panam Salud Republica*. 2013; 33(5):370–377. [STEP 1]
2. Lamberti LM, Walker CL, Chan KY, Jian WY, Black RE. Oral zinc supplementation for the treatment of acute diarrhea in children: a systematic review and meta-analysis. *Nutrient*. 2013; 5(11):4715–4740. [STEP 1]

Do non-dihydropyridine calcium channel blockers decrease progression of chronic kidney disease?

EVIDENCE-BASED ANSWER

Non-dihydropyridine calcium channel blockers, alone and in combination with other hypertensive medications, may decrease mortality in end stage renal disease by approximately 20% to 40%, decrease proteinuria by 40% to 60% and slow the worsening of creatinine clearance in patients with chronic kidney disease (SOR: **B**, small randomized controlled trials and retrospective cohort studies).

DOI 10.1097/EBP.0000000000001037

A 2002 observational retrospective cohort study sought to determine if calcium channel blockers (CCBs)

decreased mortality in patients with end stage renal disease (ESRD).[1] Data were extracted from the United States Renal Database System Dialysis Morbidity and Mortality Study Wave II, which is compiled annually. Of the 3,716 patients identified, 51.2% were taking any CCB (n=1,902, mean 58 years old, 46.8% female, 60.8% white) and 11.8% were taking an non-dihydropyridine calcium channel blockers (NDCCB) (verapamil, 2.0% or diltiazem, 9.8%). The primary endpoint was progression to death. After controlling for age, low serum albumin, diabetes, pre-existing cardiac disease, smoking status, diastolic blood pressure, undernourishment, and race, a Cox proportional hazards model analysis was performed giving an adjusted relative risk of all-cause mortality when using any CCB of 0.79 (95% CI, 0.69–0.90; P=.001) and 0.63 (95% CI, 0.49–0.81, P<.001) when using diltiazem specifically. Limitations of the study included its retrospective nature, and all patients already had ESRD.

A 1996 randomized controlled trial (RCT) (N=52) compared the effects of lisinopril (n=18), atenolol (n=16), or a NDCCB (diltiazem: n=10 or verapamil: n=8) on the progression of nephropathy.[2] The trial included patients greater than 45 years old, with hypertension over eight years duration, non–insulin-dependent diabetes mellitus associated over 8 years duration, and proteinuria greater than 2.0 g/d at baseline. Patients were followed for a mean duration of 64 months, initially weekly for one month then approximately quarterly (range, 36–73 months) with a minimum of four visits and maximum of 14 visits. Dosing was titrated to maintain similar mean blood pressures between treatment groups. The mean rate of decline in creatinine clearance (CrCl) for lisinopril was –0.98 mL/min/y/1.73 m^2, –1.44 mL/min/y/1.73 m^2 for NDCCBs, and –3.48 mL/min/y/1.73 m^2 for atenolol. After analysis of variance was performed, both the lisinopril and the NDCCB groups had a slower rate of decrease in CrCl compared with the atenolol group (P=.0001 and P=.004, respectively). No difference was noted between the lisinopril and NDCCB groups (P=.11). Albuminuria was reduced similarly between lisinopril and NDCCB groups (–0.713 g/d and –0.818 g/d, respectively; P>.99) but remained elevated in the atenolol group (0.168 g/d; P<.01). Compared with atenolol, albuminuria was significantly reduced in the lisinopril group (P=.016) and NDCCB group (P=.012). Limitations included irregular follow-up intervals and frequency, and use of alpha blockers, central alpha agonists, and hydralazine in some patients during the study was not controlled for in the analysis.

A 1998 RCT (N=21) evaluated the effects of CCBs on proteinuria and glomerular filtration rate (GFR).[3] The trial included patients >45 years old with proteinuria greater than 300 mg/d, and both type II diabetes and hypertension over 4 years duration. Initially, 24 males and four females were recruited; 96% were Caucasian. Patients were randomized to 21 months of treatment with either nifedipine (n=10) or diltiazem (n=11) after two weeks of washout from any previous antihypertensive therapy. Dosing was titrated to achieve a blood pressure goal of <140/90 mmHg or until a max of 90 mg nifedipine or 480 mg diltiazem daily was attained. Furosemide 40 mg daily, followed by clonidine 0.1 mg daily if needed, was added if pressure control proved inadequate. In the diltiazem group, baseline proteinuria of 908 mg/d was decreased to 389 mg/d, a change of 57% (P<.05). However, treatment with nifedipine resulted in no significant change in proteinuria (873 mg/d to 905 mg/d, difference of 4%). GFR was not significantly different within either diltiazem group (98 mL/min/1.73 m^2 to 101 mL/min/1.73 m^2, P>.05) or nifedipine group (94 mL/min/1.73 m^2 to 91 mL/min/1.73 m^2, P>.05) at baseline compared with 21 months. Additionally, GFR was not different between groups at the conclusion of the study.

In 2008, a multisite, single blinded RCT (N=304) compared the progression of nephropathy after combined hypertensive therapy.[4] The trial included patients with type II diabetes (64% male, 74% white, mean age 60 years old) with a baseline urine albumin creatinine ratio [UACR] greater than 0.2 g/g that were treated with either trandolapril/verapamil (T/V, n=110) or benazepril/amlodipine (B/A, n=127). Medications were titrated to goal pressure of less than 130/80 mmHg over 4 weeks and periodically thereafter. Torsemide 10 to 40 mg was added followed by an additional non-CCB, non-ARB, non-mineralocorticoid to achieve goal. At 36 weeks, the UACR was increased in both groups (T/V, 29.3% vs B/A, 8.5%, P=.34 between groups). Mean eGFR declined in both groups (T/V, –4.8 mL/min/1.73 m^2 vs B/A, –2.1 mL/min /1.73 m^2, P=.48 between groups). Limitations included different ACEIs combined with each CCB, a high dropout rate, and variable addition of torsemide. EBP

David Moss, MD
Gino Gaddini, DO
Nellis Family Medicine Residency
Nellis AFB, 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 US Air Force Medical Department, the Air Force at large, or the Department of Defense.

REFERENCES

1. Kestenbaum B, Gillen DL, Sherrard DJ, Seliger S, Ball A, Stehman-Breen C. Calcium channel blocker use and mortality among patients with end-stage renal disease. *Kidney Int*. 2002; 61(6):2157–2164. [STEP 2]
2. Bakris GL, Copley B, Vicknair N, Sadler R, Leurgans S. Calcium channel blockers versus other antihypertensive therapies on progression of NIDDM associated nephropathy. *Kidney Int.* 1996; 50:1641–1650. [STEP 2]
3. Smith AC, Toto R, Bakris GL. Differential effects of calcium channel blockers on size selectivity of proteinuria in diabetic glomerulopathy. *Kidney Int*. 1998; 54(3)889–896. [STEP 2]
4. Toto RD, Tian M, Fakouhi K, Champion A, Bacher P. Effects of calcium channel blockers on proteinuria in patients with diabetic nephropathy. *J Clin Hypertens*. 2008; 10(10) 761–769. [STEP 2]

Is HPV self-swab collection as effective as clinician collection for cervical cancer screening?

EVIDENCE-BASED ANSWER

Human papillomavirus (HPV) self-swab collection is as accurate as clinician collection for detection of CIN2+ or worse if using high-risk HPV (hrHPV) assays based on polymerase chain reaction (SOR: **A**, meta-analysis).

