GEMs
GOOD EVIDENCE MATTERS

GEMs of the Week

Volume 5 - Issue 21

FPIN
Family Physicians Inquiries Network

What's in this week's issue?

Week of June 9 - June 12, 2025

SPOTLIGHT:
Breaking Up with IBS: It’s Not You, It’s Your Microbiome!

Breaking Up with IBS: It's Not You, It's Your Microbiome!

GEMs GOOD EVIDENCE MATTERS SPOTLIGHT

The Role of Gut-Microbiota in the Pathophysiology and Therapy of Irritable Bowel Syndrome: A Systematic Review

Shrestha B, Patel D, Shah H, et al. The Role of Gut-Microbiota in the Pathophysiology and Therapy of Irritable Bowel Syndrome: A Systematic Review. *Cureus*. 2022;14(8):e28064. Published 2022 Aug 16. doi:10.7759/cureus.28064

KEY TAKEAWAY: Gut dysbiosis contributes to irritable bowel syndrome (IBS) symptom severity, and microbiota-directed therapies show potential for symptom relief, though the evidence remains limited.

STUDY DESIGN: Systematic review of five review articles, two randomized control trials, two case-control studies, and one cross-sectional study (N=329)

LEVEL OF EVIDENCE: STEP 3 (downgraded due to lack of statistical analysis)

BRIEF BACKGROUND INFORMATION: IBS is a chronic gastrointestinal (GI) disorder characterized by abdominal pain and changes in bowel movements. While there are no specific diagnostic tests, treatment focuses on symptom management through medications, dietary changes, and lifestyle adjustments. This study aimed to explore the connection between gut microbiome imbalances and IBS pathophysiology and severity.

PATIENTS: Patients with IBS

INTERVENTION: Microbiota directed therapies

CONTROL: Usual care, healthy controls, placebo

PRIMARY OUTCOME: IBS severity, microbiota composition, quality of life, psychological symptoms

METHODS (BRIEF DESCRIPTION):

- Human studies published in English between 2017–2022 were included in the review.
- Patients were diagnosed with IBS based on the Rome IV criteria, and the severity of the disease ranged from mild to severe.
- The Rome IV diagnostic criteria includes recurrent abdominal pain, at least one day per week in the last three months associated with >2 of the following: Pain improved or worsened by bowel movements, change in stool frequency, and change in stool form.
    - Symptom onset at least six months prior to diagnosis
    - Symptoms not explained by other medical conditions
    - Severity based on patient-reported GI and extraintestinal symptoms, degree of disability, and illness-related perceptions and behaviors
- Gut dysbiosis is defined as an imbalance of the gut microbiome characterized by loss of baseline biome diversity, overgrowth of pathogenic organisms, and a reduction in beneficial microbes.
- The primary outcomes were measured mainly through clinical assessments (anxiety, depression, IBS severity), microbiota profiling, and self-reported questionnaires regarding psychological symptoms and quality of life.
- Microbiota-directed therapies such as prebiotics, probiotics, synbiotics, and rifaximin were reviewed.

INTERVENTION (# IN THE GROUP): Not available

COMPARISON (# IN THE GROUP): Not available

FOLLOW-UP PERIOD: Not available

RESULTS:

Primary Outcome –

- Gut dysbiosis increased IBS severity compared to balanced microbiota and healthy microbiota composition (no numerical or statistical results reported).
- Enteric infections such as Norwalk virus, *Clostridium dificile*, *Campylobacter*, and *Giardia* increased the risk of IBS (no numerical or statistical results reported).
- Gut dysbiosis reduced patient's quality of life compared to balanced microbiota and healthy microbiota composition (no numerical or statistical results reported).
- Depression and anxiety increased the risk of IBS, and conversely, IBS increased the risk for development of depression or anxiety (no numerical or statistical results reported).
- *Lachnospiraceae*, *Bacteroidaceae*, and *Ruminococcaceae* were linked to IBS symptoms such as diarrhea, anxiety, and depression (no numerical or statistical results reported).

