Epidemiology and Predictors of Failure of the Infant Car Seat Challenge

WHAT'S KNOWN ON THIS SUBJECT: The American Academy of Pediatrics recommends neonates born at <37 weeks' gestation receive a predischarge Infant Car Seat Challenge, meaning up to 500,000 infants qualify annually. However, little is known about incidence and risk factors for failure in this group.

WHAT THIS STUDY ADDS: This is the largest study to date to examine incidence and risk factors for failure of the Infant Car Seat Challenge. We sought to identify infants most at risk for failure to narrow the scope of testing.

abstract

OBJECTIVES: The American Academy of Pediatrics recommends all neonates born at <37 weeks' gestation receive a predischarge Infant Car Seat Challenge (ICSC), a resource-intensive test with little information on failure rates and risk factors. We sought to determine incidence and predictors of failure to allow more selective testing.

METHODS: We conducted a retrospective medical record review of 1173 premature neonates qualifying for the ICSC between 2009 and 2010. We looked at ICSC result and potential risk factors and then performed bivariate and multivariable logistic analyses to evaluate for predictors of failure.

RESULTS: Overall incidence of failure was 4.3%. Infants who failed were less premature and had higher birth weights. Late-preterm infants made up 60% of our study population but accounted for 78% of failures (P = .019). Infants who passed had older chronologic ages at time of testing, were more likely to have been exposed to caffeine, and were more likely to have required some type of respiratory support than those that failed. Final multivariable model demonstrated that increasing birth gestational age (GA) increased the odds of failure when corrected for gender, race, and small for GA status. For every 1-day increase in birth GA the odds ratio of failure was 1.03 (95% confidence interval 1.01−1.05).

CONCLUSIONS: We found that increasing birth GA was a significant predictor of failure, and that late-preterm infants comprised a significant percentage of infants who failed. This suggests that limiting testing to extremely premature infants would miss most cases of ICSC failure.
Determining readiness for hospital discharge in preterm neonates is extremely challenging. The American Academy of Pediatrics recommends not discharging neonates until “physiologically mature and stable cardiorespiratory function has been documented for a sufficient duration,” whether an infant is supine in a crib, car bed, or upright in a car seat.1,2 This applies to early- and late-preterm infants, all of whom are at risk for morbidity and mortality because of respiratory and neurologic immaturity.3 The best way to determine when neonates reach this level of physiologic maturity is the subject of ongoing debate. One attempt to assess maturation of respiratory control and safety for discharge home is the Infant Car Seat Challenge (ICSC), which has become part of the routine predischarge assessment for preterm infants across the United States, resulting in the testing of up to 500,000 infants annually.4–7 The ICSC is a period of monitoring for apnea, bradycardia, or desaturations while in the semi-upright position in a car safety seat before discharge. It was first recommended based on studies from the 1980s, which observed that preterm neonates had significant desaturations when positioned in a car seat, which persist once replaced supine.8–10

The presumed rationale for performing ICSCs assumes that (1) the degree and duration of cardiopulmonary events seen in ICSC “failures” are clinically relevant, and (2) testing will identify infants who are at risk for subsequent avoidable adverse cardiopulmonary events while in a car seat. These assumptions have never been proven, however, so the predictability of ICSC testing remains unclear. With the growing concern over the effect of intermittent desaturations on long-term neurodevelopmental outcomes, further evaluation of this test that screens for intermittent desaturations becomes more important.11

Current American Academy of Pediatrics guidelines recommend ICSCs for all neonates born at <37 weeks’ gestation, lasting 90 to 120 minutes or duration of travel in a vehicle, whichever is longer.12 Even if ICSC failure accurately predicts potential for subsequent events, it still occupies large amounts of nursing, respiratory therapist, and other professionals’ time. Testing qualifying infants for 90 minutes uses ~750,000 hours of skilled observers’ time annually. More focused screening strategies would decrease unnecessary testing and lead to overall savings to the health care system. Additionally, current information on ICSC failure rates would allow better estimation of sample sizes needed for future studies to determine the correlation between failed ICSC testing and later outcomes. Unfortunately, previous studies attempting to identify incidence and potential risk factors for failure were limited by small sample size, focused on infants with specific comorbidities, such as congenital heart disease, or were performed in the 1980s, compromising their generalizability to today’s NICU graduates.10,13–16

We therefore performed this study to determine the incidence of ICSC failure in a large cohort of premature infants and to characterize significant variables associated with ICSC failure. We hypothesized that infants with evidence of more severe clinical illness and those born more premature would be more likely to fail.