DOI 10.1097/EBP.0000000000001058

A 2014 meta-analysis examined 36 accuracy studies of 154,556 women to determine the utility of high-risk human papillomavirus (HPV) (hrHPV) testing on self-collected samples compared with clinician-collected samples.[1] Eligible studies included the following three criteria: the vaginal sample was self-collected by participants followed by a clinician-collected sample or two-armed randomized control trials with one arm using self-collected samples and the other arm using clinician-collected samples; hrHPV testing was performed on both samples or the clinician-collected sample was examined microscopically; and all enrolled women or those with a positive test underwent colposcopy and biopsy to evaluate for cervical intraepithelial neoplasia grade 2 or worse (CIN2+). This meta-analysis included studies that used both signal amplification and polymerase chain reaction (PCR)–based HPV assays to determine the accuracy for finding CIN2+ or CIN grade 3 or worse (CIN3+). In primary screening, self-collected samples had a pooled sensitivity estimate of 76% for CIN2+ and 84% for CIN3+. The pooled sensitivity of self-collected HPV tests for primary screening was lower than clinician-collected HPV samples for detecting CIN2+ (relative sensitivity 0.88; 95% CI, 0.85–0.91) and CIN3+ (relative sensitivity 0.89; 95% CI, 0.83–0.96). The pooled specificity estimate of self-collected primary screening HPV samples in excluding CIN2+ was 86%. The pooled specificity of self-collected HPV tests compared with clinician-collected HPV samples was slightly less for detecting CIN2+ (relative specificity 0.96; 95% CI, 0.95–0.97) and CIN3+ (relative specificity 0.96; 95% CI, 0.93–0.99). The authors noted test effects, with self-collected HPV assays using signal amplification having lower sensitivity and specificity compared with HPV testing performed by clinicians (data unable to be pooled). The authors concluded hrHPV testing on self-collected samples using PCR has potential for use in routine cervical cancer screening.

The meta-analysis above was updated in 2018 to include 56 accuracy studies, with 22 new diagnostic studies.[2] This meta-analysis again evaluated hrHPV testing on self-collected versus clinician-collected samples. Studies were eligible with similar criteria as noted above: a vaginal sample was self-collected by the participant followed by a clinician-collected sample, the samples were tested with the same hrHPV assay, and colposcopy and biopsy were used to evaluate for CIN2+ in all enrolled participants or those with one or more positive tests.

The updated meta-analysis found that PCR-based hrHPV testing using self-collected samples was as sensitive as clinician-collected samples in detecting CIN2+ (pooled relative sensitivity ratio 0.99; 95% CI, 0.97–1.02) and CIN3+ (pooled relative sensitivity ratio 0.99; 95% CI, 0.96–1.02); however, it was slightly less specific than clinician-collected samples at excluding CIN2+ (pooled relative specificity ratio 0.98; 95% CI, 0.97–0.99) and CIN3+ (pooled relative specificity ratio 0.98; 95% CI, 0.97–0.99). The PCR-based tests had a pooled absolute sensitivity of 96% and a pooled absolute specificity of 79% for both self-collected and clinician-collected samples.

Similar to the 2014 analysis, the updated meta-analysis found HPV assays using signal amplification were less sensitive and slightly less specific on tests from self-

collected samples in comparison with clinician-collected samples for CIN2+ (pooled relative sensitivity ratio 0.85; 95% CI, 0.80–0.89; pooled relative specificity ratio 0.96; 95% CI, 0.93–0.98) and CIN3+ (pooled relative sensitivity ratio 0.86; 95% CI, 0.76–0.98; pooled relative specificity ratio 0.97; 95% CI, 0.95–0.99). The pooled absolute sensitivity of hrHPV testing using signal amplification on self-collected samples was 77% in comparison with 93% sensitivity in clinician-collected samples, whereas the pooled absolute specificity to exclude CIN2+ was 84% in self-collected samples compared with 86% in clinician-collected samples. The authors concluded that using hrHPV assays based on PCR with self-collected samples was as accurate as clinician-collected samples.

The latest recommendations from both the United States Preventive Services Task Force (USPSTF)[3] and the American College of Gynecology (ACOG)[4] include hrHPV screening as an option for cervical cancer screening for women age 30 to 65 years old (USPSTF Grade A recommendation; no strength of recommendation by ACOG).[3,4] In addition, the USPSTF stated that hrHPV testing has the "potential to be collected by the patient and mailed to health programs for analysis." EBP

Kelley Collier, MD
Rebecca Mullen, MD, MPH
University of Colorado Family Medicine Residency
Denver, CO

The authors declare no conflicts of interest.

REFERENCES

1. Arbyn M, Verdoodt F, Snijders PJ, et al. Accuracy of human papillomavirus testing on self-collected versus clinician-collected samples: a meta-analysis. *Lancet Oncol*. 2014; 15(2):172–183. [STEP 1]
2. Arbyn M, Smith SB, Temin S, et al. Detecting cervical precancer and reaching underscreened women by using HPV testing on self samples: updated meta-analyses. *BMJ*. 2018; 363:k4823. [STEP 1]
3. US Preventive Services Task Force, Curry SJ, Krist AH, Owens DK, et al. Screening for cervical cancer: US Preventive Services Task Force Recommendation Statement. *JAMA*. 2018; 320(7):674–686. [STEP 1]
4. American College of Obstetricians and Gynecologists. *Cervical Cancer Screening Update. ACOG Practice Advisory*. Washington, DC: American College of Obstetricians and Gynecologists; 2018. Available at: https://www.acog.org/clinical/clinical-guidance/practice-advisory/articles/2018/08/cervical-cancer-screening-update. [STEP 5]

Is intermittent supplementation with oral iron more effective than daily dosing in reducing anemia and associated outcomes?

EVIDENCE-BASED ANSWER

It depends on the indication. Daily oral iron supplementation is more effective than intermittent iron for reducing anemia in children younger than 12 years, whereas intermittent iron is as effective as daily iron in menstruating or pregnant women (SORT **B**, meta-analysis of lower-quality randomized controlled trials [RCTs]). Intermittent iron supplementation has fewer adverse side effects (mostly gastrointestinal) in menstruating or pregnant women, but no difference in side effects is noted in children younger than 12 years (SORT **B**, meta-analysis of lower-quality RCTs).

