jmp
Statistical Discovery.™ From SAS.

CHALLENGE

Help Kaiser Permanente reduce adverse events like shock, respiratory failure and brain damage in newborn babies.

SOLUTION

JMP's modeling and recursive partitioning identified predictors of respiratory distress that then were developed into a scoring system, the Richardson Score, that hospitals can use to identify at-risk newborns.

RESULTS

Nurses or physicians in the Kaiser Permanente Medical Care Program assign appropriate staffing or transfer at-risk newborns to a neonatal intensive care unit for care.

MORE INFORMATION

www.jmp.com

www.kp.org

Kaiser Permanente research team seeks to prevent catastrophic outcomes in babies

JMP® software plays an essential role in developing predictive models

Dr. Gabriel Escobar's work is not only about helping newborns enter the world. It's also about helping them enter it safely.

"Our primary research objective is to develop tools that physicians and nurses can use to prevent newborns from suffering catastrophic consequences," Escobar says.

Escobar is the Director of the Perinatal Research Unit at Kaiser Permanente's Division of Research in Oakland, CA. One of his current research projects aims to develop models to predict which newborns with respiratory distress are at risk for shock, respiratory failure or death.

"Nowadays," Escobar explains, "the number of newborns who die is extremely small. For the most part, we've made tremendous progress in terms of infant mortality for babies of all gestations. But where we haven't made as much progress is in preventing infant morbidity. So our ultimate goal is not only to increase the number of survivors, but to improve the quality of life of the survivors. We want to prevent babies from going into

shock or respiratory failure and developing brain damage."

To do that, physicians and nurses need to know when, how and for which babies to intervene. That's where statistical modeling comes in, and where JMP statistical discovery software from SAS is now playing a major role in Escobar's research.

The importance of the first pass at the data

"In our research, we develop mathematical models to predict the occurrence of adverse outcomes in patients. The problem is that sometimes we have so much data that it is hard to know just where to begin. Also, because hindsight is so powerful, human beings have a tendency to try to replicate past analyses, rather than taking fresh looks," Escobar says.

One critical solution is recursive partitioning, which randomly splits the data set into learning and validation data sets. It has the advantage of finding data patterns that can be missed, especially if one wants to avoid biases based on past experience.

STATISTICAL DISCOVERY.™ FROM SAS.

"Our primary use of recursive partitioning is to identify clusters of patients for whom the likelihood of a bad outcome is extremely high. We then characterize those clusters to define the components of more sophisticated models. They help us zero in on these important subsets, often identifying predictors or combinations of predictors that we would never have considered," Escobar explains.

Take, for example, his recent study of term and near-term babies with respiratory distress. Evaluation of newborns with respiratory distress is one of the most common problems in neonatal care. Between 2 percent and 3 percent of all newborns will experience some form of respiratory distress.

Arterial blood gas measurements can be used to assess newborns with respiratory problems. But, as Escobar wrote in the report on his study, "interpreting physiologic data remains more of an art than a science."

Using data from a large cohort of Kaiser Permanente newborns, Escobar first split the data into derivation and validation data sets. Following bivariate analyses, the team performed a number of exploratory analyses on the derivation data set using recursive partitioning. Recursive partitioning identified clusters of infants with high rates of the adverse outcomes (prolonged assisted ventilation and/or death). Escobar and his team then employed cutoff values for predictors identified with recursive partitioning in "traditional" logistic regression models.

The coefficients from their final model were then employed to define a simple scoring system that was validated on the data set held in reserve for that purpose.

The score is based on an infant's gestational age, the worst arterial blood gas result in the first 12 hours of observation and the worst mean arterial blood pressure in those first 12 hours. It's called the Richardson Score, in honor of the late Dr. Douglas K. Richardson, a pioneer in neonatal health services research.

Escobar says there are several potential uses for the Richardson Score, which has been disseminated throughout the Kaiser Permanente Medical Care Program.

The first is to serve as an objective measure of risk. "Because the score is so simple," Escobar writes in his report, "nurses or physicians can assign it quickly and get a good sense as to a given baby's likelihood of deterioration. Knowing that a baby is likely to require prolonged assisted ventilation is also important from a staffing standpoint."

Knowing that a baby's risk of requiring additional care is high could play an important role in the decision to transport the infant to a different hospital, one with a neonatal Intensive Care Unit (ICU). Escobar points out that more than half of US hospitals do not have neonatal intensive care units. In the Kaiser Permanente Medical Care Program, pediatricians receive laminated cards that explain the scoring

system, which they can use to decide whether a baby should be transported to a hospital with a neonatal ICU.

"Another potential use for this score is as a risk-adjustment tool," Escobar continues. "Finally, the score could be used to assign babies with respiratory distress to different arms of randomized clinical trials."

SAS' JMP® software: a scalable and cost-effective analytic alternative

Given how tight funding has become, resource allocation is a critical component of a research team's success. This is one of the primary reasons Escobar has embraced JMP.

"The cost of some proprietary software is simply unattainable for somebody with a small research unit," he explains. Ease of use is another key consideration. "For example," he adds," if we were to use standalone recursive partitioning software, that would mean that in addition to the cost of the license, we would need to expend money on application-specific training for such software.

"On the other hand, the reason I like JMP so much is that my project managers can learn it quickly, including its recursive partitioning component. Moreover, it turns out that they can also use JMP for another dozen or so analytic tasks that fulfill critical functions for the group."