METHODS

Study Population

We conducted a retrospective medical record review of premature neonates born over a 2-year period at Beth Israel Deaconess Medical Center (BIDMC) in Boston, MA, who qualified for ICSC. Inclusion criteria included the following: (1) inborn at BIDMC between January 1, 2009, and December 31, 2010; (2) survived to discharge; and (3) met ICSC testing criteria of birth gestational age (GA) <37 weeks. Exclusion criteria included the following: (1) discharge weight less than the minimum allowed weight for the family’s car safety seat (<5 lb in a standard car seat, <4 lb in a premature car seat), and (2) discharged on a home ventilator. The study was approved by the BIDMC Committee on Clinical Investigation Institutional Review Board. Approximately 20% of the qualifying infants were transferred from BIDMC to local NICUs and special care nurseries (SCNs) before discharge. We therefore obtained institutional review board approval to identify ICSC results from the 5 institutions with the highest census of BIDMC-transferred patients.

ICSC Criteria

ICSCs were performed by the discharging facility per the individual institution’s guidelines and were physically performed in the NICU or SCN regardless of admission location. Infants were placed in their family’s personal car safety seat and positioned per manufacturer’s instructions. Car seat fit assessment was done by staff trained in car seat testing before beginning the test. Heart rate, respiratory rate, saturations, respiratory status, color, and work of breathing were directly observed throughout the duration of the test.

Different monitor models were used at each institution with varied averaging times. BIDMC used Philips Intellivue MP90 monitors (Philips Electronics, Andover, MA) with 20-second averaging times; Boston Children’s Hospital used Philips Intellivue MP70 (Philips Electronics, Andover, MA) with 10-second averaging times; South Shore Hospital used Spacelabs Ultraview 1600 (Ardu Medical, Inc, Fetus, MO) with 8-second averaging times; Beverly Hospital used Philips MP50 (Philips Electronics, Andover, MA) with 10-second averaging
times; Winchester Hospital used Dash 5000 (General Electric Company, Fairfield, CT) with 8-second averaging times; Brockton Hospital used Spacelabs Ultraview 2400 (Ardis Medical, Inc, Festus, MO) with 8-second averaging times. All averaging times are as reported by the manufacturing company and verified with each institution’s biomedical engineering department.

Failure criteria were similar but not identical at the different institutions. All tests lasted a minimum of 90 minutes. The 6 facilities used similar oxygen saturation target cutoffs (4 used <90% and 2 used <88%), with 2 requiring >10-second duration and 1 requiring >20-second duration. All used an identical heart rate cutoff of <80 beats per minute to determine failure; 4 required any drop below 80 beats per minute, and 2 required >10-second duration. If the observer noted alteration in work of breathing or respiratory distress, in consultation with the physician or nurse practitioner, this could also count as an ICSC failure.

Statistical Methods

We compared baseline demographics between subjects whose ICSC results were and were not available. We also compared baseline demographic and clinical information between subjects who passed and those who failed the ICSC. We used Fisher Exact testing, \( \chi^2 \) testing, t testing, and nonparametric Wilcoxon Rank Sum testing as appropriate for binary, categorical, and continuous variables. Logistic regression models were used to model predictors of ICSC failure. Effects are reported as odds ratios and 95% confidence intervals. We included gender, race, and SGA (small for gestational age) status a priori as covariables because of their clinical importance in the outcomes of neonates. SGA status was determined using Fenton Growth Charts.\(^7\) We performed bivariate logistic regression analysis of the crude relationship between the outcome of failure of the ICSC and each of the demographic predictors (birth weight, birth GA, gender, race, singleton versus multiple gestation, SGA status); clinical predictors (mode of delivery, respiratory support requirements, prior treatment with caffeine for apnea of prematurity, postnatal steroids, diuretics at the time of ICSC, antireflux medications at the time of ICSC); and infant characteristics at the time of testing (correct GA [CGA], chronologic age, and weight). We retained those predictors with \( P \leq .2 \) and added back each of the previously eliminated covariables to assess for confounding. We assessed for collinearity in our model by evaluating the effect of each covariable on SE and \( P \) values of the others. All statistical analysis was done by using SAS 9.2 (SAS Institute, Cary, NC).

