Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Translating Best Evidence into Best Care

EDITOR'S NOTE: Studies for this column are identified using the Clinical Queries feature of PubMed, "hand" searching *JAMA*, *JAMA Pediatrics*, *Pediatrics*, *The Journal of Pediatrics*, *The New England Journal of Medicine*, AAP daily briefing, and from customized EvidenceAlerts.

EBM PEARL: CHI-SQUARE (CS) TEST: The CS test detects differences between categorical variables. Categorical variables have 2 or more possible, fixed values, such as yes/no-type variables (eg, COVID-19–yes or no), or political parties, but not continuous variables such as blood pressure measurements. A statistical way of describing the CS test is, it detects whether categorical variables are independent or dependent, ie, corresponding to accepting or rejecting the null hypothesis (no difference between variables), respectively. Here is an example from the Hoberman et al randomized study (see commentary below).

Data (with additions) taken from the demographics table in the study by Hoberman et al.

Sex	Tympanostomy	Medical	*Total*
Female	**42** (46.4)	**48** (43.6)	*90*
Male	**87** (82.6)	**73** (77.4)	*160*
Total	*129*	*121*	*250*

Notice the CS formula.

$$x^2 = \sum \frac{(O_i - E_i)^2}{E_i}$$

x^2 = *chi squared*
O_i = *observed value*
E_i = *expected value*

The expected value (E) is calculated for each cell by multiplying its row and column totals and then dividing by the entire table total (250 in this case). The expected values are in parentheses in the table. One other piece of information is needed to run the CS test: the degrees of freedom (df). The df is a measure of how many of the 4 data values in the table above are actually independent. The df is calculated by multiplying the number of rows minus 1 by the number of columns minus 1. In our example, (2-1) × (2-1) = 1 df. Only 1 value (any one of the 4) is actually independent. This means that if you know the totals in the rows and columns, all you need is one of the 4 data numbers and the other 3 can be calculated. Using an online CS calculator, the CS test equals 1.37 corresponding to a $P > .05$, so any differences are not statistically significant. CS testing was used by Hoberman et al to check for equal distribution to 2 groups in the study after randomization.

—Jordan Hupert, MD

Tympanostomy tubes vs medical management for recurrent otitis

Hoberman A, Preciado D, Paradise JL, Chi DH, Haralam M, Block SL. Tympanostomy Tubes or Medical Management for Recurrent Acute Otitis Media. *N Engl J Med* 2021;384:1789-99. https://doi.org/10.1056/NEJMoa2027278

Question Among young children with frequent acute otitis media (AOM) infections, what is the therapeutic efficacy of tympanostomy-tube (TT) placement, compared with episodic antibiotic treatment of AOM infections, in reducing the subsequent number of AOMs?

Design Age-stratified randomized controlled trial.

Setting Children's Hospital of Pittsburgh and affiliated practices, Children's National Medical Center in Washington, D.C., and Kentucky Pediatric and Adult Research in Bardstown, Kentucky.

Participants Two hundred fifty children 6 to 35 months of age with 3 or more AOM episodes in the last 3 months.

Intervention TT placement vs medical management (episodic antibiotic treatment).

Outcomes Primary outcome: Mean number of AOM episodes per child-year over a 2-year period.

Main Results Intention-to-treat and per-protocol analyses (per-protocol analysis was employed, as 10% of the TT placement and 45% of the medical management groups crossed over) demonstrated AOM rate risk ratios of 0.97 (95% CI, 0.84-1.12) and 0.82 (95% CI, 0.69-0.97), respectively.

Conclusions Depending on the type of analysis (intention to treat or per protocol), TT placement was either non-superior or superior to medical management.

Commentary Hoberman et al, in a trial with stringent inclusion/exclusion criteria, compared TT placement with medical management for the treatment of recurrent AOM. Ten percent of the tubes group and 45% of the medical management group crossed treatment arms. The intention-to-treat analysis revealed no significant difference in the primary outcome, AOM rates over 2 years. However, the per-protocol analysis revealed a statistically significant 15% lesser AOM incidence in the TT group. Time to first AOM, treatment failures, days with otitis-related symptoms, and days of antimicrobial treatment, but not days with tube otorrhea, favored the TT group in the intention-to-treat analysis. Hearing and speech were not assessed. However, a previous study found no difference in speech or language development between children randomized to immediate versus delayed tube insertion.[1] Continuing medical management as an option for children with recurrent AOM should be counter-balanced with the risk of frequent AOM necessitating tube placement

Amber D. Shaffer, PhD
UPMC Children's Hospital of Pittsburgh
Pittsburgh, Pennsylvania