- Microbiota-directed therapies, including probiotics and prebiotics, improved IBS symptoms, psychological symptoms, and quality of life compared to placebo and controls (no numerical or statistical results reported).
- Prebiotics improved microbiota composition compared to placebo and controls (no numerical or statistical results reported).

LIMITATIONS:

- The review lacked detailed statistical analysis, limiting interpretation of significance.
- Only English language studies were included.
- Few randomized control trials and longitudinal studies were available.
- No standardized protocols were used for microbiome-based therapies.

Ethan Gordon, MD
Central Michigan University FMRP
Saginaw, MI

Arm Position: Is There a Difference During Blood Pressure Readings?

Arm Position and Blood Pressure Readings: The ARMS Crossover Randomized Clinical Trial

Liu H, Zhao D, Sabit A, et al. Arm Position and Blood Pressure Readings: The ARMS Crossover Randomized Clinical Trial. *JAMA Intern Med*. 2024;184(12):1436-1442. doi:10.1001/jamainternmed.2024.5213

KEY TAKEAWAY: Nonstandard arm positions such as arm at the side or on the lap during blood pressure (BP) measurement artificially inflate BP readings.

STUDY DESIGN: Randomized, crossover clinical trial

LEVEL OF EVIDENCE: STEP 2

BRIEF BACKGROUND INFORMATION: BP is a vital health metric for the diagnosis and treatment of hypertension (HTN). Accurate BP measurement with arm support on a desk is recommended. However, nonstandard arm positions such as arm on the side or lap are common in busy primary care offices and measurements taken at home. Few studies have looked rigorously at standard vs nonstandard BP measurement, and even fewer with randomization. Inaccurate BP measurement can result in misdiagnosis of HTN and the erroneous use of the readings in common risk stratification calculators before starting patients on adjunct medications.

PATIENTS: Adults 18–80 years old

INTERVENTION: BP readings taken with unsupported arm at the side and on the lap

CONTROL: BP reading taken with arm supported on a desk

PRIMARY OUTCOME: Systolic BP (SBP) and diastolic BP (DBP)

Secondary Outcome: SBP and DBP stratified by healthcare use and BP

METHODS (BRIEF DESCRIPTION):

- Participants were recruited from BP screening at a public food market, direct mailing, informational brochure, and physician recommendations.
- Mean age of participants was 57 years old and 53% were women.
- Participants with rashes, gauze dressings, casts, edema, paralysis, tubes, open sores or wounds, or arteriovenous shunts on both arms; mental impairment; pregnancy; or a mid-upper arm circumference of >55 cm were excluded from the study.
- Each participant received a total of 12 BP readings as follows:
  - Three sets of three BP readings were taken in random order, comprising arm supported on a desk with midcuff at midheart level (desk 1), hand supported on the lap (lap), and arm unsupported on the side (side).
  - The 4th set of BP readings was taken in the same method as desk one (called desk 2).
  - Prior to each set of BP readings, the participant took a two-minute walk and a five-minute seated rest.
- A validated oscillometric BP device was used.
- Participants were randomized into one of six groups for order of the arm position BP reading.
- Participants provided self-reported metrics that included demographics, medical history, height and weight from which body mass index [BMI] was calculated, use of antihypertensive medication, and any healthcare use in the past year.
- The outcomes were assessed by examining the difference in mean SBP and DBP between the lap and desk one and side and desk one compared to the difference between desk two and desk one:
  - Lap arm position (lap–desk 1) – (desk 2–desk 1)
  - Side arm position (side–desk 1) – (desk 2–desk 1)

INTERVENTION (# IN THE GROUP): 133

COMPARISON (# IN THE GROUP): 133

FOLLOW-UP PERIOD: Not applicable

RESULTS:

