The Impact of Environmental and Genetic Factors on Neonatal Late-Onset Sepsis

Matthew J. Bizzarro, MD, Yuan Jiang, PhD, Naveed Hussain, MD, Jeffrey R. Gruen, MD, Vineet Bhandari, MD, DM, and Heping Zhang, PhD

Objective To assess the genetic contribution to late-onset sepsis in twins in the newborn intensive care unit.
Study design A retrospective cohort analysis of twins born from 1994 to 2009 was performed on data collected from the newborn intensive care units at Yale University and the University of Connecticut. Sepsis concordance rates were compared between monozygotic and dizygotic twins. Mixed-effects logistic regression analysis was performed to determine the impact of selected nongenetic factors on late-onset sepsis. The influence of additive genetic and common and residual environmental effects were analyzed and quantified.
Results One hundred seventy monozygotic and 665 dizygotic twin pairs were analyzed, and sepsis identified in 8.9%. Mean gestational age and birth weight of the cohort was 31.1 weeks and 1637 grams, respectively. Mixed-effects logistic regression determined birth weight (regression coefficient, -0.001; 95% CI, -0.003 to 0.000; P = .028), respiratory distress syndrome (regression coefficient, 1.769; 95% CI, 0.943 to 2.596; $P < .001$), and duration of total parenteral nutrition (regression coefficient, 0.041; 95% CI, 0.017 to 0.064; $P < .001$) as significant nongenetic factors. Further analysis determined 49.0% (P = .002) of the variance in liability to late-onset sepsis was due to genetic factors alone, and 51.0% (P = .001) the result of residual environmental factors.
Conclusions Our data support significant genetic susceptibility to late-onset sepsis in the newborn intensive care unit population. *(J Pediatr 2011;158:234-8)*.

Bloodstream infections (BSI) are a relatively common problem in the newborn intensive care unit (NICU) population, particularly in premature neonates.[1-4] Late-onset sepsis (BSI occurring at >72 hours of life) comprises the majority of episodes in this population, with a high rate of associated morbidity and mortality, longer hospital stay, and increased costs.[1-10]

Traditionally, neonatologists attribute the high prevalence of late-onset sepsis in the NICU population to a combination of environmental and host factors including but not limited to the immature neonatal immune system, a compromised skin barrier, the need for invasive procedures, the prolonged use of invasive life-support apparatus such as endotracheal tubes and central venous catheters, and prolonged hospital stay.[9,11] There is significant individual variability among the NICU population with respect to the susceptibility, response to, and outcome associated with late-onset sepsis that may not be explained by these factors alone.[2,11] We hypothesized that in addition to environmental effects, genetic factors play a major role in predisposing neonates toward developing late-onset sepsis. Using data from a large cohort of monozygotic (MZ) and dizygotic (DZ) twins, we analyzed and quantified the genetic and environmental contributions to late-onset sepsis.

Methods

Data on all twin pairs born from January 1, 1994, to December 31, 2009, were collected from two medical centers: the University of Connecticut and Yale University. The institutional review boards of each participating center approved the contribution of data to this study.

BSI	Bloodstream infections	NICU	Newborn intensive care unit
BW	Birth weight	RDS	Respiratory distress syndrome
DZ	Dizygotic	SNP	Single nucleotide polymorphisms
GA	Gestational age	TPN	Duration of total parenteral nutrition
TNF	Tumor necrosis factor		
INST	Treating institutions	VENT	Duration of (invasive) mechanical ventilation
MELR	Mixed effects logistic regression		
MZ	Monozygotic	VLBW	Very low birth weight

From the Department of Pediatrics (M.B., J.G., V.B.), Yale University School of Medicine, New Haven, CT; the Department of Epidemiology and Public Health (Y.J., H.Z.), Yale University School of Medicine, New Haven, CT; the Department of Pediatrics (N.H.), University of Connecticut School of Medicine, Farmington, CT; and the Department of Genetics, Investigative Medicine and the Yale Child Health Research Center, New Haven, CT (J.G.)

Supported by the National Institutes on Drug Abuse (R01 DA016750 to Y.J and H.Z.), the National Institute of Neurological Disorders and Stroke (R01 NS43530 to J.G.), and the National Heart, Lung, and Blood Institute (K08 HL 074195 to V.B.). The authors declare no conflicts of interest.

 10.1016/j.jpeds.2010.07.060

The zygosity of each twin pair was determined by histopathologic examination of the placenta with additional confirmation using sex concordance or discordance. Late-onset sepsis was specifically chosen as the outcome of interest because, in our (and in most) NICUs, it describes the vast majority of BSI.[1-4] Late-onset sepsis was defined as a blood culture obtained at >72 hours of life that yielded a traditional neonatal pathogen (eg, *Escherichia coli*) or a commensal species (eg, *Staphylococcus epidermidis*).[4] Blood cultures that yielded common skin flora such as coagulase-negative staphylococci, which comprised the majority of commensal species-related BSI, were reviewed using specific inclusion criteria from the Centers for Disease Control and Prevention.[12] Although this definition has recently been modified,[13] the previous definition[12] was used to maintain consistency throughout the study period. Because late-onset sepsis was analyzed as a dichotomous outcome, if an infant had multiple episodes only the first was included.

