

Supplementary Online Content

Josephson CD, Caliendo AM, Easley KA, et al. Blood transfusion and breast milk transmission of cytomegalovirus in very low-birth-weight infants: a prospective cohort study. *JAMA Pediatr*. Published online September 22, 2014.
doi:10.1001/jamapediatrics.2014.1360.

eMethods. Additional methods for laboratory testing and statistical analysis

eFigure 1. Percentage of longitudinal blood and urine samples with CMV NAT expression

eFigure 2. ROC curve for maximum \log_{10} CMV NAT expression in breast milk as a predictor of mother-to-infant CMV transmission among 189 CMV seropositive mothers who fed their infants CMV-positive maternal breast milk

eResults. Results for CMV-infected infant laboratory abnormality, statistical model adequacy

eTable. Factors associated with mother-to-infant CMV transmission by breast milk among CMV seropositive mothers

This supplementary material has been provided by the authors to give readers additional information about their work.

eMethods

Laboratory

Whole blood samples were tested allowing the detection of both cell free and cell associated virus, which will maximize the sensitivity of the assay. The limit of detection of the assay is 205 IU/ml and the limit of quantification is 300 IU/ml. The limit of detection is defined as the lowest concentration of DNA detected in 95% of replicates, while the limit of quantification is determined based on assay reproducibility.

Competing risks and censoring

The two competing endpoints were (1) onset of CMV infection with right censoring at last negative blood or urine CMV NAT test and (2) death with right censoring at discharge or 90 days. All censoring is assumed non-informative.

Adjustment for multiple births

The CMV infection data occurs in clusters due to multiple births. CMV incidence estimation is valid and consistent not considering the correlation due to clusters; however, variance estimation used to construct 95% confidence intervals does account for this correlation. Bootstrapping by clusters was used to obtain valid variance estimation and to calculate 95% confidence intervals.

The cause-specific hazard rate (CSHR) refers to the instantaneous rate of occurrence of CMV among infants still CMV-free and the subdistribution hazard rate (SHR) is the probability of CMV by time, t .¹

Cause-specific hazard ratios (CSHR) were calculated to measure the degree of association between baseline characteristics and CMV infection and between baseline characteristics and death by fitting a stratified Cox proportional-hazards regression model for competing risks.

Covariates for the univariable competing risk Cox regression model (Table 3a)

Eleven risk factors were included as potentially independent prognostic risk factors in the univariable competing risk analyses of CMV and mortality.

Eight baseline covariates

Infant baseline risk factors included birth weight, SNAP, gender and leukopenia at birth (WBC < 5000 cells/ μ L). Maternal baseline covariates included isolated spontaneous labor, chorioamnionitis (clinical and/or histological diagnosis), antenatal steroids and rupture of membranes. Rupture of membranes was defined as premature rupture of membranes (yes or no) and rupture of membranes greater than 18 hours or \leq 18 hours.

Three time-dependent covariates (TDC)

To examine the temporal relation between late onset sepsis (>7 days) and outcome, occurrence of late onset sepsis was included as a binary (0,1) time-dependent covariate in the Cox regression model. Breast milk feeding days and \log_{10} CMV NAT expression in the breast milk were defined as time-dependent covariates that were defined for four study time intervals (weeks 1, 2-3, 4-5 and after 5 weeks). Last value carried forward was used when a TDC was missing for a study time interval.

Covariate selection for the multivariable Cox regression analysis

Bootstrap bagging was used to identify stable and reliable predictors of infant CMV infection.² Nine covariates were included in the analysis [gender, birth weight, leukopenia at birth, isolated spontaneous labor, chorioamnionitis, premature rupture of membranes (yes or no), late onset sepsis, breast milk feeding days, and \log_{10} CMV NAT expression in breast milk].