Primary Outcome –

- The lap arm position showed significantly higher BP readings for both SBP and DBP compared to desk readings:
  - SBP (mean difference [MD] 3.9 mmHg; 95% CI, 2.5–5.2)
  - DBP (MD 4.0 mmHg; 95% CI, 3.1–4.9)
- The side arm position showed significantly higher BP readings for both SBP and DBP compared to desk readings:
  - SBP (MD 6.5 mmHg; 95% CI, 5.1–7.9)

- DBP (MD 4.4 mmHg; 95% CI, 3.4–5.4)

Secondary Outcome –

- The lap arm position showed significantly larger SBP difference among those with no healthcare use compared those with healthcare use in the past one year:
  - No healthcare use (MD 9.6 mmHg; 95% CI, 6.4–13)
  - Healthcare use (MD 3.6 mmHg; 95% CI, 2.0–5.2)
- The side arm position showed a significantly larger SBP difference among those with SBP ≥130 mmHg compared to those with <130 mmHg:
  - SBP ≥130 (MD 8.5 mmHg; 95% CI, 5.7–11)
  - SBP <120 (MD 5.3 mmHg; 95% CI, 3.8–6.9)

LIMITATIONS:

- Unequal randomization of participants to each of the six groups, although sensitivity analysis showed similar results as the planned analysis.
- The study was conducted in one geographical area.
- The sample sizes of some of the subgroups were small.
- Due to the use of the automated BP device, the generalizability of the results to other BP devices was uncertain.

Reece Tuckerman, MD
Womack Army Medical Center FMRP
Fort Bragg, NC

The views expressed herein are those of the author and do not necessarily reflect the official policy of the Department of the Army, Defense Health Agency, Department of Defense, or the U.S. Government.

Is 15 the New 10? Longer Interval for Screening Colonoscopy

Longer Interval Between First Colonoscopy with Negative Findings for Colorectal Cancer and Repeat Colonoscopy. A Retrospective Cohort Study.

Liang Q, Mukama T, Sundquist K, et al. *JAMA Oncol.* 2024;10(7):866–873. doi:10.1001/jamaoncol.2024.0827I

KEY TAKEAWAY: After first negative colonoscopy, waiting up to 15 years for repeat colonoscopy appears to be as good as repeating colonoscopy at 10 years.

STUDY DESIGN: Retrospective cohort study

LEVEL OF EVIDENCE: STEP 3

BRIEF BACKGROUND INFORMATION: Colonoscopy is one of the preventive screening options for colorectal cancer (CRC) and is recommended every 10 years starting at age 45-50 irrespective of ethnicity, race, or gender. This study used a large dataset to evaluate whether a longer colonoscopy interval time after a first negative colonoscopy in individuals without a family history of CRC would change patient outcomes.

PATIENTS: Adults 45-69

INTERVENTION: Negative initial colonoscopy

CONTROL: No screening colonoscopy

PRIMARY OUTCOME: CRC diagnosis or CRC-Specific death

METHODS (BRIEF DESCRIPTION):