RESULTS

Over the 2-year study period, 1173 infants met inclusion criteria. Of those, 1135 were ICSC eligible. We were able to do complete chart reviews and obtain ICSC results on 91.4% \( (n = 1036) \). We found a 4.3% \( (n = 45) \) failure rate (Fig 1). Of those who failed, 43 were tested at BIDMC (96%) and 2 were from 1 of the transfer facilities (4%). There was no difference in the proportion of early-versus late-preterm infants who failed between the sites.

Initial demographic information was available on the 8.6% \( (n = 97) \) of infants who qualified for ICSC but did not have available results data. Most of these (86%) were transferred to outside facilities for further convalescent care to other facilities and their final discharge plan is unavailable. Eight remained inpatient for an average of 6.8 additional days after ICSC failure (range 5–12 days) because of respiratory immaturity and desaturations with feeds and at rest. All 8 underwent a 5-day spell count and subsequently passed ICSCs. In total, 51% of those who failed their ICSC \( (n = 23) \) were discharged from the hospital in a car bed, 42% \( (n = 19) \) went on to pass an ICSC and were discharged from the hospital in a car seat, and 7% \( (n = 3) \) were transferred to outside facilities for further care and their discharge information was unavailable.

Predictors of ICSC Failure

There was no statistically significant or clinically important difference in the weight or the CGA at the time of testing between those who passed and those who failed (36\(^{4/7}\) vs 36\(^{5/7}\) respectively) (Table 1). Those who failed were less premature with median birth GA of 36\(^{4/7}\) vs 34\(^{1/7}\) weeks in those who passed \( (P = .0008) \). Consistent with these findings, infants who failed had significantly higher birth weights (2467 vs 2169 g, \( P = .0045) \). Although only 60% of our
The study population were late preterm (34+0/7–36+6/7 weeks birth GA), this group accounted for 78% of failures (P = .019). Infants who passed, therefore, had older chronologic (postnatal) ages at the time of testing when compared with those who failed (P = .0002) and were more likely to have a history of caffeine treatment of apnea of prematurity (Table 2). This represents infants who had been treated in the past for apnea of prematurity. At the time of ICSC, only 5 infants remained on caffeine, 2 for nonclinical indications, as they were participating in a trial of extended caffeine use. Those who passed were more likely to have required continuous positive airway pressure and were overall more likely to have required any type of respiratory support than those who failed (P = .029). Gender, race, mode of delivery, gestation number, or use of antireflux medications, postnatal steroids, or diuretics did not vary between the 2 groups.

The final multivariable model demonstrated that increasing birth GA increased the odds of failure when corrected for gender, race, SGA status. For every one day increase in birth GA, the odds ratio of failure = 1.03 (95% CI 1.01–1.05; Table 3).

**DISCUSSION**

This is the largest study of ICSC testing reported to date, with 91.4% of eligible medical records able to be analyzed. Of those infants who qualified for ICSC based on prematurity (birth GA <37 weeks), we found an overall ICSC failure incidence of 4.3%. This is lower than previously published reports, although most other studies were smaller and focused on particular groups of neonates, such as late-preterm infants or term neonates with specific comorbidities, such as hypoxic cardiac disease.15,16,18

The ICSC is time consuming and costly; limiting unnecessary testing would be beneficial to both individual nurseries and to the overall health care system. To preserve resources, many NICUs limit testing to those born at lower GAs, presumably because of a belief that only extremely premature infants are at risk for failure. In fact, our hypothesis was that infants born more premature would be more likely to fail; however, our study demonstrated that this assumption is incorrect. We found no difference in the CGA or weight at the time of testing between those who passed and those who failed. Those who passed were statistically more likely to be born at smaller birth weights and lower GAs and therefore had older postnatal chronologic ages at the time of testing. They seemed to have been