A dataset was constructed of size equal to the original by random sampling of cases with replacement (bootstrap sampling). On average, approximately one-third of infants were not sampled, whereas some infants were sampled more than once. The bootstrap sample was analyzed using the Cox model with an automated forward stepwise algorithm with entry criterion of $p < 0.10$ and a retention criterion of $p < 0.05$. The result was stored. This process of sampling, automated analysis and storing was repeated 1000 times. The number of times a risk factor appeared in these 1000 analyses was taken as reflection of the reliability (signal). Following Breiman's median rule (devised to balance type I and type II errors), risk factors were

retained if they appeared in at least 50% of the models, with the interpretation that there is a least a 50% chance the risk factor is statistically significant ($p < 0.05$).^{2,3} The cause-specific hazard ratio and its 95% confidence interval were calculated for each factor in the presence of the others in the final model identified with bootstrap bagging (505 infants without missing covariate data; listwise deletion).

Bootstrap bagging was also used to identify the predictors of mortality. Ten covariates were included in the analysis [birth weight, SNAP, leukopenia at birth, isolated spontaneous labor, chorioamnionitis, premature rupture of membranes (yes or no), receipt of antenatal steroids, late onset sepsis, breast milk feeding days, and \log_{10} CMV NAT expression in breast milk].

Assessment of model adequacy for the Cox model

Time dependent covariates (interaction between each baseline risk factor and log time) were included in the Cox regression model to test whether the hazard rates for each risk factor were proportional over time. Significant interactions indicate violation of the assumption. Although the Cox proportional hazards model is the most efficient or powerful when the hazards are proportional the model is still valid as long as the cumulative incidence curves do not cross. Regression diagnostics for the Cox model included Cox-Snell residuals to examine the overall fit of the final model and martingale residuals to compare the observed number of events to the expected number of events based on the fitted model.

Methods for Figure 2c

Repeated-measures analysis of \log_{10} breast milk NAT expression was done using a means model with SAS Proc Mixed (version 9), providing separate estimates of the means by time on study (weeks 1, 2-3, 4-5 and after 5 weeks) for CMV-transmitting mothers and CMV non-transmitting mothers. The model included CMV-transmitting status, time on study and the statistical interaction between the two predictors. A compound-symmetric variance-covariance form among the repeated measurements was assumed for the outcome and robust estimates of the standard errors of parameters were used to perform statistical tests and construct 95% confidence intervals.⁴ The model-based means are unbiased with unbalanced and missing data, so long as the missing data are non-informative (missing at random). All statistical tests were 2-sided and a P value ≤ 0.05 was considered statistically significant.

Methods for eFigure 1

The generalized estimating equations approach was used to analyze the repeated binary data on CMV NAT expression (detected or not detected) in the blood or urine of VLBW infants. The percentage of samples with detectable virus over time was estimated using SAS Proc Genmod with an exchangeable correlation binomial-logit model. Results are reported as the model-based estimates and 95% confidence intervals of the percentage of samples with detectable virus at each of 5 time intervals (1-3 weeks, 4-6 weeks, 7-9 weeks, 10-12 weeks and 13-15 weeks).

Methods for eTable 1 and eFigure 2

Univariable and multivariable logistic regression was used to examine the independent effects of premature rupture of membranes, caesarean delivery and maximum \log_{10} CMV NAT expression in breast milk on maternal-infant CMV transmission among seropositive mothers who fed their infants CMV-positive maternal breast milk. The odds ratio and its 95% confidence interval were calculated for each risk factor in the presence of the others in the final multivariable model. The goodness-of-fit of the logistic model was evaluated using the Hosmer-Lemeshow test. These same models were refit to examine the independent effects of rupture of membranes greater than 18 hours or ≤ 18 hours, caesarean delivery and maximum \log_{10} CMV NAT expression in breast milk on maternal-infant CMV transmission among seropositive mothers who fed their infants CMV-positive maternal breast milk.