DOI 10.1097/EBP.0000000000000870

A 2019 systematic review and meta-analysis of 25 randomized controlled trials (RCTs) and quasi-RCTs (N=10,996) compared intermittent oral iron supplementation (dosing on nonconsecutive days one, two, or three days of the week) with daily oral iron or placebo in menstruating women.[1] Study participants were from predominately low- or middle-income populations in 15 different countries, including studies from Europe (1), Latin America (4), Africa (5), and Asia (15). Of note, multiple studies included folic acid or other supplements in both the intervention and control groups. Oral weekly dosing of elemental iron ranged from 10 to 120 mg in the intermittent group and from 50 to 300 mg in the daily dosing groups. The follow-up period for 13 studies was three months or less, and for 11 studies, it was more than three months to one year. Compared with daily oral iron supplementation (with folic acid or other supplements or alone), intermittent oral iron supplementation (with folic acid or other supplements or alone) demonstrated similar results for the primary outcomes of anemia, hemoglobin concentration, and iron deficiency, although ferritin was statistically higher in the daily supplementation group (see **TABLE 1**). Possibly most importantly, from

TABLE 1. Intermittent[a] versus daily oral iron supplementation for reducing iron deficiency anemia and associated outcomes in menstruating women[1]

Outcome	Participants (studies)	Analysis of difference (95% confidence interval)
Anemia	1,749 (8)	RR 1.1 (0.93–1.3)
Hemoglobin concentration	2,127 (10)	MD 0.04 g/dL (–0.14 to 0.23)
Ferritin concentration	988 (4)	**MD –6.1 ng/mL (–11 to –1.5)**
Iron deficiency	198 (1)	RR 4.3 (0.56–33.2)
Any adverse side effect	1,166 (6)	**RR 0.42 (0.21–0.82)**

Statistically significant differences in BOLD. [a] Intermittent=dosing on nonconsecutive days between one and three times weekly. CI=confidence interval; MD=mean difference; RR=relative risk.

a patient-oriented perspective, women receiving intermittent dosing experienced significantly fewer side effects (NNT=8, primarily gastrointestinal) (see **TABLE 1**). Study bias was difficult to ascertain as methodology was unclear in most studies, although high attrition rates, and inadequate blinding was noted in many of the studies contributing to a GRADE working group assignment of "low" or "very low" quality for the primary outcomes.

A 2019 systematic review and meta-analysis of 21 RCTs and quasi-RCTs (N=5,490) compared intermittent oral iron supplementation (same definition as preceding review) with daily oral iron or placebo in pregnant women.[2] At the time of randomization, participants could be of any gestational age and parity, and studies targeting specific illnesses such as AIDS and tuberculosis were excluded. Studies took place in Latin America (6) and Asia (15). The dosing of elemental iron in the intermittent groups ranged from 80 to 300 mg and in the daily groups (40 to 120 mg). Compared with daily oral iron supplementation (with folic acid or other supplements or alone), intermittent oral iron supplementation (with folic acid or other supplements or alone) demonstrated similar results for the primary infant outcomes of birth weight, low birth weight (<2,500 g), premature birth (<37 weeks' gestation), and neonatal death within 28 days of birth (see **TABLE 2**). No difference was observed in the maternal outcome of anemia at term, although again, there was a significant reduction in side effects with the intermittently dosed women (NNT=10) (see **TABLE 2**). Using GRADE working group criteria, the primary outcomes in this review were assessed as being from studies of "low" or "very low quality."

TABLE 2. Intermittent[a] versus daily oral iron supplementation for reducing anemia and associated outcomes in pregnant women[2]

Outcome	Participants (studies)	Analysis of difference
Infant outcomes		
Low birth weight (<2,500 g)	1,898 (8)	RR 0.82 (0.55–1.2)
Birth weight (g)	1,939 (9)	MD 5.2 g (–30 to 40)
Premature birth (<37 wks' gestation)	1,177 (5)	RR 1.03 (0.76–1.4)
Neonatal death (within 28 d after delivery)	795 (1)	RR 0.49 (0.04–5.4)
Maternal outcomes		
Anemia at term	676 (4)	RR 1.22 (0.84–1.8)
Side effects (any)	1,777 (11)	**RR 0.56 (0.37–0.84)**

Statistically significant differences in BOLD. [a] Intermittent=dosing on nonconsecutive days between one and three times weekly. CI=confidence interval; MD=mean difference; RR=relative risk.

TABLE 3. Intermittent[a] versus daily oral iron supplementation for reducing iron deficiency anemia and associated outcomes in children of 0 to 12 years[3]

Outcome	Participants (studies)	Analysis of difference (95% confidence interval)
Anemia	980 (6)	**RR 1.2 (1.04–1.5)**
Hemoglobin concentration	2,851 (19)	MD –0.06 g/dL (–0.15 to 0.04)
Ferritin concentration	902 (10)	MD –4.2 ng/mL (–9.4 to 1.1)
Iron deficiency	76 (1)	**RR 4.0 (1.2–13)**
Any adverse side effect	895 (4)	RR 0.6 (0.19–1.9)

Statistically significant differences in BOLD. [a] Intermittent=dosing on nonconsecutive days between one and three times weekly. CI=confidence interval; MD=mean difference; RR=relative risk.

A 2011 systematic review and meta-analysis of 33 RCTs and quasi-RCTs (N=13,114) compared intermittent oral iron supplementation (same definition as preceding reviews) with daily oral iron or placebo in children younger than 12 years.[3] Study participants were from predominately low- or middle-income populations, including studies from Latin America (7), Africa (8), and Asia (18). The dosing of elemental iron in the intermittent groups ranged from 7.5 to 200 mg with the daily group's dosing not fully reported because most dosage schedules were reported as milligrams per kilogram per day and comparative weighs or average total dosages were not reported. Eleven studies lasted up to three months and eight greater than three months. Compared with daily oral iron supplementation (with folic acid or other supplements or alone), intermittent oral iron supplementation (with folic acid or other supplements or alone) demonstrated similar results for adverse side effects but improved outcomes for hemoglobin and ferritin concentrations (see **TABLE 3**). Anemia and iron deficiency anemia were significantly improved in the daily regimen groups (see **TABLE 3**). This review lacked variance in blinding, allocation, selective reporting, attrition, and other sources of potential bias. Using GRADE working group criteria, the primary outcomes in this review were assessed as being from studies of "low" or "very low quality." EBP

Timothy Mott, MD
Ethan Sellers, DO
South Baldwin Regional Medical Center Family Medicine Residency, Foley, AL

The authors declare no conflicts of interest.

References

1. Fernández-Gaxiola AC, De-Regil LM. Intermittent iron supplementation for reducing anaemia and its associated impairments in adolescent and adult menstruating women. *Cochrane Database Syst Rev.* 2019; (1):CD009218. [STEP 1]
2. Peña-Rosas JP, De-Regil LM, Gomez Malave H, Flores-Urrutia MC, Dowswell T. Intermittent oral iron supplementation during pregnancy. *Cochrane Database Syst Rev.* 2015; (10):CD009997. [STEP 1]
3. De-Regil LM, Jefferds MED, Sylvetsky AC, Dowswell T. Intermittent iron supplementation for improving nutrition and-development in children under 12 years of age. *Cochrane Database Syst Rev.* 2011; (12):CD009085. [STEP 1]

How accurate is HbA1c in diagnosing diabetes in children?

EVIDENCE-BASED ANSWER

HbA1c is an inaccurate method for identifying prediabetes and type 2 diabetes in obese pediatric populations (strength of recommendation [SOR]: **B**, based on two large cohort studies).