Reference

1. Paradise JL, Feldman HM, Campbell TF, Dollaghan CA, Rockette HE, Pitcairn DL, et al. Tympanostomy tubes and developmental outcomes at 9 to 11 years of age. N Engl J Med 2007;356:248-61. https://doi.org/10.1056/NEJMoa062980

Robust antibody response in children to acute COVID-19 infection and lasts for months

Garrido C, Hurst JH, Lorang CG, Aquino JN, Rodriguez J, Pfeiffer TS, et al. Asymptomatic or mild symptomatic SARS-CoV-2 infection elicits durable neutralizing antibody responses in children and adolescents. *JCI Insight* 2021. https://doi.org/10.1172/jci.insight.150909

Question Among children and adolescents with asymptomatic or mildly symptomatic SARS-CoV-2 infection, how do the specific antibody and neutralizing responses compare with those of adults?

Design Two-group cohort study.

Setting Duke University, Durham North Carolina.

Participants Sixty nine children and adolescents (<21 years of age) with mild or asymptomatic acute SARS-CoV-2 infection, compared with 24 adults.

Intervention Sera collected during acute infection and approximately 2 and 4 months later.

Outcomes SARS-CoV-2–specific humoral immune and neutralizing responses.

Main Results Both acute infection and 2 and 4 months later elicited strong IgM, IgG, and IgA antibody responses against SARS-CoV-2 antigens and were associated with neutralizing activity. These responses were as good or better than those elicited by adult sera.

Conclusions Children and adolescents generate strong and durable antibody responses to SARS-CoV-2 infection.

Commentary This study provided important data on the longitudinal change of SARS-CoV-2 antibody responses in children and adolescents with asymptomatic or mildly symptomatic infection. Similar to adults,[1] SARS-CoV-2 IgM, IgG, IgA and neutralizing antibody responses persisted after 4 months of the acute infection. It would be interesting to follow up for a longer period of time as such evidence would be valuable for pediatric patient management and vaccination strategies. In addition to antibody levels, it is also important to evaluate the quality of antibody responses, such as dynamic changes of antibody binding avidity and T-cell responses in pediatric patients with COVID-19. Furthermore, this report demonstrated that SARS-CoV-2 specific antibody responses in children and adolescents were generally more robust and durable than those of adults with mild symptomatic infection. This finding was consistent with a previous study,[2] which demonstrated higher IgG and surrogate neutralizing antibody activity as well as stronger antibody binding avidity in children compared with young adults. However, the antibody responses vary remarkably from young children, adolescences, and young adults, to elderly adults and adults with metabolic syndrome and chronic cardiovascular disease who exhibit higher IgG activity.[3] The authors' conclusion would be strengthened if the patient cohort could be stratified based on more refined age ranges and comorbidities in both pediatric and adult populations using a larger sample size. It is also recommended to report assay analytical performance (eg, cutoff, linear range, precision, specificity) to demonstrate assay reliability.

Zhen Zhao, PhD
He S. Yang, PhD
Weill Cornell Medicine
New York, New York

References

1. Jiang XL, Wang GL, Zhao XN, Yan FH, Yao L, Kou ZQ, et al. Lasting antibody and T cell responses to SARS-CoV-2 in COVID-19 patients three months after infection. Nat Commun 2021;12(1):897. https://doi.org/10.1038/s41467-021-21155-x

2. Yang HS, Costa V, Racine-Brzostek SE, Acker K, Yee J, Chen Z, et al. Association of age with SARS-CoV-2 antibody responses. JAMA Network Open 2021;4:e214302. https://doi.org/10.1001/jamanetworkopen.2021.4302
3. Racine-Brzostek SE, Yang HS, Jack GA, Chen Z, Chadburn A, Ketas TJ, et al. Postconvalescent SARS-CoV-2 IgG and neutralizing antibodies are elevated in individuals with poor metabolic health. J Clin Endor Metab 2021;106:e2025-34. https://doi.org/10.1210/clinem/dgab004

Get SET Early autism screening program

Pierce K, Gazestani V, Bacon E, Courchesne E, Cheng A, Barnes CC, et al. Get SET Early to Identify and Treatment Refer Autism Spectrum Disorder at 1 Year and Discover Factors That Influence Early Diagnosis. *J Pediatr* 2021;236:179-88. https://doi.org/10.1016/j.jpeds.2021.04.041

Question Among toddlers, how does early autism-spectrum-disorder (ASD) screening affect referral and diagnosis?

Design Cohort study.

Setting Pediatric offices throughout San Diego County.

Participants Two hundred three providers screened toddlers ages 12-24 months.

Intervention Early ASD screening program, Get SET Early, employing the Communication and Symbolic Behavior Scales Infant-Toddler Checklist at ages 12, 18, and 24 months.