The predictive power or predictive ability of the logistic regression model with \log_{10} CMV NAT expression in breast milk as a risk factor of maternal-infant CMV transmission (outcome) among 189 CMV seropositive mothers who fed their infants CMV-positive maternal breast milk was summarized by constructing the receiver operator characteristic curve (ROC) curve.

References

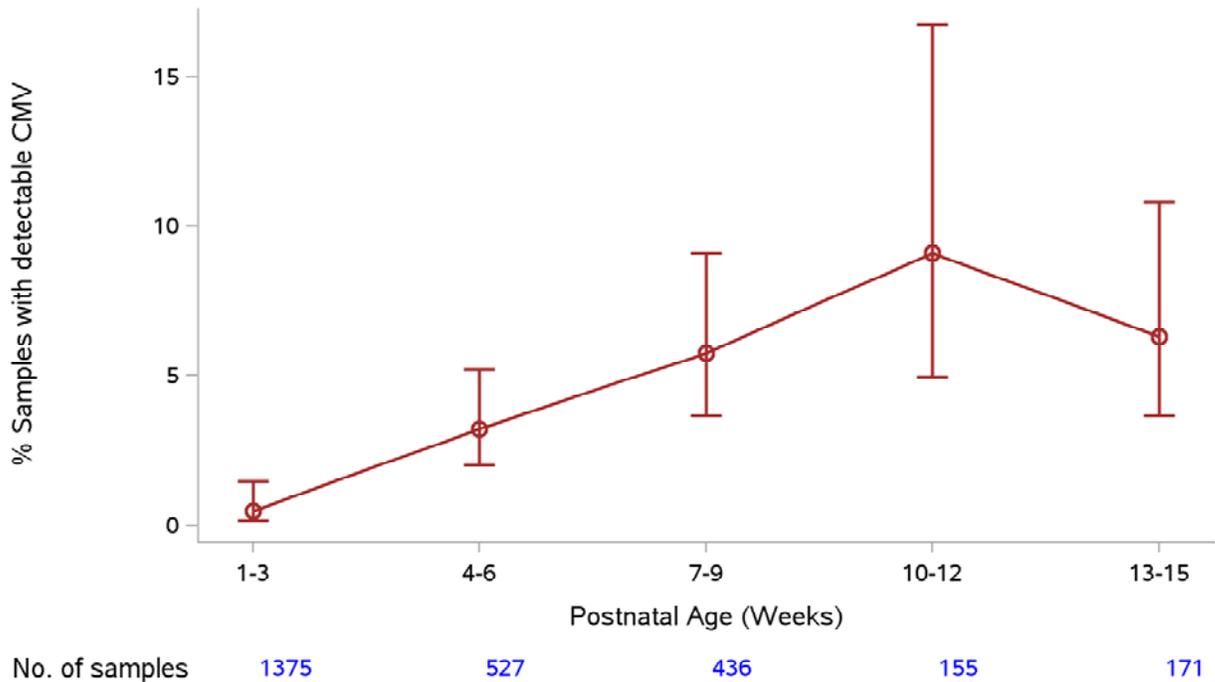
¹Latouche A, Allignol, Beyersmann J, Labopin M and Fine JP. A competing risks analysis should report results on all cause-specific hazards and cumulative incidence functions. *Journal of Clinical Epidemiology* 2013;66:648-653.

²Breiman L. Bagging predictors. *Machine Learning*. 1996;24:123-140.

³Blackstone EH. Breaking down barriers: helpful breakthrough statistical methods you need to understand better. *J Thorac Cardiovasc Surg*. 2001;122:430-439.