DOI 10.1097/EBP.0000000000001061

A large cohort study evaluated the accuracy of HbA1c for identifying pediatric patients at risk for type 2 diabetes and prediabetes. Patients (N=1,156) were recruited from a pediatric obesity clinic in the United States from 2005 to 2010. All patients were obese (>95th percentile for age and sex), had no diabetes history, and took no medications that would affect glucose

metabolism. Type 2 diabetes was defined as a fasting glucose of >125 mg/dL or a two-hour oral glucose tolerance test (OGTT) of ≥200 mg/dL. Prediabetes was defined as a fasting glucose of 100 to 125 mg/dL or a two-hour OGTT of 140 to 199 mg/dL. Testing was administered after a 10-hour fast. A standard OGTT was administered with a 1.75 g/kg dose of glucose (maximum of 75 g) on all subjects. Fasting and two-hour plasma glucose were measured using radioimmunoassays, and HbA1c was measured on the same day using an assay based on latex immunoagglutination inhibition methodology. Of the 1,156 patients, 31 had type 2 diabetes according to the OGTT criteria and only 10 of these patients had an HbA1c >6.4. An additional 347 patients were prediabetic per OGTT, but only 103 of those patients had an HbA1c in the range of 5.7 to 6.4. Based on a threshold HbA1c of 5.8% for identifying risk for type 2 diabetes, HbA1c showed only a 78% specificity and 68% sensitivity when compared with the gold standard OGTT.

Another cohort study in 2012 in the United Kingdom (n=266) sought to determine if HbA1c could accurately identify those at risk for prediabetes and type 2 diabetes in obese children. Patients were excluded if pregnant, previously diagnosed with diabetes or glucose intolerance, they took any chronic medication, or had a history of hypoglycemia or hemoglobinopathy. Type 2 diabetes and prediabetes were defined by the same standards as the first study. Testing was administered after a 12-hour fast. A standard OGTT was administered with the same glucose load as the previous study to all subjects. Fasting and two-hour plasma glucose were measured using the hexokinase/G6PD method, and HbA1c was measured per the National Glycohemoglobin Standardization Program via the monoclonal anti-HbA1c antibody method. Per the OGTT, 13 subjects were identified as prediabetic, but only three of those patients had an HbA1c between 5.7 and 6.4. No patients were identified as diabetic per OGTT. Based on a threshold HbA1c of 5.7% for identifying risk for type 2 diabetes, HbA1c had an 88% specificity and 23% sensitivity when compared with the standard-of-care OGTT. EBP

Jared Eames, DO
Jedda Rupert, MD
Nellis AFB Family Medicine Residency
Nellis AFB, Las Vegas, NV

The author declares 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 Air Force Medical Department, the Air Force at large, or the Department of Defense.

References

1. Galhardo J, Shield J. The role of haemoglobin A1c in screening obese children and adolescents for glucose intolerance and type 2 diabetes. *Acta Med Port*. 2015; 28(3): 307–315. [STEP 2]
2. Nowicka P, Santoro N, Liu H, et al. Utility of hemoglobin A(1c) for diagnosing prediabetes and diabetes in obese children and adolescents. *Diabetes Care*. 2011; 34(6):1306–1311. [STEP 2]

What is the best pharmacotherapy for periodic limb movement disorder?

EVIDENCE-BASED ANSWER

Rotigotine patch reduces the number of periodic limb movements per hour in adults with periodic limb movement disorder (SOR: **A**, meta-analysis of randomized controlled trials [RCTs]). Pramipexole and ropinirole are also likely effective (SOR: **B**, extrapolated from a meta-analysis of RCTs for restless leg syndrome). Levodopa with a DOPA decarboxylase inhibitor and gabapentin enacarbil may also be considered after trying pramipexole and ropinirole (SOR: **B**, evidence-based guideline).

DOI 10.1097/EBP.0000000000001076

A 2018 systematic review and meta-analysis of five randomized controlled trials (RCTs) examined transdermal rotigotine patch use in 197 adult patients with periodic limb movement in sleep (periodic limb movement disorder [PLMD]).[1] Rotigotine therapy lasted between three weeks and 20 weeks. Maximal dosages ranged from 3 mg/24 h to 16 mg/24 h. The primary outcome was the change in frequency of PLMD. Rotigotine decreased the number of periodic limb movements per hour on polysomnogram compared with the placebo (three trials; N=177; weighted mean difference [WMD] –32, 95% CI, –43 to –22).

A 2016 meta-analysis of 12 multicenter RCTs with 3,286 adult patients with moderate-to-severe restless leg

syndrome (RLS) evaluated the efficacy of pramipexole in the treatment of RLS.[2] Patients included were older than 18 years, with a formal diagnosis of RLS by the International Restless Legs Syndrome Study Group diagnostic criteria. Patients received treatment from three weeks to 26 weeks (mean duration 11 weeks/person), with doses ranging from 0.125 mg/d and 1.5 mg/d. The primary outcome was measured by posttreatment change in the International Restless Leg Syndrome Study Group Rating Scale (IRLS) score (10 items to rate the severity of symptoms in the past week, scored 0–4, highest score of 40—most severe). Secondary outcomes were measured using several other scales, including the Clinical Global Impression of Improvement (CGI-I) scale (one to seven, very much improved to very much worse) the Medical Outcomes Study sleep disturbance score (MOS, a subjective scale with six primary sleep domains, range 0–100 for each domain), and the RLS-Quality of Life (QOL) score (an 18-item scale scored 0–100, with higher scores correlating to higher QOL). More patients in the pramipexole group experience at least a 50% reduction in IRLS score compared with the placebo (eight trials; N=2,188; risk ratio [RR] 1.6; 95% CI, 1.4–1.7). Compared with the placebo, more patients in the pramipexole group were positive responders (scored "very much improved" or "much improved") using the CGI-I scale (11 trials; N=3,234; 66% vs 44%; RR 1.5; 95% CI, 1.3–1.7) and on the Patient Global Impression scale (nine trials; N=2,568; 63% vs 41%; RR 1.5; 95% CI, 1.3–1.8). Pramipexole also improved QOL as compared to placebo (four trials; N=1,397; WMD 5.4; 95% CI, 2.3–8.5). Finally, pramipexole was associated with decreased daytime tiredness compared with the placebo (four trials; N=1,411; WMD –0.61; 95% CI, –1.2 to –0.01), as measured by the MOS sleep disturbance score.

A 2009 meta-analysis of six double-blinded RCTs in the United States and Europe examined the effectiveness of ropinirole in 1.679 adult patients between 18 and 79 years of age with primary moderate-to-severe RLS.[3] Patients were given ropinirole immediate-release 0.25 to 6 mg/d or placebo over a 12-week duration. Outcomes included the end point changes in both MOS sleep disturbance score and CGI-I scale when compared with the placebo. Patients were more likely to be labeled as positive responders (scored "very much improved" or "much improved") to ropinirole compared with the placebo using the CGI-I scale (63% vs 47%; odds ratio 1.9; 95% CI, 1.5–2.3). Ropinirole resulted in an increase in sleep per night (44 min/night vs 22 min/night; $P<.001$) compared with the placebo. Using the MOS scale, patients treated with ropinirole show improvement in sleep adequacy (WMD 7.2; 98.75% CI, 3.8–11), sleep disturbance (WMD –8.5; 98.75% CI, –11.5 to –5.5), and daytime somnolence (WMD –3; 98.75% CI, –5.3 to –0.7) as compared to placebo.