Outcomes ASD detection rate and factors that influence the "SET," screen-evaluate-treat sequence.

Main Results Two hundred three pediatricians administered 57 603 screens. Only one-third of screen-positive children were referred for treatment, due to provider concern for false positive results. However, parental concern doubled the referral probability. Eight hundred ninety seven toddlers were evaluated, of whom one-half were diagnosed with ASD, approximately a 1% overall prevalence.

Conclusions Get SET Early implementation detected ASD and initiated early treatment.

Commentary Innovative community-based screen-evaluate-treat solutions are critical to ameliorating delays and disparities in ASD diagnosis and intervention. Pierce et al achieved broad community impact with their Get SET Early model. Strengths of this model include effective training and partnerships with community pediatricians serving demographically diverse children, digital screening with systematic data collection on clinical decision-making and referral tracking, and collaboration with Part C agencies to ascertain treatment engagement. This large, rigorously designed study has important practice implications. Most notably, sequential broadband screening starting at 12 months expedites detection of ASD and is associated with diagnosis and treatment enrollment at younger ages. As these authors and others[1] have highlighted, the combination of caregiver concern, pediatrician judgement, and use of screening results requires further evaluation to determine how ASD detection can be optimized. This line of study is especially important given limitations in the accuracy of ASD screening tools.[2] A potential limitation regarding generalizability of the Get SET Early model is its reliance on a laboratory-grade, federally-funded evaluation center. Given widespread problems with access to ASD diagnostic specialists,[3] the scalability and sustainability of Get SET Early may be limited unless it can be successfully embedded within community systems of ASD evaluation.

Rebecca McNally Keehn, PhD
Mary Ciccarelli, MD
Indiana University School of Medicine
Indianapolis, Indiana

References

1. McNally Keehn R, Tang Q, Swigonski N, Ciccarelli M. Associations Among Referral Concerns, Screening Results, and Diagnostic Outcomes of Young Children Assessed in a Statewide Early Autism Evaluation Network. J Pediatr 2021;233:74-81.e8. https://doi.org/10.1016/j.jpeds.2021.02.063
2. Guthrie W, Wallis K, Bennett A, Brooks E, Dudley J, Gerdes M, et al. Accuracy of Autism Screening in a Large Pediatric Network. Pediatrics 2019;144:e20183963. https://doi.org/10.1542/peds.2018-3963
3. Bridgemohan C, Bauer NS, Nielsen BA, DeBattista A, Ruch-Ross HS, Paul LB, et al. A workforce survey on developmental-behavioral pediatrics. Pediatrics 2018;141:e20172164. https://doi.org/10.1542/peds.2017-2164

Updated Bhutani bilirubin curves include 12-hour-age levels and demonstrate gestational age and racial differences

Bahr TM, Henry E, Christensen RD, Minton SD, Bhutani VK. A New Hour-Specific Serum Bilirubin Nomogram for Neonates ≥35 Weeks of Gestation. *J Pediatr* 2021;236:28-33.e1. https://doi.org/10.1016/j.jpeds.2021.05.039

Question Among 35+ week gestational age babies, how do gestational age and race affect hour-specific total serum bilirubin (TSB) levels?

Design Retrospective data-set analysis.

Setting Twenty Intermountain Healthcare hospitals.

Participants All newborns ≥35 weeks gestational age with at least 1 TSB.

Intervention First serum bilirubin level.

Outcomes Hour-specific TSB nomogram.

Main Results The new nomogram included 2 orders of magnitude more patient data and generally agreed with the 1999 nomogram. Differences included higher 75th and 90th percentile TSB values after 60 hours, higher values for lower gestational age, lower values for Black babies, and higher values for Asian babies.

Conclusions An updated version of the original 1999 Bhutani bilirubin curves include 12-hour-age levels and demonstrates differences based on gestational age and race.

Commentary Updating the 1999 Bhutani nomogram using a cohort of nearly 400 000 bilirubin results seems like a great idea; however, does it improve management of hyperbilirubinemia? The authors fail to provide any information on the predictive accuracy of their risk zones or if they represent an improvement over the prior nomogram. Although they offer data on bilirubin results obtained in the first 12 hours after birth, it is unclear how useful these data are. These 12-hour-after-birth data are likely based on an unrepresentative sample of infants who had additional risk factors or early jaundice, prompting the early testing. The authors confirm the shortcomings of using a single nomogram for all infants, as nomograms differed based upon race and gestational age. Use of a single nomogram for treatment decisions is particularly hard to justify, because infants born prematurely and those with hemolysis are at higher risk of bilirubin toxicity, and therefore should be treated at lower levels. Over 2 decades have passed since the development of the Bhutani nomogram, and we now have electronic health records and smartphones in our pockets. It is time to move beyond simple 4 category prediction models on a piece of paper and use more robust risk prediction models that incorporate multiple factors to improve our risk prediction and clinical decision support.[1-3]