---

![Flow diagram of study population.](image1)

**FIGURE 1**
Flow diagram of study population.

![Flow diagram of ICSC failure outcomes.](image2)

**FIGURE 2**
Flow diagram of ICSC failure outcomes.
more critically ill, with a higher likelihood of requiring caffeine treatment for apnea of prematurity during their course and requiring some form of respiratory support during their hospitalization. Those who failed were more likely to be born with higher birth weights, older GAs, and had younger postnatal ages at the time of testing. In fact, failures occurred most often in late-preterm neonates. These findings are supported by previous studies that found late-preterm infants (or “near-term” infants born at 35–36 weeks) had a 12% rate of apnea, bradycardia, and desaturations during their ICSC.\textsuperscript{18} We found an overall failure rate of 4.3%, which was made up of a 2.4% failure rate in infants born at <34 weeks vs 5.6% in infants born late preterm. Omitting late-preterm infants from ICSC testing misses most of those infants at risk for desaturation events during ICSCs.

Notably, after correcting for clinically important variables of gender, race, and SGA status, the 1 covariate that significantly predicted ICSC failure in our cohort was advancing birth GA, with the more premature infants being more likely to pass. We found that for each additional day of birth GA, the odds ratio for failure was 1.03 (95% confidence interval 1.01–1.05). In other words, the odds of failing an ICSC increase by 21% for every additional week of birth gestation.

It is unclear why otherwise “healthy” late-preterm infants, who did not require respiratory support or caffeine, were more likely to fail. We believe this may be because of 2 factors: (1) chronologic age (and therefore postnatal maturity) at the time of testing, and (2) the differences in timing and duration of cardiopulmonary monitoring for extremely premature versus late-preterm infants. It is possible that chronologically older infants have had additional time to mature compared with infants with younger chronologic age. The prevalence of many NICU morbidities increases with decreasing GA and therefore extremely premature infants have extended lengths of hospitalization, whereas late-preterm infants often have very short hospitalizations. We know that the incidence of intermittent hypoxic events is likely a function of maturation. Martin et al\textsuperscript{19} showed that the number of intermittent hypoxic events in preterm infants increases in number over the first few weeks of life followed by a decrease in weeks 6 to 8. Because infants in our study cohort were found to have been tested at similar CGA regardless of whether they passed or failed (36+6/7 vs 36+5/7 weeks, respectively), those born more prematurely were actually chronologically older at the time of ICSC testing compared with those born less premature.

In addition, our findings may not be related simply to specific GAs or chronologic ages, but also to how closely infants are monitored. We do know that late-preterm neonates are not physiologically mature even if they do not require NICU admission.\textsuperscript{3} Unlike “healthy” late-preterm neonates, those who are more premature or distressed are admitted to NICUs and placed on monitors that supply continuous information on their saturations, heart rates, and indicators of respiratory immaturity. Once they are deemed

### TABLE 1 Age and Weight Predictors of ICSC Failure

<table>
<thead>
<tr>
<th>Predictor</th>
<th>ICSC Pass, n = 991, Estimate (error)</th>
<th>ICSC Fail, n = 45, Estimate (error)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Birth gestational age, wk\textsuperscript{a}</td>
<td>34.7 (3.9)</td>
<td>36.0 (2.3)</td>
<td>.0008</td>
</tr>
<tr>
<td>Birth weight, g\textsuperscript{a}</td>
<td>2169 (694)</td>
<td>2467 (653)</td>
<td>.0045</td>
</tr>
<tr>
<td>Chronologic age at test, d\textsuperscript{a}</td>
<td>16 (33)</td>
<td>4 (16)</td>
<td>.0002</td>
</tr>
<tr>
<td>GA at test, wk\textsuperscript{a}</td>
<td>36.0 (1.6)</td>
<td>36.9 (2.6)</td>
<td>.08</td>
</tr>
<tr>
<td>Weight at test, g\textsuperscript{a}</td>
<td>2565 (615)</td>
<td>2565 (577)</td>
<td>.71</td>
</tr>
</tbody>
</table>

\textsuperscript{a} Denotes effect estimate reported in median with SE reported via interquartile range.