A 2012 American Academy of Sleep Medicine evidence-based guideline for the management of RLS and PLMD in adults recommended pramipexole and ropinirole as standard first-line treatment with benefits clearly outweighing harms in patients with moderate-to-very severe symptoms (Level of Evidence—high).[4] Off-label use of levodopa with a DOPA decarboxylase inhibitor and gabapentin enacarbil were also given as options after a trial of pramipexole and ropinirole (Level of Evidence—high). The guideline also mentioned the Food and Drug Administration off-label use of opioids, pregabalin, carbamazepine, and clonidine as options, although with an unclear benefit and harm/benefit balance (Level of Evidence—low). The use of supplemental iron in patients with low ferritin levels was noted as having an unclear benefit/harm balance (Level of Evidence—very low). EBP

Daffer Ghanim, MD
Kenneth Herring, MD
Corey Lyon, DO
University of Colorado FMR, Denver, CO

The authors declare no conflicts of interest.

References

1. Wu MN, Tseng PT, Chen TY, et al. The efficacy and tolerability of rotigotine on patients with periodic limb movement in sleep: a systematic review and meta-analysis. *PLoS One*. 2018; 13(4):e0195473. [STEP 1]
2. Liu GJ, Wu L, Lin Wang S, Xu LL, Ying Chang L, Fu Wang Y. Efficacy of pramipexole for the treatment of primary restless leg syndrome: a systematic review and meta-analysis of randomized clinical trials. *Clin Ther*. 2016; 38(1):162–179. [STEP 1]
3. Hansen RA, Song L, Moore CG, et al. Effect of ropinirole on sleep outcomes in patients with restless legs syndrome: meta-analysis of pooled individual patient data from randomized controlled trials. *Pharmacotherapy*. 2009; 29(3): 255–262. [STEP 1]
4. Aurora RN, Kristo DA, Bista SR, et al. The treatment of restless legs syndrome and periodic limb movement disorder in adults—an update for 2012: practice parameters with an evidence-based systematic review and meta-analyses:

an American Academy of Sleep Medicine Clinical Practice Guideline. *Sleep*. 2012; 35(8):1039–1062. [STEP 1]

Does motivational interviewing in primary care appointments improve HbA1c in patients with type 2 diabetes mellitus?

EVIDENCE-BASED ANSWER

No. Among patients with diabetes, motivational interviewing does not significantly improve HbA1c (SOR: **B**, extrapolated from a meta-analysis of patients with type 1 and type 2 diabetes). Among patients with type 2 diabetes, the addition of motivational interviewing does not further improve HbA1c, body mass index, blood pressure, and total cholesterol or psychological distress when compared with standard care alone (SOR: **B**, randomized controlled trial).

DOI 10.1097/EBP.0000000000001071

A 2014 systematic review and meta-analysis evaluated the evidence for the efficacy of motivational interventions in promoting glycemic control in diabetic patients compared with usual care.[1] Inclusion criteria for randomized controlled trials (RCTs) completed between 1983 and 2013 were patient samples diagnosed with diabetes, an intervention arm aimed at increasing motivation versus a control arm of usual care, and HbA1c values pretreatment and post-intervention. Usual care varied by study and ranged from basic standard of care at the practice to scheduled diabetes education and counseling classes. Thirteen RCTs consisting of 3,327 patients (1,223 DM-1; 1,895 DM-2; 209 unspecified) were selected. Baseline HbA1c ranged from 6.8% to 11.9% and interventions ranged from eight to 18 months. The analysis assessed the difference in mean change in HbA1c and was presented so that a positive effect favored the intervention and a negative effect favored the control. The pooled mean difference in HbA1c change between groups was 0.17% (95% CI, –0.09% to 0.43%), showing a positive, though not statistically significant, effect.

A 2018 cluster RCT (N=334) investigated whether integrating motivational interviewing into diabetes care was more effective at improving glycemic control compared with standard care alone.[2] Patients were adults 18 to 79 years old (mean age 59), recruited from inner-city general practices in London who had type 2 diabetes for at least two years, a HbA1c greater than 8.0% over the preceding 18 months, and who were on at least two oral diabetes medications or insulin. The control group (12 practice clusters, N=170) received nurse-led standard diabetes care per national guidance for 12 sessions, 30 minutes each, over 12 months. In the same number of sessions, the intervention group (12 practice clusters, N=164) received standard care plus integrated motivational interviewing techniques by nurses trained in six skills drawn from motivational interviewing. The six skills included active listening, managing resistance, directing change, supporting self-efficacy, addressing health beliefs, and shaping behaviors. The primary outcome was change in HbA1c and the secondary outcomes were changes in blood pressure, body mass index, total cholesterol, depressive symptoms using Patient Health Questionnaire-9 (PHQ-9), and diabetes-specific psychological burden using the Diabetes Distress Scale, measured at study enrollment and follow-up at approximately 18 months. HbA1c was measured by a fasting blood draw, and a difference between groups of 1% was the minimum needed to be considered significant. Intention-to-treat analysis was performed with 103 patients lacking HbA1c at follow-up. At the end of the study period, motivational interviewing integration did not improve HbA1c compared with standard care. No significant effects were noted on secondary outcomes, and costs were higher in the intervention group because of training expenses. Limitations of this study included being slightly underpowered at 77% compared with the proposed 80% (because of high attrition rate), and intervention nurses failed to show significantly higher proficiency in motivational interviewing skills compared with control nurses. EBP

Broc Parker, DO
Jennifer Svarverud, 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 US Air Force Medical Department, the Air Force at large, or the Department of Defense.

References

1. Jones A, Gladstone B, Lübeck M, Lindekilde N, Upton D, Vach W. Motivational interventions in the management of HbA1c levels: a systematic review and meta-analysis. *Prim Care Diabetes*. 2014; 8(2):91–100. [STEP 1]
2. Ismail K, Winkley K, de Zoysa N, et al. Nurse-led psychological intervention for type 2 diabetes: a cluster randomized controlled trial (Diabetes-6 study) in primary care. *Br J Gen Pract*. 2018; 68(673):e531–e540. [STEP 2]

What examination findings are best for differentiating central causes of vertigo from peripheral causes of vertigo?

EVIDENCE-BASED ANSWER

A series of oculomotor examinations evaluating the integrity of primary vestibular pathways, gaze holding circuits, and otolithic pathways from the pons to the cerebellum are effective at distinguishing central from peripheral vertigo (SOR: **B**, consistent results from a small cohort and cross-sectional study). Examination findings evaluating only stroke risk factors may be less helpful (SOR: **B**, cross-sectional study).

DOI 10.1097/EBP.0000000000001050

A 2013 cross-sectional study (N=190) compared the accuracy of detecting central and peripheral vertigo using the HINTS physical examination (Head Impulse, Nystagmus type, and Test of Skew) and a summed stroke risk score ABCD2 exam (Age, Blood pressure, Clinical features, Duration of symptoms, and Diabetes).[1] Patients (124 central, 66 peripheral) included those with at least one hour of acute vestibular syndrome symptoms, within one week of onset, along with at least one of the stroke risk factors. Patients had a median age of 61 years old and 61% were men. The gold standard for determining central or peripheral cause was neuroimaging (97% by diffusion-weighted magnetic resonance imaging [MRI-DWI]) along with signs and symptoms over three days with repeat imaging where indicated. Screening was performed by a neuro-ophthalmologist. A positive HINTS examination indicated a central cause with absence of observable catch-up saccade on a bilateral head impulse test, presence of vertical nystagmus, absence of horizontal nystagmus, or upward or downward deviation of an uncovered pupil when focused on a point (skew deviation). A lack of all these features suggested a peripheral cause. An ABCD2 score with four or more risk factors was assumed to indicate possible stroke. The HINTS test had better test characteristics with sensitivity of 97% (95% CI, 92%–99%) and specificity of 99% (95% CI, 93%–99.9%), whereas the ABCD2 risk score only had sensitivity of 58% (95% CI, 49%–67%) and a specificity of 61% (95% CI, 49%–72%). Initial MRI alone had sensitivity of 87% when compared with the gold standard. A key limitation included the exclusion of those without stroke risk factors.