⁴Diggle PJ, Liang KY, and Zeger SL. *Analysis of Longitudinal Data*. New York: Oxford Science Publications; 1994. p 68-77.

eFigure 1. Percentage of longitudinal blood and urine samples with CMV NAT expression

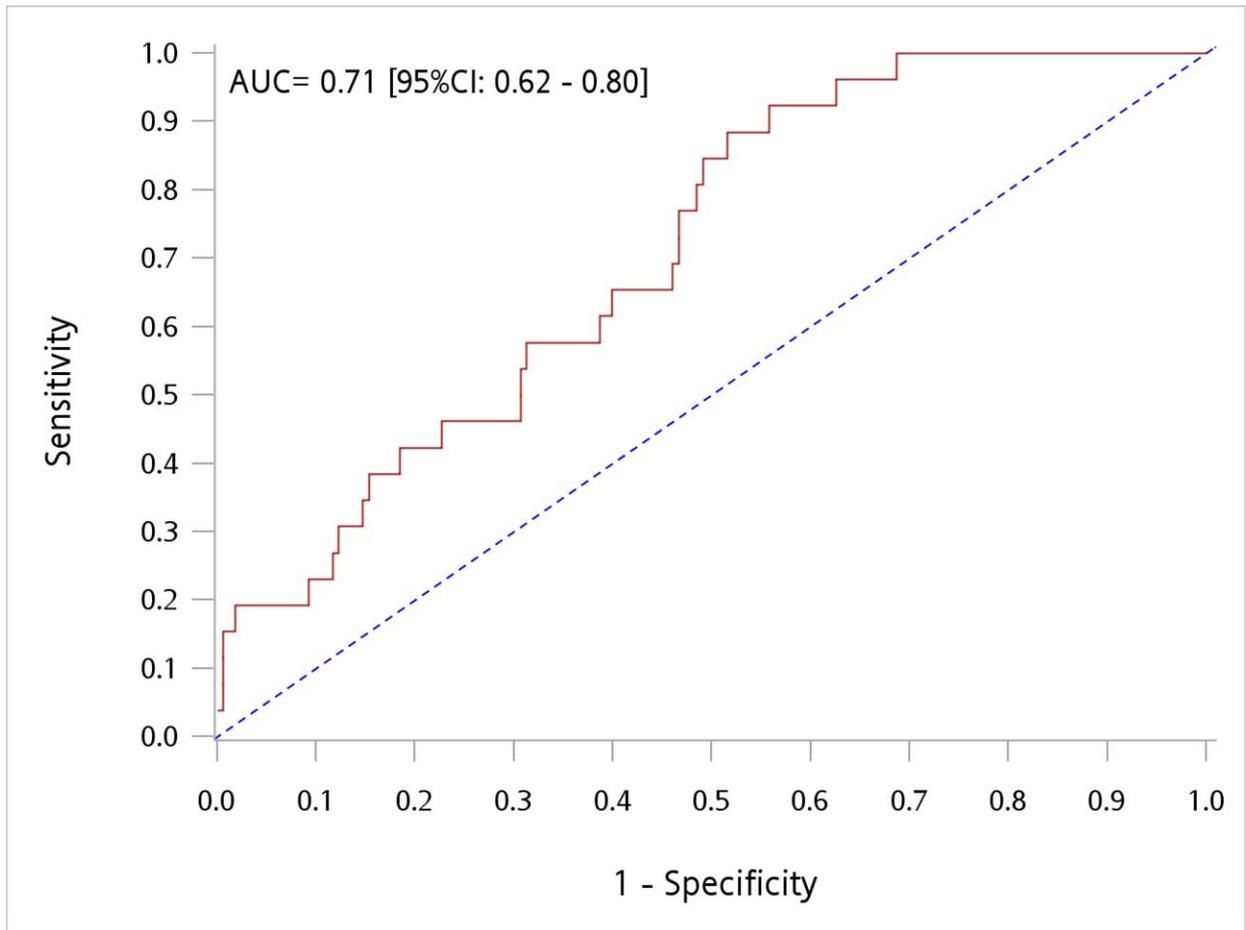


Abbreviations: VLBW, very low birth weight; CMV, cytomegalovirus; NAT, nucleic acid testing; No, number.