### TABLE 2 Unadjusted Demographic and Clinical Predictors of ICSC Failure

<table>
<thead>
<tr>
<th>Predictor</th>
<th>ICSC Pass, n = 991, n (%)</th>
<th>ICSC Fail, n = 45, n (%)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td>.13</td>
</tr>
<tr>
<td>Male</td>
<td>538 (54)</td>
<td>30 (67)</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>453 (46)</td>
<td>15 (33)</td>
<td></td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td></td>
<td>.15</td>
</tr>
<tr>
<td>White</td>
<td>614 (62)</td>
<td>26 (58)</td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>135 (14)</td>
<td>3 (7)</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>242 (24)</td>
<td>16 (35)</td>
<td></td>
</tr>
<tr>
<td>Delivery mode</td>
<td></td>
<td></td>
<td>.27</td>
</tr>
<tr>
<td>Vaginal</td>
<td>375 (38)</td>
<td>13 (29)</td>
<td></td>
</tr>
<tr>
<td>Cesarean delivery</td>
<td>616 (62)</td>
<td>32 (71)</td>
<td></td>
</tr>
<tr>
<td>Gestation number</td>
<td></td>
<td></td>
<td>.75</td>
</tr>
<tr>
<td>Singleton</td>
<td>624 (63)</td>
<td>30 (67)</td>
<td></td>
</tr>
<tr>
<td>Multiples</td>
<td>367 (37)</td>
<td>15 (33)</td>
<td></td>
</tr>
<tr>
<td>Respiratory support required</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Any</td>
<td>410 (41)</td>
<td>11 (24)</td>
<td>.029</td>
</tr>
<tr>
<td>Ventilator</td>
<td>204 (21)</td>
<td>5 (11)</td>
<td>.13</td>
</tr>
<tr>
<td>CPAP</td>
<td>345 (35)</td>
<td>7 (16)</td>
<td>.065</td>
</tr>
<tr>
<td>Nasal canula</td>
<td>124 (13)</td>
<td>8 (18)</td>
<td>.56</td>
</tr>
<tr>
<td>Postnatal Medications</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caffeine</td>
<td>214 (22)</td>
<td>3 (7)</td>
<td>.014</td>
</tr>
<tr>
<td>Diuretics</td>
<td>17 (2)</td>
<td>0 (0)</td>
<td>1</td>
</tr>
<tr>
<td>Antireflux</td>
<td>34 (3)</td>
<td>2 (4)</td>
<td>.67</td>
</tr>
<tr>
<td>Postnatal Steroids</td>
<td>9 (1)</td>
<td>1 (2)</td>
<td>.36</td>
</tr>
<tr>
<td>SGA</td>
<td>80 (8)</td>
<td>3 (7)</td>
<td>1</td>
</tr>
</tbody>
</table>

CPAP, continuous positive airway pressure.
mature and appropriate for discharge, they then have their ICSC performed. Although all preterm infants are at risk for desaturations related to respiratory immaturity, the only time healthy preterm neonates are screened for desaturations is during the ICSC. In fact, we found that 35% of preterm infants who failed their ICSC while admitted to the well-infant nursery required subsequent NICU admission for an average of almost 1 week because of persistent respiratory immaturity, which was identified based solely on the requirement for a screening ICSC, not because of noted respiratory distress at baseline. These findings are concerning, given recent work indicating that late-preterm infants have worse neurodevelopmental outcomes than their term counterparts. It is possible that unrecognized respiratory immaturity in this population could be contributing to these findings, but more research will need to be done to evaluate this hypothesis.

A limitation of this study is that we were unable to obtain results on all eligible subjects. The number of subjects who were eligible for ICSC but did not have ICSC results was small (8.6%). Those who did not have results were more likely male, born via cesarean delivery, and born at smaller birth weights and lower GAs. Because this is a small minority of the eligible patients, we do not feel that these missing data would significantly alter our findings.