Michael W. Kuzniewicz, MD, MPH
Thomas B. Newman, MD, MPH
University of California, San Francisco
San Francisco, California

References

1. Chou JH. Predictive Models for Neonatal Follow-Up Serum Bilirubin: Model Development and Validation. JMIR Med Inform 2020;8:e21222. https://doi.org/10.2196/21222
2. Daunhawer I, Kasser S, Koch G, Sieber L, Cakal H, Tutsch J, et al. Enhanced early prediction of clinically relevant neonatal hyperbilirubinemia with machine learning. Pediatr Res 2019;86:122-7. https://doi.org/10.1038/s41390-019-0384-x
3. Kuzniewicz MW, Park J, Niki H, Walsh EM, McCulloch CE, Newman TB. Predicting the Need for Phototherapy After Discharge. Pediatrics 2021;147:e2020019778. https://doi.org/10.1542/peds.2020-019778

New tubeless, automated insulin pump improves HgA1c and time in pre-defined glucose range

Brown SA, Forlenza GP, Bode BW, Pinsker JE, Levy CJ, Criego AB, et al. Omnipod 5 Research Group. Multicenter Trial of a Tubeless, On-Body Automated Insulin Delivery System with Customizable Glycemic Targets in Pediatric and Adult Participants with Type 1 Diabetes. *Diabetes Care* 2021;44:1630–40. https://doi.org/10.2337/dc21-0172

Question Among children and adults with type 1 diabetes, what are the safety and effectiveness benefits of a tubeless, automated insulin delivery system?

Design Prospective, single cohort, multicenter.

Setting Seventeen sites across the U.S.

Participants Six to 70 years old, diagnosed with type 1 diabetes for at least 6 months. Two cohorts: 111 children (ages 6-13.9 years) and 129 adolescents/adults (ages 14-70 years).

Intervention A tubeless insulin pump (Pod) with an embedded automated insulin delivery algorithm (Omnipod 5), an interoperable glucose sensor, and a mobile app (Omnipod 5 app) on a locked-down Android phone. Two weeks of usual insulin regimen preceded 3 months of automated insulin delivery.

Outcomes Primary safety outcomes: incidence of severe hypoglycemia and diabetic ketoacidosis. Primary effectiveness outcomes: change in HbA1c and percent time in sensor glucose range 70-180 mg/dL (time in range).

Main Results HbA1c was significantly reduced in children by 0.68% (mean ± SD: 7.67 ± 0.95% to 6.99 ± 0.63%, $P < .0001$) and in adults by 0.38% (7.16 ± 0.86% to 6.78 ± 0.68%), $P < .0001$). Time in range improved by 3.7 h/day in children and 2.2 h/day in adults (both $P < .0001$). Time in hypoglycemia <70 mg/dL among adults (median [interquartile range] decreased from 2.00% [0.63-4.06] to 1.09% [0.46-1.75], $P < .0001$). There was no change in children.

Conclusions The Omnipod 5 appears safe and demonstrated modest improvement in HgB A1c and time in a pre-defined glucose range. Adults experienced less time in hypoglycemia.

Commentary The trial was designed for FDA approval (now granted) and is soon to be commercially launched as Insulet's Omnipod HORIZONTM automated insulin dosing (AID). This will be the fourth commercial system to market, following Medtronic's Smartguard, Tandem's Control IQ, and CamAPSR FX, but is different by virtue of the tubeless pump. The single-arm design limits effectiveness analyses. Best glucose time in range was overnight, and greatest improvements were in participants with poorer baseline control. The prima facie evidence is that this system is safe, with similar effectiveness to other AID systems.[1] The short study duration limits ability to detect rare events such as diabetic ketoacidosis or severe hypoglycemia. Generalizability is unknown, the study participants had excellent baseline glucose control, and there was lack of representation from ethnic minorities. This aside, the commercial release will provide another option for a personalized AID-system choice.

Mercedes Burnside, MB, ChB
Martin de Bock, PhD
University of Otago
Christchurch, New Zealand

Reference

1. Weisman A, Bai J-W, Cardinez M, Kramer CK, Perkins BA. Effect of artificial pancreas systems on glycaemic control in patients with type 1 diabetes: a systematic review and meta-analysis of outpatient randomised controlled trials. Lancet Diabetes Endocrinol 2017;5:501-12. https://doi.org/10.1016/s2213-8587(17)30167-5