- The authors conducted a retrospective cohort study using data from Swedish national registries from 1990 to 2018.
- Patients were excluded who had a colonoscopy prior to 1990; had a diagnosis of CRC, polyp, adenoma, or carcinoma in situ prior to 1990; had a sigmoidoscopy or proctoscopy before the end of the follow up period; had positive findings on initial colonoscopy; or had a family history of CRC or inflammatory bowel disease.
- Patients aged 45-69 were included in the exposed group if they had a negative initial colonoscopy between 1990 and 2016.
- All patients aged 45-69 were eligible to be part of the control group until their first colonoscopy, sigmoidoscopy or proctoscopy prior to end of the follow up, or were diagnosed with CRC, polyp, adenoma, or carcinoma in situ before 1990
- The authors matched controls to exposed patients in an 18:1 ratio by gender, birth year +/- 2 years, and baseline.
- Both male and female groups and both exposed and control groups had an average age of 59 years. Most identified as female (59.2 %) and there were no significant differences in socioeconomic status, educational attainment, or geographic region.
- The end points for participants in both groups included precancerous findings, CRC diagnosis or carcinoma in situ, receiving other endoscopic procedures, death, emigration, or a family member being diagnosed with CRC.
- Follow-up was halted for all subjects at the end of the year prior to receiving a diagnosis of a precancerous lesion, to avoid the influence of any additional preventative procedures on risk estimates.
- Cumulative risk was calculated as the total risk of an event occurring over a specific period, reported as the ten-year Standardized Incidence Ratios (SIR) for CRC diagnosis, and Standardized Mortality Ratios (SMR) for deaths.
- Analysis adjusted for birth year, sex, socioeconomic status, area of residence, level of job and level of education residence, and level of education.
    - SIR compares the observed number of disease cases in a specific population to the number of cases expected based on a reference population (controls).
    - SMR compares the observed number of deaths in a study population to the number of deaths expected based on the age and sex distribution of a standard population.

INTERVENTION (# IN THE GROUP): 110,074

COMPARISON (# IN THE GROUP): 1,981,332

FOLLOW-UP PERIOD: 29 years

RESULTS:

- Primary Outcome
    - During the follow up period, there were fewer CRC cases and CRC specific deaths in the exposed vs. control group (CRC cases: 0.44% vs. 1.1% and CRC-specific deaths: 0.10% vs. 0.28%), respectively.
    - As compared to the control group, exposure to a negative initial colonoscopy was associated with a significantly lower 10-year cumulative

risk of CRC diagnosis (SIR = 0.72; 95% CI, 0.54-0.94) through year 15.
    - As compared to the control group, exposure to a negative initial colonoscopy was associated with a significantly lower 10-year cumulative risk of CRC-specific deaths (SMR= 0.55;95% CI, 0.29-0.94) through year 15.

- Secondary Analysis
    - Duration of cumulative CRC risk reduction differed by gender, with risk remaining low for 14 years for women vs 15 years for men.
    - Extending the time interval between first and second colonoscopy from 10 to 15 years in exposure group (compared to the control group) was estimated to result in 2.4 more CRC cases and 1.4 more CRC-specific deaths per 1000 individuals.
    - Extending the interval between colonoscopies from 10 to 15 years would allow each person in exposure group to avoid one colonoscopy over their lifetime.

LIMITATIONS:

- Study is based on only one national population (Swedish) which is predominantly White, limiting generalizability.
- No data on use of stool-based screening tests was available.
- Data was not adjusted for some confounders such as colonoscopy quality, including bowel preparation or adenoma detection rates.
- Accuracy and validity of colonoscopy codes in Swedish medical registries has not been specifically established, though is likely to be high.

Sumanth Danda, MD
Eastern Main Medical Center Family Medicine Residency Program
Bangor, ME

Can AI Revolutionize Dermatological Lesion Detection?

Evaluation of an artificial intelligence-based decision support for the detection of cutaneous melanoma in primary care: a prospective real-life clinical trial

Papachristou P, Söderholm M, Pallon J, Taloyan M, Polesie S, Paoli J, Anderson CD, Falk M. Evaluation of an artificial intelligence-based decision support for the detection of cutaneous melanoma in primary care: a prospective real-life clinical trial. Br J Dermatol. 2024 Jun 20;191(1):125-133. doi: 10.1093/bjd/ljae021. PMID: 38234043.

KEY TAKEAWAY: The Dermalyser® app supports melanoma detection in primary care by effectively identifying benign lesions as non-melanoma. However, it also flags many non-melanoma lesions as suspicious, which may lead to unnecessary follow-up or referrals.

STUDY DESIGN: Prospective cohort study

LEVEL OF EVIDENCE: STEP 3

BRIEF BACKGROUND INFORMATION: Primary care physicians (PCPs) are often the first evaluators of skin lesions in patients, though their diagnostic accuracy is lower than that of dermatologists. Few studies have investigated artificial intelligence (AI) tools on patients with skin lesions in real time clinical settings.