A 2011 small cohort study (N=24) examined the effectiveness of evaluating patients with a four-step oculomotor sign examination (head impulse test, nystagmus assessment, vertical smooth pursuit, and skew deviation) to distinguish vestibular neuritis from stroke.[2] Patients were adults who presented to the emergency department within 72 hours of an acute isolated vertigo event. Exclusion criteria included vestibular migraines, Horner syndrome, or other visual field defects. Participants had a mean age of 64 years old and 63% were men. MRI-DWI was used as the gold standard for diagnosing stroke. Two medical professionals were educated about neuro-otology with a three-hour video lecture and one-hour group tutorial before implementation. Sensitivity for the presence of any central signs was 100% (95% CI, 79%–100%) and absence of all signs had specificity of 90% (95% CI, 69%–100%). The study had a small sample size and examiners were not masked to the clinical history of patients. EBP

Kenneth Kenneth-Nwosa, MD
Chinyere Ogamdi, MD
Aubrey Palmer, MD
Susan C. Weller, PhD
University of Texas Medical Branch, Galveston, TX

The authors declare no conflicts of interest.

References

1. Newman-Toker DE, Kerber KA, Hsieh YH, et al. HINTS outperforms ABCD2 to screen for stroke in acute continuous vertigo and dizziness. *Acad Emerg Med.* 2013; 20(10): 986–996. [STEP 3]
2. Chen L, Lee W, Chambers BR, Dewey HM. Diagnostic accuracy of acute vestibular syndrome at the bedside in a stroke unit. *J Neurol.* 2011; 258(5):855–861. [STEP 3]

What is the least invasive approach to evaluation of asymptomatic lymphadenopathy in children?

EVIDENCE-BASED ANSWER

The least invasive approach to evaluation of asymptomatic lymphadenopathy is a thorough history and physical examination. Laboratory evaluation and chest x-ray can give additional information. Signs and symptoms that should prompt consideration for biopsy include node size greater than 2 cm, abnormal chest radiograph, systemic symptoms such as night sweats or weight loss, and supraclavicular adenopathy. In addition, elevated lactate dehydrogenase (LDH), C-reactive protein (CRP), and erythrocyte sedimentation rate (ESR) can be associated with malignancy (strength of recommendation [SOR]: C based on two retrospective cohort studies).

DOI 10.1097/EBP.0000000000001084

A 2006 retrospective cohort study (N=457) looked at characteristics associated with malignancy in patients less than 19 years old with peripheral adenopathy referred to the Pediatric Oncology Department at the Gazi University Medical School. On chart review, researchers collected patient age, sex, medical history, recent upper respiratory infection symptoms, travel history, and insect bite or animal exposure. Detailed physical examination data included location, size, and duration of lymph nodes. Laboratory evaluation included complete blood count, peripheral smear, ESR, CRP, LDH, uric acid, and numerous infectious serologies. All patients had a chest x-ray performed, one-third had an abdominal ultrasound performed, and 29% underwent excisional biopsy. Patients were 37% women, and ages ranged from two months to 19 years (median: seven years). Overall, 346 (76%) had benign disorders and 111 (24%) had malignancies. All lymph nodes less than 1 cm were benign. Of the malignant lymph nodes, 86% were greater than 3 cm and 14% were between 1 and 3 cm. Acute lymphadenopathy, defined as duration of less than four weeks, was benign 98% of the time. All supraclavicular lymph nodes identified were malignant. Chest x-ray was normal in 98% of the benign cases and 72% of the malignant cases. Of the abnormal chest x-ray findings, mediastinal lymphadenopathy was only seen in the malignant group. When abdominal ultrasound was performed, lymphadenopathy, hepatosplenomegaly, and the presence of a mass strongly correlated with malignancy (38%, 22%, 8%, respectively, of malignant group vs 1%, 0.2%, 0%, respectively, of benign group; $P<.0001$). Of the laboratory studies, LDH>430 IU/L, CRP>6 mg/L, and ESR>20 mm/L occurred more often in the malignant group (80%, 73%, 83%, respectively, of malignant group vs 12%, 23%, 26%, respectively, of benign group; $P<.0001$). Fever was more common in the benign group (49% benign vs 23% malignant; $P<.001$), whereas night sweats, weight loss, and hepatosplenomegaly were more common in the malignant group (24%, 31%, 30% malignant vs 6%, 4%, 6% benign; $P<.001$).[1]

A 1984 retrospective cohort study (N=163) looked at clinical variables that would appropriately select patients who would benefit from excisional biopsy and developed a model to assist with clinical decision making. The study identified 9 to 25-year-old patients who underwent lymph node biopsy or excision at the Hospital of the University of Pennsylvania and The Children's Hospital of Philadelphia. Exclusion criteria included previous biopsy that revealed histopathology, absence of palpable peripheral adenopathy, or lack of a medical record or pathology slides that could be reviewed. Cases in the derivation group (N=123) were divided into seven diagnostic categories (normal lymph node, reactive hyperplasia, miscellaneous [nonmalignant], granulomatous reaction, Hodgkin disease, non-Hodgkin lymphoma, and metastatic cancer). Categories 1 to 3 were considered the "no treatment" group for the purposes of the model and categories 4 to 7 were considered the "treatment" group (meaning would benefit from excisional biopsy). Patients' charts were reviewed for 22 clinical variables that could predict the need for "treatment" or "no treatment" and three clinical variables were ultimately identified who strongly correlated with the outcome: abnormal chest x-ray (16% no treatment group vs 66% treatment group) and size greater than 2 cm (18% no treatment group vs 62% of treatment group) correlated with the need for "treatment," whereas ear, nose, throat (ENT) symptoms (32% no treatment group vs 10% treatment group) correlated with "no treatment" needed ($P<.05$). In the predictive model, 5 points were assigned for abnormal chest x-ray, 3 points for size greater than 2 cm, and –3 points for ENT symptoms. An additional –2 points were used as a constant. The points were added to determine the final score. A total score of greater than zero indicated the "treatment" group and less than zero indicated the "no treatment"

group. When applied retrospectively to the derivation group, the model correctly assigned 95% of cases, exhibiting a sensitivity and positive predictive value of 95% and a specificity and negative predictive value of 96%. When the model was applied prospectively to 33 new cases, it allocated 32 of 33 patients appropriately (97%). History of night sweats and weight loss also were associated with granuloma or tumor ($P<.05$), but these factors were not included in the model. Study limitations included the small number of cases, and children younger than nine were not included in analysis.[2] EBP

Christina Zarza, MD
St. Anthony North Family Medicine Residency
Westminster, CO

The author declares no conflicts of interest.

References

1. Oguz A, Karadeniz C, Temel E, Citak E, Visal F. Evaluation of peripheral lymphadenopathy in children. *Pediatr Hematol Oncol.* 2006; 23(7):549–561. [STEP 3]
2. Slap G, Brooks J, Sanford Schwartz J. When to perform biopsies of enlarged peripheral lymph nodes in young patients. *JAMA*. 1984; 252:1321–1326. [STEP 3]

Does a sex education curriculum decrease unintended pregnancy in adolescents compared with an abstinence curriculum?