Escobar says JMP is suitable for people who are not full-time researchers. But this does not mean that JMP cannot be employed for what he calls "heavy lifting."

"In my unit, we have professional programmer analysts who do the most complex programming. Let's say, for example, that we need to look at the Kaiser Permanente birth cohort over four years — which is 120,000 records — and we want to see how many of those had mothers who were screened for Group B streptococcus, which is a bacterium that causes newborn infections. Well, that involves linking 120,000 baby records to 120,000 mommy records to maybe 200,000 laboratory test results. That kind of work is reserved for our professional programmers. "

"On the other hand," continues Escobar, "once those programmers have built the huge data set and we need to know what the rate of screening for Group B streptococcus was in, for example, 20- to 25-year-old women at facility X, we can use our project managers, who know less programming, to use JMP for those simple data queries and analyses, or for recursive partitioning."

This is possible because JMP is graphical and easy to use.

"Given the shortage that we have of professional programmer time, this is really a lifesaver," Escobar says.

KAISER PERMANENTE

Big Baby Breathing Hard

CRUCIAL MEASURES *for the* **EARLY IDENTIFICATION** *&* **TIMELY TREATMENT** *of* **RESPIRATORY DISTRESS** *in* **INFANTS ≥34 WEEKS GESTATION**

- Identify all infants with one hour or more of respiratory distress.
  (Definition: persistent grunting, persistent retractions and/or flaring, SaO_2 <94% in room air, often coupled with poor perfusion and/or central cyanosis)
  - With severe respiratory distress begin treatment immediately.
  - Just an increased respiratory rate doesn't predict a poor outcome.
  - Many normal infants will breathe 60-90 BPM after birth.
- Place in 100% O_2 by hood.
- Place an oximeter.
- Order a chest x-ray.
- Measure, correct, and follow the mean blood pressures (MBPs).
- Obtain an arterial blood gas via an arterial stick. If necessary place a UAC. Capillary or venous blood gases cannot be used to assess these infants. This applies to:
  - Infants with persistent respiratory distress ≥ one hour.
  - Infants requiring supplemental oxygen ≥ one hour.
- Obtain a blood culture and CBC and start on antibiotics.
- Check, correct, and follow glucose levels.
- Correct hypothermia.
- Assess the Richardson Score *(See below).*

NB: ***Call your covering/consulting Neonatologist*** when the infant is ***between 2-3 hours of age***. Be ready to discuss whatever test or x-ray results are available.

NB: Even if you have been unable to generate all the necessary data or do a Richardson Score ***a phone call should be made to the Neonatologist between 2-3 hours of age***. This applies 24/7.

RICHARDSON SCORE			
Gestational age	38 or more	=	0
	<38 weeks	=	1
Worst arterial pH	≥7.30, any pCO_2	=	0
(and concomitant pCO_2)	7.25-7.29, pCO_2 <50	=	0
	7.25-7.29, pCO_2 ≥50	=	1
	<7.25, any pCO_2	=	2
Ratio of	2.5 or more	=	0
Arterial PaO_2	1.2-2.49	=	2
% Inspired O_2	<1.2	=	4
Lowest mean BP	≥30 mm Hg	=	0
	<30 mm Hg	=	1
TOTAL RICHARDSON SCORE		=	___

©KP Nursery Directors of Northern California

INTERPRETATION of RICHARDSON SCORE		
SCORE	RISK *of* ASSISTED VENTILATION FOR >3 DAYS *or* DEATH	RISK *of* ECMO, NITRIC OXIDE *and/or* DEATH
0	1.2%	0.1%
1	5.3%	0.0%
2	4.4%	0.8%
3	15.0%	1.9%
4	17.8%	3.0%
5	31.5%	3.2%
6	43.8%	13.5%
7	51.4%	8.6%
8	50.0%	20.0%

Adapted from G. Escobar, Revised 10/21/03

Kaiser Permanente disseminates instructions for calculating the Richardson Score and using it to determine if a baby is likely to require prolonged assisted ventilation. The Richardson Score was named in honor of Dr. Douglas K. Richardson, a pioneer in neonatal health services research. Poster courtesy of KP Nursery Directors of Northern California.

Exploring additional research opportunities

Escobar also has been using JMP for his research on adults.

"It is very similar to what we have done with the newborns," states Escobar. "We are developing models to predict which adults will experience deterioration while in the hospital and may then end up needing to go into the ICU. So, in the same way that we developed the models for newborns, we use recursive partitioning to identify these clusters of patients."

JMP is not just handy for model building and recursive partitioning. It is also quite adept at displaying different kinds of data. These can range from different hospitals' percent capacity on a given day in relation to patients' severity of illness — what Escobar calls "a macro level view of hospital kinetics" — to a single display showing five vital-signs time series for an individual patient.

Going forward, Escobar has plenty of plans for JMP: "We're now developing models for predicting the risk of an adult being hospitalized and/or dying within six months from some T-zero. Ideally, since some causes of hospitalization or death due to chronic illness are preventable through evidence-based strategies, being able to detect patients at risk earlier could be used to improve outcomes.

"So our next generation of projects is going to be on predicting risk for multiple adult outcomes, and we certainly intend to be using JMP for them as well," Escobar says.