PATIENTS: Adults with at least one skin lesion with any suspicion of melanoma

INTERVENTION: AI -based Smartphone app Dermalyser®

CONTROL: Clinical or histopathological diagnosis

PRIMARY OUTCOME: Diagnostic accuracy

METHODS (BRIEF DESCRIPTION):

- Included were adults >18 years old with ≥1 skin lesions suspicious of melanoma in 36 Swedish primary care centers.
- Excluded were adults with lesions related to other malignancy, damaged or tattooed skin, located on inaccessible body spot, covered by dense hair, poor quality dermoscopic images, melanin rich skin type (Fitzpatrick skin type V-VI).
- The Dermalyser® app evaluated each lesion and classified it as either "melanoma" or "none detected." All lesions were also assessed via teledermoscopy using polarized light contact dermoscopes for suspected skin tumors.
- The Dermalyser® app classification of the lesion was compared to the final clinical diagnosis or, when available, the histopathological diagnosis following excision.
- Treating physicians excised lesions based on standardized clinical criteria.
- Final diagnoses were determined by comparing the physician's clinical assessment with histopathological findings, when histology was available.

INTERVENTION (# IN THE GROUP): 228

COMPARISON (# IN THE GROUP): The same 228 patients

FOLLOW-UP PERIOD: Not available

RESULTS:

Primary Outcome –

- The AI-based smartphone application improved diagnostic accuracy for all melanomas compared to the clinical or histopathological diagnosis (area under the receiver operating curve [AUROC] 0.96; 95% CI, 0.93–0.98).
  - Sensitivity: 95%
  - Specificity: 85%
  - Positive Predictive Value (PPV): 36%
  - Negative Predictive Value (NPV): 99.5%
- The AI-based smartphone application improved diagnostic accuracy for invasive melanomas compared to the clinical or histopathological diagnosis (AUROC 0.99; 95% CI, 0.97–1.0).
  - Sensitivity: 100%
  - Specificity: 93%
  - PPV: 38%
  - NPV: 99%

LIMITATIONS:

- The study did not incorporate comprehensive patient and family medical history, which are crucial for assessing melanoma risk.
- Limited ability to view lesion changes over time could reduce the ability to detect melanoma indicators and result in missed diagnoses.
- Data quality could affect the accuracy and lead to reduced reliability of the app.
- The AI's diagnosis in a test setting can affect how useful it is in real life, possibly causing over-reliance.

Brian Chung, MD
Northside Hospital – Gwinnett
Lawrenceville, GA

Effect of Yoga on Frailty in Older Adults: A Systematic Review

Loewenthal J, Innes KE, Mitzner M, Mita C, Orkaby AR. Effect of Yoga on Frailty in Older Adults : A Systematic Review. *Ann Intern Med*. 2023;176(4):524-535. doi:10.7326/M22-2553

KEY TAKEAWAY: Effects of yoga may in frailty of elderly may be similar to effects seen with active interventions like exercise.

STUDY DESIGN: Systematic review of 33 randomized controlled trials (RCTs) (N=2,384)

LEVEL OF EVIDENCE: STEP 2 (downgraded due to high heterogeneity and lack of meta-analysis)

BRIEF BACKGROUND INFORMATION: Current options available for improving frailty in older adults include exercise, nutrition, cognitive exercise, and education. This study assessed how yoga compares to both inactive treatment options like education and active treatment options when looking at markers of frailty in older adults. This can change medical management, by providing a low impact option that can be more palatable and achievable for patients, making it more likely they will do it.

PATIENTS: Adults ≥65 years old

INTERVENTION: Yoga-based interventions

CONTROL: Active control, usual care, active control, and active education

PRIMARY OUTCOME: Gait speed, timed up and go test, six-minute walk test, handgrip strength, serum albumin, balance, lower extremity strength and endurance.

