GEMs
GOOD EVIDENCE MATTERS

GEMs of the Week

Volume 4 - Issue 7

FPIN
Family Physicians Inquiries Network

What's in this week's issue?

Week of February 12 - 16, 2024

SPOTLIGHT: The Pills and Needles of Morning Sickness

The Pills and Needles of Morning Sickness

GEMs GOOD EVIDENCE MATTERS SPOTLIGHT

Acupuncture and Doxylamine-Pyridoxine for Nausea and Vomiting in Pregnancy: A Randomized, Controlled, 2x2 Factorial Trial

Wu XK, Gao JS, Ma HL, et al. Acupuncture and Doxylamine-Pyridoxine for Nausea and Vomiting in Pregnancy: A Randomized, Controlled, 2×2 Factorial Trial. *Ann Intern Med.* 2023;176(7):922-933. doi:10.7326/M22-2974

KEY TAKEAWAY: Doxylamine-pyridoxine and acupuncture (alone and together) are effective for treating moderate to severe nausea and vomiting in pregnancy.

STUDY DESIGN: Randomized, double-blinded, controlled trial

LEVEL OF EVIDENCE: STEP 2

BRIEF BACKGROUND INFORMATION: Nausea and vomiting in pregnancy complicate 10% of wanted pregnancies. The current practice of doxylamine-pyridoxine pharmacology is not well studied in clinical practice. Acupuncture frequently treats nausea and vomiting in pregnancy. In Asia, limited evidence supports this treatment option, including a systematic review with few studies, small sample sizes, high risk of bias, lack of proper controls, and inconsistent or limited treatment sessions.

PATIENTS: Women in early pregnancy with moderate to severe nausea and vomiting in pregnancy (NVP)

INTERVENTION: Daily active acupuncture and/or daily doxylamine–pyridoxine

CONTROL: Daily sham acupuncture and daily placebo

PRIMARY OUTCOME: NVP severity

Secondary Outcome: Quality of life, adverse events, maternal and perinatal complications

METHODS (BRIEF DESCRIPTION):

- Women in early pregnancy (7–14 weeks gestation) with moderate to severe NVP (PUQE scores 6–12 for moderate and >13 for severe), from 13 tertiary hospitals in mainland China from June 21, 2020, to February 2, 2022.
- Inclusion criteria: Maternal age 20–45 years old, viable intrauterine singleton or multiple pregnancy confirmed by ultrasound.
- Exclusion criteria: Weight loss greater than 20% compared to before conception, a disease associated with vomiting, conservative treatment was effective therefore not needing an acupuncture recommendation.
- Computer randomized groups for:
  - Daily 30-minute active acupuncture + daily doxylamine-pyridoxine
  - Daily 30-minute active acupuncture + daily placebo
  - Daily 30-minute sham acupuncture + daily doxylamine-pyridoxine
  - Daily 30-minute sham acupuncture + daily placebo
- For active acupuncture PC6 and ST36 were chosen as core acupoints with adjuvant acupoints (personalized per patient based on three traditional Chinese patterns) of CV12 for stomach deficiency, LR3 for liver heat, and ST40 for phlegm dampness.
- Patients received two tablets (10 mg doxylamine + 10 mg pyridoxine per tablet vs placebo) packaged by a commercial pharmaceutical group that were identical in color, size, odor, and taste.
  - The dosage was increased by one tablet if PUQE scores did not improve after two days with a maximum daily dosage of four tablets.
- The following was collected at baseline and on day 15: PUQE scores, weight difference, electrolyte levels, IV fluid treatment, use of additional medications, hospitalizations, termination of pregnancy, quality of life and well-being, pregnancy complications, treatment adherence, and neonatal outcomes.
  - PUQE score: 24 items with a 5-point scale, higher scores indicate more severe NVP
  - NVPQOL scale: 30 items with a 7-point scale, higher scores indicate worse quality of life
  - Zung self-rating depression scale: 20 items with a 4-point scale, higher scores indicate more severe depression
  - Zung self-rating anxiety scale: 20 items with a 4-point scale, higher scores indicate more severe anxiety

INTERVENTION (# IN THE GROUP):

- Active acupuncture + doxylamine-pyridoxine: 88
- Active acupuncture + placebo: 88
- Sham acupuncture + doxylamine-pyridoxine: 88

COMPARISON (# IN THE GROUP):

- Sham acupuncture + placebo: 88

FOLLOW-UP PERIOD: 15 days

RESULTS:

Primary Outcome –

- Patients who received a combination of both active acupuncture + doxylamine-pyridoxine had the largest reduction in episodes of emesis compared to the control group (mean difference [MD] –1.6; 95% CI, –2.2 to –0.9).
- Patients who received doxylamine–pyridoxine alone had a significant reduction in episodes of emesis compared to the control group (MD –1.0; 95% CI, –1.6 to –0.4).
- Patients who received active acupuncture had a significant reduction in episodes of emesis compared to the control group (MD –0.7; 95% CI, –1.3 to –0.1).