The graph provides model-based estimates and 95% confidence intervals of the percentage of samples with detectable virus at each time interval. The analysis included data from 2664 samples (1793 blood samples and 871 urine samples) from 539 VLBW infants (29 infants with CMV infection). The vertical bars are the 95% confidence intervals.

Postnatal age (weeks)	No. of samples	No. of infants with CMV Infection	No. of CMV positive pests	Model-based percentage	95% confidence interval
1-3	1375	5	9	0.5	0.2 - 1.4
4-6	527	10	11	3.2	2.0 - 5.2
7-9	436	20	29	5.8	3.7 - 9.1
10-12	155	8	17	9.1	4.9 - 16.8
13-15	171	8	11	6.3	3.7 - 10.8
Total	2664		77		

eFigure 2. ROC curve for maximum \log_{10} CMV NAT expression in breast milk as a predictor of mother-to-infant CMV transmission among 189 CMV seropositive mothers who fed their infants CMV-positive maternal breast milk



Analysis based on data from 189/462 (40.9%) mothers who fed their infants CMV-positive maternal breast milk (26 transmitter mothers and 163 non-transmitter mothers).

Odds ratio per 1 \log_{10} increase in maximum breast milk CMV NAT expression: 2.75 (95% CI: 1.57-4.84), $p=0.0004$. The calculated risk of mother-to-infant transmission is as follows: $e^w/(1+e^w)$, where $w = -5.67 + 1.012 \cdot X$ and X is \log_{10} maximum breast milk CMV NAT expression.

Estimated intercept \pm standard error and $\beta \pm$ standard error from univariable logistic regression: -5.67 ± 1.16 and 1.012 ± 0.287 .

eResults

Laboratory values for 24 infants determined to have asymptomatic CMV infection

Hepatic enzymes were normal (AST range 17-41; ALT range 5-13) as was hepatic function (direct bilirubin range 0.2-0.5) and no significant hematologic abnormalities attributable to viral infection were observed (platelet range 194-642 x 10³; WBC range 4.5-27.8 x 10³; ANC range 2,210-13,632).

Laboratory values for 5/29 infants with CMV infection and laboratory abnormalities

1 infant had elevated AST (808 units/L), ALT (1140 units/L) and direct bilirubin (2.7 mg/dL); 1 infant had elevated AST (74 units/L), an abnormal platelet count (49 x 10³ cells/ μ L) and low ANC (1000 cells/ μ L) and developed ; Both of these infants received antiviral therapy for CMV infection. In addition, 1 infant had a mildly elevated direct bilirubin (1.1 mg/dL) and two infants had abnormal platelet counts (87 and 66 x 10³ cells/ μ L).

eFigure 1

The percentage of longitudinal blood and urine samples with detectable CMV virus (above 300 IU/ml) increased over the first three months of life ($p < 0.001$, test for linear trend). The percentage of samples with detectable virus increased from 0.5% at 1-3 weeks to 3.2% at 4-6 weeks. By 10-12 weeks, 9.1% (95% CI: 4.9% to 16.8%) of samples had detectable virus.

eTable 1

Both prolonged rupture of membranes (greater than 18 hours versus \leq 18 hours) and premature rupture of membranes (yes or no) were identified as risk factors for maternal-infant CMV transmission among 189 CMV seropositive mothers in separate multiple logistic regression models where breast milk CMV NAT expression and Caesarean delivery were also included as covariates. The adjusted OR for prolonged rupture of membranes was 3.09 (95% CI: 1.18 - 8.08; $p=0.02$) and the adjusted OR for premature rupture of membranes was 4.44 (95% CI: 1.74-11.36; $p=0.002$).

eFigure 2

The ROC curve summarizes the diagnostic accuracy of maximum \log_{10} CMV NAT expression in breast milk as a predictor of maternal-infant CMV transmission among 189 CMV seropositive mothers (26 CMV transmitter mothers and 163 CMV non-transmitter mothers) who fed their infants CMV-positive maternal breast milk. Although maximum \log_{10} CMV NAT expression in breast milk was associated with maternal-infant CMV transmission among CMV seropositive mothers, the diagnostic accuracy statistics (test sensitivity and 1-specificity) as reflected by the ROC curve were poor. For example, at 3.50 maximum \log_{10} CMV NAT expression in breast milk, test accuracy was 73% sensitivity and 53% specificity (1-specificity = 47%).