EVIDENCE-BASED ANSWER

Yes. Comprehensive sex education is associated with decreased adolescent pregnancy rates compared with abstinence curricula. The incidence of adolescent pregnancy seems to be higher with more emphasis on abstinence and lower when less emphasis is on abstinence (SOR: **B**, cohort and retrospective cohort trials).

DOI 10.1097/EBP.0000000000001051

A 2002 cohort study surveyed 1,179 never-married heterosexual adolescents 15 to 19 years old to compare incidence of pregnancy with comprehensive sex education (CSE) or abstinence-only sex education versus no formal sex education.[1] For abstinence-only education, the survey asked the participants if they received any formal instruction on how to say "no" to sex from school, church, or community center. For sex education exposure, the participants were asked if they received any instruction on birth control. Overall, 9.4% of adolescents stated that they had no sex education, 24% received abstinence only, and 67% received CSE. Pregnancy was reported in 7.3% of those surveyed. Significantly fewer adolescents in the sex education curriculum reported a pregnancy compared with no education (adjusted odds ratio [aOR] 0.39; 95% CI, 0.22–0.69), whereas no difference was noted between abstinence-only curriculum and no education (aOR 0.74; 95% CI, 0.38–1.5).

A 2011 retrospective study reviewing data from the Education Commission of the States evaluated the correlation between different levels of abstinence education state laws and teen pregnancy.[2] The level of emphasis on abstinence was graded on a scale from zero to three, with three having the highest emphasis on abstinence until marriage as part of sex education curriculum (21 states), level 2 promotes abstinence, but discussion of contraception is not prohibited (seven states), level 1 covers abstinence as part CSE with medically accurate education on contraception (11 states), and level 0 states do not have any specific mention of abstinence (nine states). Average teen pregnancy per every 1,000 girls age 14 to 19 years old was 73 for level 3 states, 62 per level 2 states, 56 for level 1 states, and 59 for level 0 states. After accounting for socioeconomic status, teen educational attainment, ethnic composition of the teen population, and availability of Medicaid waivers for family planning services in each state, the authors concluded the more abstinence education is emphasized in state laws and policies, the higher the teen pregnancy rates.

An ACOG Committee Opinion on adolescent sexuality education stated that CSE should be evidence based and medically accurate, including the benefit of delaying sexual intercourse, but also including contraception (no strength of recommendation provided).[3] EBP

Alexandra Targan, MD
Shannon Langner, MD
Jennifer Caragol, MD
University of Colorado FMR, Denver, CO

The authors declare no conflicts of interest.

References

1. Kohler PK, Manhart LE, Lafferty WE. Abstinence-only and comprehensive sex education and the initiation of sexual activity and teen pregnancy. *J Adolesc Health.* 2008; 42: 344–351. [STEP 3]
2. Stanger-Hall KE, Hall DW. Abstinence-only education and teen pregnancy rates: why we need comprehensive sex education in the U.S. *PLoS One.* 2011; 6(10):e24658. [STEP 4]
3. American College of Obstetricians and Gynecologists' Committee on Adolescent Health Care. Committee opinion no. 678: comprehensive sexuality education. *Obstet Gynecol.* 2016; 128(5):e227–e230. [STEP 5]

Is it safe to use another class of NSAID in patients allergic to ibuprofen?

EVIDENCE-BASED ANSWER

It is safe to use a selective cyclooxygenase-2 inhibitor and titrate to therapeutic dose in most patients allergic to ibuprofen. (SOR: **B**, based on two single-blinded, placebo-controlled oral challenges and one double-blinded placebo-controlled oral challenge).

DOI 10.1097/EBP.0000000000000755

A 2011 single-blinded, placebo-controlled oral challenge examined tolerance of etoricoxib, a selective cyclooxygenase-2 (COX-2) inhibitor, in 97 patients aged 14 years and older with a history of cross-intolerance between NSAIDs and aspirin (ASA).[1] Patients were evaluated at allergy clinics at five hospitals in Spain. Patients were given escalating doses of NSAIDs of increasing COX inhibition (acetaminophen being the weakest COX inhibitor used) or placebo. Patients were monitored for acute or worsening urticaria, as well as changes in blood pressure, heart rate, or peak expiratory flow. A total of 511 challenges were performed in 252 patients. Forty-seven were intolerant to acetaminophen (group A), and 205 were tolerant to acetaminophen but intolerant to other NSAIDs, including ibuprofen (group B). Tolerance to etoricoxib was assessed via challenge in all group A patients and a representative sample of 50 randomly selected group B patients. Of the 47 group A subjects, 12 (25.53%) showed signs of cross-intolerance, whereas only three (6%) of group B patients showed signs of cross-intolerance ($P<.03$). All reactions to etoricoxib were limited to mild cutaneous symptoms (pruritus, wheals), which resolved within one to two hours of taking an oral antihistamine.

A 2002 randomized controlled trial (n=60) examined the use of celecoxib, a selective COX-2 inhibitor, in 60 patients with asthma who demonstrated anti-inflammatory exacerbated respiratory disease (AERD).[2] A single-blinded oral challenge with placebo was performed to ensure that lung function was stable in all participating patients. Only patients in whom forced expiratory volume in one second (FEV_1) remained >70% of predicted value and hourly FEV_1 variance was <10%, were determined to have stable lung function and allowed to progress to the double-blinded celecoxib challenge. The day after the placebo challenge, patients were then given either 100 mg celecoxib or placebo at 7:00 AM and the other at noon. The next day, patients were given either 200 mg celecoxib or placebo at 7:00 AM and the other at noon. Hourly examinations of the nose, eyes, skin, and chest were performed to identify reactions. The final day, a single-blinded oral aspirin challenge was performed to confirm respiratory sensitivity anti-inflammatories. Sensitivity was defined as an FEV1 decline of ≥15% with concurrent oculonasal reaction or an FEV1 decline of ≥20% without oculonasal reaction. Of the 60 patients who completed the two-day celecoxib challenge and were confirmed to be aspirin sensitive, none experienced nasal symptoms or decline in FEV1. The authors calculated the probability of cross-reactivity to celecoxib in patients with AERD to be 0% to 5% (one-sided 95% CI, 0.00–0.05) via Wilcoxon's signed rank test. A significant limitation of this study is its small study size (n=60).