Secondary Outcome –

- Compared to placebo, patients who received doxylamine-pyridoxine had a significant improvement in:
  - Quality of life (MD –17; 95% CI, –27 to –7.8)
  - Depression (MD –3.5; 95% CI, –6.0 to –1.0)
  - Anxiety (MD –2.8; 95% CI, –5.1 to –0.4)
  - Global well-being (MD 1.0; 95% CI, 0.4–1.5)
- Compared to placebo, patients who received doxylamine-pyridoxine experienced more:
  - Maternal weight (MD 0.5 kg; 95% CI, 0.2–0.9 kg)
  - Somnolence (odds ratio [OR] 2.4; 95% CI, 1.4–4.0)
  - Dyspnea (risk difference 8.0%; 95% CI, 0.4–16%)
  - Itching and pain (OR 8.1; 95% CI, 1.8–36)
- Medication adherence was higher for doxylamine-pyridoxine compared with placebo (OR 3.7; 95% CI, 1.3–10).

LIMITATIONS:

- This study had a short intervention and follow-up period.
- Few patients in the "severe nausea and vomiting of pregnancy" group were intervened and evaluated.
- Another scoring system may have been better at evaluating treatment such as the HyperEmesis Level Prediction (HELP) score.
- This study did not include a true control for participants who had no treatment (observation only) for nausea and vomiting during pregnancy.
- It is unknown if sham acupuncture may have induced an active effect.
- Daily regimens and personalized acupoint protocols based on TCM patterns may not be as effective in different cultures or not feasible for everyone.

Xenia Brianna Gonzalez, MD
David Grant USAF Medical Center- Travis AFB
Fairfield, CA

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

Amitriptyline for IBS? Something to Consider

Amitriptyline at Low-Dose and Titrated for Irritable Bowel Syndrome as Second-Line Treatment in Primary Care (ATLANTIS): A Randomized, Double-Blind, Placebo-Controlled, Phase-3 Trial

Ford AC, Wright-Hughes A, Alderson SL, et al. Amitriptyline at Low-Dose and Titrated for Irritable Bowel Syndrome as Second-Line Treatment in Primary Care (ATLANTIS): a randomized, double-blind, placebo-controlled, phase 3 trial. *Lancet*. 2023;402(10414):1773-1785. doi:10.1016/S0140-6736(23)01523-4

KEY TAKEAWAY: Amitriptyline at low doses is more effective than a placebo and can be titrated safely to treat irritable bowel syndrome (IBS) symptoms for those who fail first-line treatments.

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

LEVEL OF EVIDENCE: STEP 2

BRIEF BACKGROUND INFORMATION: IBS is a functional bowel disorder that affects 5–10% of the global population and is often chronic as well as associated with psychiatric co-morbidity. When IBS first-line therapies are ineffective at symptomatic control, the use of tricyclic antidepressants as second-line treatment has been suggested but few studies with adequate size and power have studied their effectiveness and safety in the primary care setting.

PATIENTS: Adult patients with IBS who have failed first-line treatments with active symptoms

INTERVENTION: Amitriptyline

CONTROL: Placebo

PRIMARY OUTCOME: Global IBS symptoms

Secondary Outcome: Relief of IBS symptoms, somatic symptoms, mental health, ability to work, medication tolerability

METHODS (BRIEF DESCRIPTION):