ICSCs were not performed in the same style of car safety seat for each neonate in the study, but instead were performed in the family’s personal car seat that was planned to be used after discharge. This is standard practice for ICSCs across the nation because the aim of the study is to demonstrate safety in the infant’s own car seat.

Another potential source of variation was the difference in monitors and averaging times at each institution (from 8–20 seconds). Shorter averaging times are more sensitive to detect brief intermittent desaturations but increase the rate of false alarms because of motion artifact. Longer averaging times have less likelihood of false alarms but may underestimate the frequency of desaturations. There is growing concern regarding the effect of brief intermittent desaturations on neurodevelopmental outcomes in preterm infants. Recent prospective studies use 2-second averaging times to evaluate intermittent desaturations in neonates. Because this was a retrospective study to demonstrate current practice, we did not have control over the averaging times set at each NICU; however, national recommendations to use shorter averaging times as the standard in all NICUs do not currently exist. The standard practice seems to be the use of the default setting for each institution’s model of monitor. As a result, most units do use longer averaging times. Our results are therefore most generalizable to NICUs that have similar testing methods and that use monitors with similar averaging times.

Of note, the vast majority of subjects (>80%) were tested at BIDMC, which has the longest averaging time (20 seconds), and 96% of failed ICSCs occurred there. We feel that because of the longer averaging time, it is possible that our data underestimate the true failure rate in the population. If the averaging time were shorter, it is possible that more infants would have met clinical failure criteria. However, even if 4.3% of preterm infants failed their ICSC, this would indicate >20 000 failures in preterm infants annually in the United States alone.

An additional limitation of any study of ICSC testing is the lack of national standards to define ICSC failure and the resulting variability of failure criteria between institutions. The hospitals represented in our study have very similar but not identical failure criteria; however, on review, all of the 45 infants who failed the test would have failed at all institutions. Further studies of the ICSC may help to create nationwide guidelines for failure criteria.

We do not feel that these differences in failure criteria or averaging times biased the findings of predictors of failure. Regardless of location of testing, these differences were not specific for GA or any of the other clinical or demographic factors we evaluated. To evaluate possible bias related to these differences, we did a subgroup analysis of infants tested at BIDMC versus the entire cohort from the 6 institutions and our outcomes and conclusions were unaffected.

### Conclusions

In summary, our study identified a relatively low but significant incidence of ICSC failure in a large premature birth cohort. This number may be an underestimate of failure rates in the general population because of longer averaging times at our main site. We found that increasing birth GA was a significant predictor of failure, which is likely related to a younger chronologic age at the time of testing, and that late-preterm infants comprised a significant percentage of the infants who failed. This suggests that limiting testing
to extremely premature infants would miss most cases of ICSC failure. No other predictors of ICSC failure were identified to potentially allow more selective testing. No studies have shown that the ICSC results directly predict subsequent desaturation episodes during upright positioning in car seats. Further studies of the ICSC are ongoing at our institution to confirm or refute its prognostic significance. This study provides important epidemiologic data regarding ICSC failure that will allow planning and recruitment for future studies to confirm the relevance and validity of ICSC testing.

ACKNOWLEDGMENTS
The authors thank Dr. Ivana Culic, Dr. Susan Hwang, Dr. Karen McAlmon, and Dr. Ozlem Kasaroglu for their assistance at their respective sites; Dr. Joseph Ghou at MGH for assistance with Fenton Growth Chart data; Dr. John Zupancic, Dr. Stella Kourembanas, and the Division of Newborn Medicine at Boston Children’s Hospital; and the Scholars in Clinical Science Program at Harvard Catalyst for ongoing support of Dr. Davis.

REFERENCES

Epidemiology and Predictors of Failure of the Infant Car Seat Challenge
Natalie Louise Davis, Freeman Condon and Lawrence M. Rhein

Pediatrics 2013;131;951; originally published online April 1, 2013;
DOI: 10.1542/peds.2012-2775

The online version of this article, along with updated information and services, is located on the World Wide Web at:
/content/131/5/951.full.html