A 2004 single-blinded, placebo-controlled oral challenge examined the cross-tolerances of nimesulide, meloxicam, and rofecoxib in 140 patients with a history of aspirin and NSAID sensitivity living in Turkey.[3] Patients were selected for the study based on a reliable history of urticaria/angioedema, naso-ocular symptoms, bronchospasm, and/or anaphylactoid reaction to a prescribed ASA or NSAIDs. Asthmatic patients were deemed eligible for the study if their asthma had been stable for at least two weeks and their FEV1 was >70% predicted. One hundred twenty-

seven patients were challenged with nimesulide, 61 with meloxicam, 51 with rofecoxib, and 37 with all three drugs. Placebos were given to all patients on day 1 and then followed by one-fourth and three-fourth of therapeutic doses (nimesulide 100 mg, meloxicam 7.5 mg, rofecoxib 25 mg) of the active drug at 60-minute intervals on day 2 of the study. Patients were observed for allergic symptoms, changes in FEV1 or hypotension. The reaction rates of both nimesulide and meloxicam were found to be significantly higher than rofecoxib (29.7% vs 2.7%; $P<.0001$ and 10.8% vs 2.7%; $P=.048$). No patients reacted to placebo, and all patients with asthma tolerated rofecoxib without any adverse effects. The authors concluded that the highly-selective COX-2 inhibitor, rofecoxib, showed the most favorable tolerability among patients with aspirin and NSAID intolerance. Rofecoxib was withdrawn by the manufacturer from market in the United States in 2004. EBP

Dallas Chase, MD
Kyle Johnson, DO
Lauren Wilfling, DO
Mercy Family Medicine
St Louis, MO

The authors declare no conflicts of interest.

References

1. Doña I, Blanca-López N, Jagemann LR, et al. Response to a selective COX-2 inhibitor in patients with urticaria/angioedema induced by nonsteroidal anti-inflammatory drugs. *Allergy*. 2011; 66(11):1428–1433. [STEP 2]
2. Woessner KM, Simon RA, Stevenson DD. The safety of celecoxib in patients with aspirin-sensitive asthma. *Arthritis Rheum*. 2002; 46(8):2201–2206. [STEP 2]
3. Bavbek S, Celik G, Ozer F, et al. Safety of selective COX-2 inhibitors in aspirin/nonsteroidal anti-inflammatory drug-intolerant patients: comparison of nimesulide, meloxicam, and rofecoxib. *J Asthma*. 2004; 41(1):67–75. [STEP 2]

Are benzodiazepines with a longer half-life more effective in treating acute alcohol withdrawal compared with benzodiazepines with a shorter half-life?

EVIDENCE-BASED ANSWER

Probably not. A fixed-dose regimen with a shorter half-life benzodiazepine may have an equivalent or even a quicker rate of symptom improvement compared to benzodiazepines with longer half-lives (SOR: **C**, systematic review of randomized controlled trials [RCTs] and conflicting individual RCTs).

DOI 10.1097/EBP.0000000000000990

A 2010 cochrane review of 64 randomized controlled trials (RCTs) (N=4,309) examined the efficacy and safety of benzodiazepines for the treatment of alcohol withdrawal. RCTs on patients with alcohol dependence (as classified by *The Diagnostic and Statistical Manual of Mental Disorders* [DSM-IV]) who experienced alcohol withdrawal were included regardless of age, gender, nationality, and outpatient or inpatient therapy.[1] Primary outcomes included seizures, delirium, alcohol withdrawal symptoms as measured by pre-specified scales, global improvement of overall alcohol withdrawal syndrome as measured by pre-specified scales, and adverse events. Of those 64 RCTs, five compared outcomes between groups treated with benzodiazepines with longer half-lives (chlordiazepoxide, diazepam) with those treated with benzodiazepines with shorter half-lives (lorazepam). None of the comparisons reached statistical significance. The two trials that examined alcohol withdrawal seizures, both of which used fixed dose tapers, found no difference between benzodiazepines with longer half-lives compared with those with shorter half-lives (chlordiazepoxide vs lorazepam; 1 trial; N=50; risk ratio [RR], 0.2; 95% CI; 0.01–3.97 and diazepam vs lorazepam; 1 trial; N=40; RR, 3.0; 95% CI; 0.13–69.52). The authors of the review concluded that based on these comparisons, no strong evidence was found that certain benzodiazepines were more effective than others.

A 2013 double-blind RCT conducted in the addiction ward of a private hospital evaluated the efficacy of a fixed-dose chlordiazepoxide taper compared with a fixed-dose lorazepam taper in treating 108 male inpatients (mean age, 44.6 years) admitted with alcohol withdrawal (classified by the DSM-IV with clinical evidence of alcohol withdrawal with The Clinical Institute Withdrawal Assessment for Alcohol [CIWA]>8 on admission) without any other comorbidities.[2] Each arm received a fixed-dose regimen of a benzodiazepine that was tapered by 20% each day and discontinued by day five, delivered in four divided daily doses orally. One group received chlordiazepoxide (starting daily doses of 150 mg for The Clinical Institute Withdrawal Assessment for Alcohol, revised [CIWA-Ar]<15 and 200 mg for CIWA-Ar>15); the other group received lorazepam (starting daily doses of 6 mg for CIWA-Ar<15, 8 mg for CIWA-Ar>15). The primary outcomes were percentage improvement in CIWA-Ar after 48 hours and the duration of withdrawal. Patients were followed until their CIWA score was 0. The lorazepam arm had a significantly better mean rate of improvement at 48 hours compared with the chlordiazepoxide arm (70% vs 55%; $P<.001$) and a shorter mean duration of withdrawal (5.6 vs 6.7 days; $P<.0001$).

A 2015 double-blind, prospective, RCT in the inpatient wards of a teaching hospital in Bangalore evaluated the efficacy and safety of a fixed-dose chlordiazepoxide taper to a fixed-dose lorazepam taper in treating 60 patients (>18 years of age) admitted with mild-to-moderate alcohol dependence syndrome based on DSM-IV criteria, without any other comorbidities.[3] Each group received a fixed-dose regimen of a benzodiazepine that was tapered by 25% every two days and off by day eight, delivered in four divided daily doses orally. One group received chlordiazepoxide (starting at 80 mg/d); the other group received lorazepam (starting at 8 mg/d). No significant differences were noted in CIWA-Ar scores between the two groups on day eight ($P=.414$) or on day 12 ($P=.634$). The study also examined multiple components of liver function testing and found no difference in any components from baseline to day eight in either group. The application of this study is again limited given that a fixed-dose regimen was used rather than symptom-triggered therapy as well as the exclusion of patients with any other comorbidities.

A 2004 evidence-based practice guideline for the management of alcohol withdrawal delirium recommended benzodiazepines as the treatment of choice as opposed to other agents (grade A recommendation, supported by level I studies) and did not recommend any specific benzodiazepine.[4] The guideline stated

considerations such as time to onset, desired duration of action, comorbidities such as liver disease, and cost should guide the choice of benzodiazepine. EBP

Naomi Malam, MD
Jessica Devitt, MD
University of Colorado School of Medicine, Denver, CO

References

1. Amato L, Minozzi S, Vecchi S, Davoli M. Benzodiazepines for alcohol withdrawal. *Cochrane Database Syst Rev*. 2010; (3): CD005063. [STEP 1]
2. Rajmohan V, Sushil K, Mohandas E. A double blind randomized comparison of chlordiazepoxide and lorazepam in alcohol withdrawal. *Asian J Psychiatr*. 2013; 6(5):401–403. [STEP 2]
3. Ramanujam R, Padma L, Swaminath G, Thimmaiah RS. A comparative study of the clinical efficacy and safety of lorazepam and chlordiazepoxide in alcohol dependence syndrome. *J Clin Diagn Res.* 2015; 9(3):FC10–FC13. [STEP 2]
4. Mayo-Smith MF, Beecher LH, Fischer TL, et al. Management of alcohol withdrawal delirium: an evidence-based practice guideline. *Arch Intern Med*. 2004; 164(13):1405–1412. [STEP 2]