- Patients were recruited from 55 general practices across three regions of England.
- Eligible patients were ≥18 years old with any IBS subtype who failed first-line treatments; had active symptoms (IBS-SSS score 75+) with normal CBC, CRP, negative anti-tTG IGA; no SI/HI; and if female, committed to using contraception.
- Exclusion criteria included subjects ≥61 years old not seen by their physician in the past 12 months; concern for lower GI cancer, celiac disease, or inflammatory bowel disease; previous colorectal cancer; participation in another study; pregnancy/breastfeeding/plans for pregnancy; and use of, or allergy to, a tricyclic antidepressant.
- Subjects were randomized 1:1 pairing via minimization methods and balanced by IBS subtype, HADS score, and recruitment hub. All personnel involved in direct trial conduct and analysis were fully masked.
    - Placebo and low-dose amitriptyline medications had identical appearance, packaging, and labeling.
    - All subjects started with 10 mg in week one and titrated up to 30 mg by week three.
- Outcomes were measured via baseline questionnaires that were repeated at three, six, and 12 months regarding:
    - Global IBS symptoms were measured by IBS-SSS (ranging from 75–500 points) with a mean difference of ≥35 points being clinically significant.
    - Relief of IBS symptoms was measured by the SGA of relief tool.
    - Weekly binary response to “Have you had adequate relief of your IBS symptoms?”
    - PHQ-12, HADS, WSAS, and ASEC checklist scores as adjuncts for somatic symptoms, mental health, ability to work, and medication tolerability, respectively.

INTERVENTION (# IN THE GROUP): 232

COMPARISON (# IN THE GROUP): 231

FOLLOW-UP PERIOD: Six months

RESULTS:

Primary Outcome –

- Low-dose amitriptyline improved global IBS symptoms compared to placebo (–27 points; 95% CI, –47 to –7.1).
    - The goal mean difference was not met but the 95% CI included –35 and excluded the chance that the intervention had no effect.

Secondary Outcome –

- Low-dose amitriptyline was shown to be superior compared to placebo for:
  - Relief of IBS symptoms at six months (odds ratio [OR] 1.78; 95% CI, 1.2–2.7)
  - Global IBS symptoms at three months (–23 points; 95% CI, –42 to –4.6)
  - Relief of IBS symptoms at three months (OR 1.70; 95% CI, 1.2–2.5)
  - >50% of weeks with adequate relief from IBS at six months (OR 1.56; 95% CI, 1.2–2.0)
- There was no significant change in somatic symptoms, mental health, ability to work, and medication tolerability for either group at three and six months.

LIMITATIONS:

- The population was UK-based, predominantly White (97%), and female (68%).
- IBS-C and IBS-U only accounted for 20% of the subtype groups, potentially affecting the effectiveness of low-dose amitriptyline for those subtypes.
- Amitriptyline has constipation as a side effect, which could have discouraged IBS-C patients from participating.
- Treatment duration was reduced from 12 to six months due to the COVID-19 pandemic.

Diana Tran, MD, MPH
Loyola MacNeal FMRP
Berwyn, IL

OMT for Plagiocephaly: Does it Work?

A Randomized Controlled Trial of Osteopathic Manipulative Therapy to Reduce Cranial Asymmetries in Young Infants with Nonsynostotic Plagiocephaly

Bagagiolo D, Priolo CG, Favre EM, et al. A Randomized Controlled Trial of Osteopathic Manipulative Therapy to Reduce Cranial Asymmetries in Young Infants with Nonsynostotic Plagiocephaly. *Am J Perinatol*. 2022;39(S 01):S52-S62. doi:10.1055/s-0042-1758723

KEY TAKEAWAY: Compared to light touch therapy (LTT), osteopathic manipulative therapy (OMT) significantly decreases the risk of nonsynostotic plagiocephaly (NSP) in infants as measured by a reduction in the oblique diameter difference index (ODDI) scores after three months and at one year old.

STUDY DESIGN: Prospective, single-site, two-parallel-arm, randomized controlled trial

LEVEL OF EVIDENCE: STEP 2

BRIEF BACKGROUND INFORMATION: NSP is an acquired form of cranial asymmetry that occurs in infants due to external mechanical forces. OMT has been studied in various other pediatric conditions, such as feeding issues, acute otitis media, and asthma, but its role in treating NSP has not been extensively studied. Given the prevalence of NSP and its associated parental anxiety, offering an effective treatment option in the office, such as OMT, can be highly beneficial.

