neurodevelopmental outcomes, our results highlight minor microbial shifts that align with findings in older children with ASD, which may indicate early gut microbiome dysbiosis, but further research is needed.

119

Novel approach to meta-analysis of randomized trials with time-to-event endpoints: Application to Mrcc

Pratik Reddy and Ludovic Trinquart
Tufts Clinical and Translational Science Institute

OBJECTIVES/GOALS: Our study's primary objective is to develop a multivariate random effects meta-analysis model that estimates risk ratio at a series of clinically relevant time points. Our goal in doing so is to address gaps in traditional hazard ratio (HR) meta-analyses due to observed time-varying treatment effects. METHODS/STUDY POPULATION: Our model estimates pooled risk ratios (RRs), the ratio of cumulative event probabilities, at predefined timepoints (e.g., 12, 24, 36 months). We derive the variance of the log-RR for each trial at each time point by applying the delta method to the Kaplan-Meier (KM) estimator function. Furthermore, we will derive the covariance to account for the correlation between RR estimates at different time points within the same trial. We would then like to incorporate these derivations of within-trial variance and covariance at each available time milestone into our multivariate model to increase the precision of our pooled RR estimated for both Overall Survival (OS) and Progression-Free Survival (PFS). RESULTS/ANTICIPATED RESULTS: We identified 7 trials by consulting 5 systematic reviews published through 2021-2025 to ensure use of updated data. We reconstructed the pseudo-Individual Patient Data of these trials by applying PDFs of published KM curves to a validated survival curve reconstruction algorithm. For the outcome of overall survival, the max amount of follow-up time recorded is 108 months. On the other hand, there are 96 months of available follow-up for the outcome of progression-free survival. In a univariate random effects meta-analysis of HR for OS, the pooled HR was found to be 0.81 (CI: 0.74 – 0.89). Additionally, the pooled HR for PFS was found to be 0.68 (CI: 0.59 – 0.79). For both outcomes, we anticipate the multivariate random effects analysis of Risk Ratio to yield more precise confidence intervals than those of HR. DISCUSSION/SIGNIFICANCE OF IMPACT: The proposed multivariate meta-analysis model will improve precision of treatment effect estimates which will add to the ongoing evaluation of immunotherapies as a current standard of care. This is significant to translational science as it provides time-specific treatment effect estimates that could be more interpretable in a clinical setting.

120

Correlation between deepest vertical pocket (DVP) amniotic fluid measurements and neonatal Apgar scores at term

Noor Banihashem Ahmad[1,2], Hannah Heselton[2], William M. Curtin[2], Ravi Chokshi[2], Nayera Guirguis[2], Dung Dang[2], Allen Kunselman[3], Serena Wu[2] and Serdar H. Ural[2]
[1]Penn State College of Medicine; [2]Division of Maternal-Fetal Medicine, Department of Obstetrics and Gynecology, Penn State Milton S. Hershey Medical Center and [3]Division of Biostatistics and Bioinformatics, Department of Public Health Sciences, Penn State Milton S. Hershey Medical Center

OBJECTIVES/GOALS: Although the Deepest Vertical Pocket (DVP) method is increasingly used to assess amniotic fluid volume due to its simplicity and lower intervention rates, its relationship with neonatal outcomes remains underexplored. This study aims to determine if DVP measurements are associated with Apgar scores at 1 and 5 minutes in low-risk term pregnancies. METHODS/STUDY POPULATION: An IRB-approved retrospective cohort study of 171 singleton term pregnancies delivered at Penn State Hershey Medical Center between 2019 and 2024 was performed. Inclusion criteria included delivery at ≥37 weeks, birth weight between the 10th and 90th percentiles, absence of fetal anomalies, spontaneous labor or scheduled Cesarean, and classification as low risk based on ACOG 2019 Levels of Maternal Care guidelines. A single second-trimester DVP measurement was obtained from anatomy ultrasounds. Apgar scores at 1 and 5 minutes were abstracted from delivery records. Statistical analysis included ANOVA and Spearman correlation to assess relationships between DVP and Apgar scores. RESULTS/ANTICIPATED RESULTS: Among 171 patients, the mean DVP was 4.79 cm (SD 0.94). Apgar scores at 1 minute were most commonly 8 (73%) or 9 (15%), and 94% of neonates had a score of 9 at 5 minutes. There was no statistically significant correlation between DVP and Apgar scores at 1 minute (Spearman $r = -0.11$, $p = 0.14$) or 5 minutes ($r = 0.10$, $p = 0.19$). ANOVA results also revealed no significant differences in DVP across Apgar score categories at 1 minute ($p = 0.12$) or 5 minutes ($p = 0.41$). DISCUSSION/SIGNIFICANCE OF IMPACT: In this cohort of low-risk term pregnancies, DVP was not significantly associated with Apgar scores at 1 or 5 minutes. This study is one of the few to isolate DVP rather than AFI and to apply it to routine obstetric care. These findings support the use of DVP as a non-invasive, reliable fluid assessment method without concern for adverse neonatal outcomes.

121

Study design considerations for post-deployment monitoring of AI/ML models

Momin Malik, Dave Watson, Madison J. Beenken, Shauna M. Overgaard, Hanyin Wang, Imad Absah, Chung Il Wi and Young J. Juhn
Mayo Clinic

OBJECTIVES/GOALS: Safely and effectively translating AI classification models requires robust post-deployment monitoring. Yet there is little guidance about how to do so. Here, we outline how standard machine learning model development study designs are insufficient for post-deployment monitoring, and what specific study designs are needed. METHODS/STUDY POPULATION: We made original linkages between machine learning methodology and biostatistics, particularly to diagnostic testing, for study design guidance on post-deployment model performance and fairness assessment. RESULTS/ANTICIPATED RESULTS: The kind of case–control