METHODS (BRIEF DESCRIPTION):

- Studies from the US, Europe, Australia, Asia were included in the review. The health of participants ranged from community dwelling to patients in nursing homes.
- Eligible studies had a mean participant age of ≥65 years old and contained objective outcomes that had been validated as markers of frailty which include gait speed, timed up and go test, six-minute walk test, handgrip strength, serum albumin, balance, and lower extremity strength and endurance.
- Cross-sectional studies, uncontrolled trials; case series, and case studies; studies published only in dissertation or abstract form; and those evaluating mindfulness-based stress reduction were excluded from the review.
- Interventions included different styles of yoga ranging from power yoga to hatha yoga.
    - Treatment ranged from <8 weeks to >24 weeks.
    - Sessions ranged from 30-90 minutes, and from 1-4 times a week.
- Various controls were included active control, usual care, active control, and active education.
- The primary outcomes were measured by assessing factors including:
    - Gait speed which consisted of the six-minute walk (6MW) test, 6MW distance, timed up and go, 8 foot up and go, usual and/or maximal gait speed
    - Handgrip strength
    - Balance (Burg Balance Scale [BBS], single-leg standing time, Tinetti balance assessment tool, Short Physical Performance Battery [SPPB], mini-balance evaluation systems test, balance error scoring system, posturography, sharpened Romberg) Lower extremity strength and endurance
    - Multicomponent physical performance measures (SPPB, functional gait assessment)

INTERVENTION (# IN THE GROUP): Not available

COMPARISON (# IN THE GROUP): Not available

FOLLOW-UP PERIOD: Varied (<8 weeks to >24 weeks)

RESULTS:

Primary Outcome –

- Gait speed (6MWT or 6-minute walk distance, timed up and go, 8 foot up and go, usual and/or maximal gait speed) – total 25 studies
    - 11 studies against active control - 2 studies showed clinically significant differences
    - 7 studies against education - 2 studies showed clinically significant differences
    - 12 studies against inactive control – 3 studies showed clinically significant differences
- Hand grip – total 3 studies
    - 3 studies assessed hand grip strength in 3 or 2 group design in community dwelling,

nursing home, and chronic back pain populations; no studies reported clinically meaningful change in HGS

- Balance (BBS, single-leg standing time, Tinetti balance assessment tool, SPPB, mini-balance evaluation systems test, balance error scoring system, post urography, sharpened Romberg) – 29 studies
    - 13 studies against active control no statistical or clinically significant improvement in standing balance postintervention
    - 5 studies against education -> 2 studies were statistically significant button clinically meaningful
    - 10 studies against inactive control -> 1 study showed clinically meaningful improvement
- Lower extremity strength and endurance (30CST, 5 times sit to stand test, or component of SPPB) -> 13 studies
    - In 5 studies against active control -> 1 was statistically significant but not clinically meaningful
    - In 4 studies against education -> 2 studies were statistically significant but not clinically meaningful improvements
    - In 6 studies against inactive control, 3 studies were statistically significant but not clinically meaningful improvements.
- Multicomponent physical performance measures (SPPB, functional gait assessment) -> 5 studies
    - In 1 study against active control -> no change
    - In 1 study against education -> no change
    - In 3 studies against inactive control -> no change

LIMITATIONS:

- Findings were summarized narratively rather than through pooled analysis, limiting the ability to generate a precise estimate of effect and reducing statistical power compared to formal meta-analysis.
- Limitations in the review include heterogeneity in study design, population, yoga style, and small sample sizes with 61% of trials including 50 or fewer participants.
- Deficiencies in reporting include varied description of randomization process, concealment of treatment allocation, and lack of blinding procedures.
- Racial and ethnic backgrounds of populations were not noted in most of the studies so unclear if racial and ethnic backgrounds contribute to the results.

Shannon Yee, DO
Abrazo Central FMRP
Phoenix, AZ