PATIENTS: Infants with NSP

INTERVENTION: OMT

CONTROL: LTT

PRIMARY OUTCOME: ODDI score at three months

Secondary Outcome: ODDI score at one year

METHODS (BRIEF DESCRIPTION):

- 96 infants discharged from the Sant'Anna Hospital NICU in Turin, Italy who presented for follow-up with a diagnosis of NSP were included in the study.
- Inclusion Criteria: Infants between the ages of six months and one year old diagnosed with NSP who did not have previous intervention or congenital problems.
- Exclusion Criteria: Infants with synostotic plagiocephaly or other congenital problems (muscular torticollis, cerebral palsy, dysmorphism, etc.) and those who had previous osteopathic manipulation were excluded.
- Patients and parents were blinded and randomly assigned to one of the following treatment groups in a 1:1 ratio.
  - OMT group: Included 15 minutes of assessment and 30 minutes of treatment, plus repositioning therapy.
  - LTT group: Mimicked OMT without true manipulation of tissue and was used as a sham treatment, plus repositioning.
- Repositioning therapy is the standard management for NSP and parents/guardians in both groups were educated on strategies by pediatric physical therapists.
- Two different osteopaths administered either OMT or LTT over six different sessions and were blinded to the outcomes.
- The primary outcome was the ODDI score, which was used to measure the degree of NSP after three months of treatment.
  - This was measured using a plagiocephalometric tool with good intra/interrater reliability. Normal scores were <104%.
- The secondary outcome was the ODDI score at the follow-up at one year with a normal score being <104%.

INTERVENTION (# IN THE GROUP): 48

COMPARISON (# IN THE GROUP): 48

FOLLOW-UP PERIOD: Three months and at one year old

RESULTS:

Primary Outcome –

- There was a statistically significant increase in the proportion of infants in the OMT group with a reduced incidence of NSP compared to the LTT group at three months (absolute risk reduction [ARR] 0.41; 95% CI, 0.25–0.53).

Secondary Outcome –

- There was a statistically significant increase in the proportion of infants in the OMT group with a reduced incidence of NSP compared to the LTT group at one year (ARR 0.47; 95% CI, 0.31–0.64).

LIMITATIONS:

- The two intervention groups were not identical in terms of birth weight, gestational age, birth presentation, and severity of NSP.
- Study infants were former NICU patients and not recruited from community practices.
- There was no attempt to determine the type and frequency of repositioning done at home.
- Follow-up was limited at one year.

Kruti Patel, DO
Loyola MacNeal FMRP
Berwyn, IL

Do Interventional Weight Loss Programs Prevent Gestational Diabetes Recurrence?

Randomized Controlled Trial of Pre-Pregnancy Lifestyle Intervention to Reduce Recurrence of Gestational Diabetes Mellitus

Phelan S, Jelalian E, Coustan D, et al. Randomized controlled trial of prepregnancy lifestyle intervention to reduce recurrence of gestational diabetes mellitus. *Am J Obstet Gynecol.* 2023;229(2):158.e1-158.e14. doi:10.1016/j.ajog.2023.01.037

KEY TAKEAWAY: Pre-pregnancy intervention increases weight loss but does not significantly impact gestational diabetes mellitus (GDM) recurrence.

STUDY DESIGN: Randomized controlled trial, dual-site

LEVEL OF EVIDENCE: STEP 3 (downgraded due to low power)

BRIEF BACKGROUND INFORMATION: Gestational diabetes often reoccurs in subsequent pregnancies and impacts 7.8% of all pregnancies. Observational studies indicate that weight loss and physical activity before and during subsequent pregnancies can decrease GDM rates in later pregnancies. Randomized clinical trials are needed, particularly on how primary care providers can intervene. Due to the dramatic impact that GDM has on both mother and fetus/infant, during and after pregnancy, engaging in preventative studies is necessary.

PATIENTS: Women during preconception and prenatal periods with a previous history of GDM

INTERVENTION: Weight loss intervention + standard care + education

CONTROL: Standard care + education

PRIMARY OUTCOME: GDM recurrence in subsequent pregnancy

Secondary Outcome: Glucose, blood pressure

METHODS (BRIEF DESCRIPTION):

- 199 randomized English and Spanish-speaking overweight or obese females with eligibility based on: Verified GDM in a previous pregnancy, reported plans of becoming pregnant in the next 1–3 years, BMI >25kg/m2, >18 years old, and >3 months postpartum.
- The intervention group received standard care, education, and weight loss intervention, including 16 weeks with weekly 30-minute meetings and weight loss maintenance through biweekly meetings thereafter until conception.
- The control group received standard care before and during pregnancy and education about pre-pregnancy health at the beginning of the study and 16 weeks into the study.
- All participants were evaluated at baseline, 16 weeks gestation, 26 weeks gestation, and six weeks postpartum.
- The primary outcome was GDM diagnosis according to Carpenter and Custard criteria with two abnormal values on three 3-hour oral glucose tests in subsequent pregnancy. In most cases, diagnoses were based on provider assessment, not study measured.