An ROC AUC of 1.0 represents a perfect diagnostic predictor (one that has zero false positives and zero false negatives); and an ROC AUC of 0.5 represents discrimination no better than chance (the identify line (i.e., the dotted line) on eFigure 2). The estimated AUC for this study was 0.71 (95% confidence interval 0.62 to 0.80). The AUC for the predictor 'maximum \log_{10} CMV NAT expression in breast milk' corresponds to the probability that of two randomly chosen CMV seropositive mothers who fed their infants CMV-positive maternal breast milk, one who transmits CMV to their low birth weight infant and one who does not transmit CMV to their low birth weight infant, there is a 71% change (equivalent to a 0.71 probability) that the transmitting mother will be ranked with higher suspicion of transmitting CMV to their infant than a mother who does not transmit CMV to their infant.

Assessment of model adequacy for the Cox model

The p value for each time-dependent covariate was not significant ($p > 0.05$) indicating the proportionality assumption was met for each baseline covariate for the CMV cause-specific hazard rates. The p value for each time-dependent covariate (except gender where $p=0.04$) was not significant ($p > 0.05$) indicating the proportionality assumption was met for each baseline covariate for the mortality cause-specific hazard rates. Although the Cox proportional hazards model is the most efficient or powerful when the hazards are

proportional, the model is still valid as long as the cumulative incidence function curves do not cross. The cumulative incidence function curves for mortality did not cross for males and females during the first 2 months of follow-up. The overall conclusion from the regression diagnostic residual plots was that the final fitted Cox regression model provided a reasonable fit to the data. However for mothers with high \log_{10} CMV NAT expression in the breast milk (above \log_{10} 4.0) the martingale residual plot did not suggest a good model fit for the data in this region. However the estimated value of the regression coefficient for CMV NAT expression in the breast milk was reliability estimated ($\beta \pm$ standard error: 0.6263 ± 0.13).

eTable. Factors associated with mother-to-infant CMV transmission by breast milk among CMV seropositive mothers				
Univariable logistic regression analysis of risk factors among 189 CMV seropositive mothers				
	N (%)			
Factors	CMV Transmitters	CMV Non-transmitters	Beta (SE)	Odds ratio (95% CI); P
<i>ALL</i>	26 (14%)	163 (86%)		
Premature rupture of membranes	18 (69%)	53 (33%)	1.560 (0.457)	4.76 (1.94 - 11.6); 0.0006
Rupture of membranes (greater than 18 hours)	10 (38%)	29 (18%)	1.103 (0.454)	3.014 (1.24 - 7.33); 0.015
Caesarean Delivery	17 (65%)	125 (76%)	-0.555 (0.452)	0.57 (0.24 - 1.39); 0.22
Maximum Log ₁₀ CMV NAT expression in breast milk (<i>per 1 log₁₀ increase</i>)			1.012 (0.288)	2.75 (1.57 - 4.84); 0.0004
Multivariable logistic regression analysis of risk factors among 189 CMV seropositive mothers				
Factors			Beta (SE)	Odds ratio (95% CI); P
Premature rupture of membranes			1.491 (0.479)	4.44 (1.74 - 11.36); 0.002
Caesarean Delivery			-0.435 (0.499)	0.65 (0.24 - 1.72); 0.38
Maximum Log ₁₀ CMV NAT expression in breast milk (<i>per 1 log₁₀ increase</i>)			1.024 (0.310)	2.79 (1.52 - 5.12); 0.001
Abbreviations: CMV, cytomegalovirus; SE, standard error; CI, confidence interval.				