Respiratory distress syndrome (RDS) was defined as the presence of respiratory distress with an oxygen requirement in the first 6 hours of life, accompanied by a characteristic chest radiograph. Duration of mechanical ventilation (VENT) was defined as the total number of days that an infant required invasive (ie, via an endotracheal tube) positive pressure ventilation while in the NICU. Positive pressure ventilation included high frequency ventilation and/or intermittent mandatory ventilation. Duration of total parental nutrition (TPN) was defined as the total number of days that the infant required intravenous nutrition while in the NICU. Treating institutions (INST) were the two medical centers where the data were collected: the University of Connecticut and Yale University.

Statistical Analyses

Demographic data were analyzed using the Student *t* test, Wilcoxon rank sum test, or χ^2 analysis where appropriate.

χ^2 analyses of the zygosity data were performed to compare sepsis concordance rates between MZ and DZ twins. The observed numbers of twin pairs with both infants affected, with only one infant affected, and with neither infant affected were determined for MZ and DZ groups. These observed numbers formed a 2×3 contingency table and the analog expected numbers of twin pairs were calculated from the corresponding marginal totals. The observed to expected distributions of concordance were compared using χ^2 analysis.

Mixed effect logistic regression (MELR) analysis was next performed to identify the impact of selected putative risk factors on sepsis. The covariates utilized in the model included birth sequence, male sex, gestational age (GA), birth weight (BW), RDS, VENT, TPN, and INST. The status of the outcomes from twin pairs was treated as a correlated event. A MELR model was fitted to assess the relationship between the covariates listed and the outcome of interest (sepsis) and to incorporate the correlation between twin pairs.

The A (additive genetic) C (common environment) E (unique environment) model described by Feng et al[14] was then used to estimate the variance in liability for sepsis. This mixed-effects probit model included covariate effects, an additive genetic effect, a common environmental effect shared by a twin pair (no matter which zygosity it has), and a residual environmental effect. Unique "environmental" effects specific to the NICU population included nongenetic risk factors for late-onset sepsis, such as birth weight, invasive mechanical ventilation, and TPN use. The additive genetic effect, the common environmental effect, and the residual environmental effect were assumed to be independent and normally distributed. Because MZ twins are genetically identical, their additive genetic effects are equal. For DZ twins, the covariance of the additive genetic effects is half that for MZ twins.[15] The covariates adjusted in the ACE model included all significant covariates utilized in the MELR analysis. Genetic heritability could then be estimated using the ratio of estimated genetic variance and the total variance of the trait. To confirm the results of the ACE model, we fitted the AE model (without the shared environmental factors) to compare the genetic effects and the residual environmental effects.

An empirical power calculation of the genetic effect analysis was performed to assess the reliability of the model and results. We randomly simulated 100 data sets with the same sample size as our collected data, using the estimates for covariate effects, genetic effects, and environmental effects obtained from the above AE model. Then, we used the AE model again to fit each simulated data set and recorded whether the genetic effect can be significantly identified. An empirical power is the percentage of the times of significant identifications out of these 100 inferences.

Statistical analyses were performed using SAS 9.1 (PROC NLMIXED; SAS Inc., Cary, North Carolina). A *P* value of less than .05 was considered statistically significant.

Results

Late-onset sepsis was diagnosed in 149 of 1670 (8.9%) infants from our cohort, which represented 59 of 800 (7.4%) infants from the University of Connecticut and 90 of 870 (10.3%) from Yale University. The incidence of sepsis was determined to be inversely proportional to BW. In the infants with BW <1000 grams, 84 of 296 (28.4%) were diagnosed with late-onset sepsis, as compared with 47 of 362 (13.0%) in those with BW 1000 to 1499 grams, and 12 of 540 (2.2%) in infants with BW 1500 to 1999 grams. We also noted that only 6 of 469 (1.3%) infants were diagnosed in the subpopulation with BW ≥2000 grams.

There were 139 (93.3%) episodes of monomicrobial and 10 (6.7%) episodes of polymicrobial late-onset sepsis identified in our cohort. Coagulase-negative staphylococci were the most common organisms isolated (63 of 149 episodes; 42.2%), followed by *S aureus* (18 of 149 cases; 12.1%), *Enterococcus* species (13 of 149; 8.7%), and *Klebsiella pneumoniae* (8 of 149; 5.4%).

Zygosity data comprising 170 MZ and 665 DZ twin pairs were used for analysis. The 835 twin pairs had a mean GA