INTERVENTION (# IN THE GROUP): 105 (66 excluded due to no viable pregnancy)

COMPARISON (# IN THE GROUP): 94 (69 excluded due to no viable pregnancy)

FOLLOW-UP PERIOD: Through six weeks postpartum

RESULTS:

Primary Outcome –

- Weight loss interventions had no significant effect on GDM recurrence compared to control (odds ratio [OR] 1.8; 95% CI, 0.59–5.8).

Secondary Outcome –

- No significant group x time interaction was found for glucose.
- There was a group x time interaction for diastolic blood pressure with an improvement in the intervention group after 16 weeks of 7.1 (—13.0 to —1.3 mmHg decrease; *P*=.02).

LIMITATIONS:

- Very low rates of pregnancy in the 199 randomized participants.
- The study is not population-wide, so rates and benefits of GDM recurrence in intervention populations are not fully known.
- Small sample size.

Joseph Call, MD
University of Wyoming FMR
Casper, WY

AI Is Here to Stay: How It Can Help Us All Day

Diagnostic Accuracy of Differential-Diagnosis Lists Generated by Generative Pretrained Transformer 3 Chatbot for Clinical Vignettes with Common Chief Complaints: A Pilot Study

Hirosawa T, Harada Y, Yokose M, Sakamoto T, Kawamura R, Shimizu T. Diagnostic Accuracy of Differential-Diagnosis Lists Generated by Generative Pretrained Transformer 3 Chatbot for Clinical Vignettes with Common Chief Complaints: A Pilot Study. *Int J Environ Res Public Health.* 2023;20(4):3378. Published 2023 Feb 15. doi:10.3390/ijerph20043378

KEY TAKEAWAY: Artificial intelligence (AI), such as ChatGPT, can create accurate differential diagnoses for common conditions, however physician differential diagnoses are superior.

STUDY DESIGN: Cohort study

LEVEL OF EVIDENCE: STEP 3

BRIEF BACKGROUND INFORMATION: AI is increasingly used worldwide to improve patient outcomes with physician guidance. However, there is limited research on AI's effectiveness in healthcare triage, even though ChatGPT-3 outperforms other AI systems with higher rates of top differentials.

PATIENTS: Clinical vignettes

INTERVENTION: AI (ChatGPT-3) number of differential diagnoses

CONTROL: Internal medicine physicians

PRIMARY OUTCOME: Accuracy of differential diagnoses

METHODS (BRIEF DESCRIPTION):

- Two independent co-researcher general internal medicine physicians at the Department of Dokkyo Medical University Hospital in Japan were a part of this study, however, it is unclear if they were board-certified.
- 30 clinical vignettes were created based on 10 common chief complaints.
- A researcher typed "Tell me the top 10 suspected illnesses for the following symptoms: (added each clinical vignette)" into ChatGPT. This was the standardized delivery of the chief complaints.
- Each vignette had one ideal correct answer and a binary approach to scoring was utilized.
    - If the differential diagnosis contained the correct answer a point of 1 was assigned. If the answer was not represented a score of 0 was assigned.
- Aggregate differentials between a total of two physicians for all 30 vignettes were averaged and compared to ChatGPT scores.
    - High scores clinically mean a high rate of correct diagnosis generated by either the physician or ChatGPT.
    - Low scores clinically mean a lower rate of correct diagnosis by physicians or ChatGPT.
    - A higher aggregate score represented an increased rate of correct diagnoses.

INTERVENTION (# IN THE GROUP): Not applicable

COMPARISON (# IN THE GROUP): Not applicable

FOLLOW-UP PERIOD: Not applicable

RESULTS:

Primary Outcome –

- Physicians were significantly more accurate in diagnosing the top diagnosis compared to AI (93.3% vs 53.3%; p<.001).
- Physicians were significantly more accurate in providing the top five diagnoses compared to AI (98.3% vs 83.3%; p=.03).
- However, physicians underperformed in diagnosing the top 10 diagnoses compared to AI (58.7% vs 70.5%).

LIMITATIONS:

- ChatGPT limitations included biased unreliable data depending on the programmer.
- Vignettes were not based on real cases so there is a lack of broad applicability to different populations.
- ChatGPT had no access to EMRs, so they had limited accessible information.

Idiata Iredia, MD
Capital Health FMRP
Pennington, NJ