

# SNAP-II and SNAPPE-II and the Risk of Structural and Functional Brain Disorders in Extremely Low Gestational Age Newborns: The ELGAN Study

Olaf Dammann<sup>a–c</sup> Mary Naples<sup>d</sup> Francis Bednarek<sup>d</sup> Bhavesh Shah<sup>e</sup>  
Karl C.K. Kuban<sup>f</sup> T. Michael O’Shea<sup>g</sup> Nigel Paneth<sup>h</sup> Elizabeth N. Allred<sup>b</sup>  
Alan Leviton<sup>b</sup> on behalf of the ELGAN Study Investigators

<sup>a</sup>Division of Newborn Medicine, Floating Hospital for Children at Tufts Medical Center, and

<sup>b</sup>Neuroepidemiology Unit, Departments of Neurology, Children’s Hospital and Harvard Medical School, Boston, Mass., USA; <sup>c</sup>Perinatal Neuroepidemiology Unit, Departments of Gynecology and Pediatrics, Hannover Medical School, Hannover, Germany; <sup>d</sup>Division of Neonatology, University of Massachusetts Medical Center, Worcester, Mass., <sup>e</sup>Baystate Medical Center and Tufts Medical School, Springfield, Mass.,

<sup>f</sup>Division of Pediatric Neurology, Boston Medical Center, Boston, Mass., <sup>g</sup>Department of Pediatrics, Wake Forest University Health Sciences, Winston-Salem, N.C., <sup>h</sup>Department of Epidemiology, College of Human Medicine, Michigan State University, East Lansing, Mich., USA

## Key Words

Scores for Neonatal Acute Physiology · Extremely Low Gestational Age Newborns · Neurodevelopmental dysfunctions

## Abstract

**Background:** Illness severity measures predict death and illnesses in the newborn. It is unknown how well they predict brain lesions evident on ultrasound scans or neurodevelopmental dysfunctions in preterm infants. **Methods:** A total of 1,399 inborn infants born before the 28th week of gestation were given Scores for Neonatal Acute Physiology (SNAP-II and SNAPPE-II) based on data collected within the first 12 h of admission to the intensive care unit and had a protocol brain ultrasound scan read independently by 2 sonologists. Of the surviving 1,149 infants, 1,014 (88%) had a neurologic examination at approximately 24 months post-term equivalent, and 975 (85%) had a Bayley Scales of Infant Development assessment. SNAP-II and SNAPPE-II were dichotomized at arbitrary cut-offs (30 for SNAP-II and 45 for SNAPPE-II), using the highest quartile and decile of the week of gestation

as a cut-off, and at a Z score of >1 standard deviation from an external mean. **Results:** After adjustment for gestational age, high SNAP-II and SNAPPE-II scores predicted intraventricular hemorrhage, moderate/severe ventriculomegaly and echodense lesions in cerebral white matter. Only 2 SNAP-II extremes, the highest decile for gestational age and a Z score >1, also predicted echolucent lesions in the white matter. Neither SNAP-II nor SNAPPE-II predicted any statistically significant diagnosis of cerebral palsy. MDI and PDI scores <55 were consistently predicted by both high SNAP-II and SNAPPE-II, whereas scores in the 55–69 range were inconsistently predicted. High SNAP-II and SNAPPE-II inconsistently predicted a positive screen for autism spectrum disorder and small head circumference at 24 months. **Conclusion:** The physiologic instability in the first 12 post-natal hours identified by illness severity scores conveys information about the risks of brain damage and neurodevelopmental dysfunctions. This risk information might reflect postnatal characteristics in the causal chain. On the other hand, high SNAP scores might be indicators of immaturity and vulnerability.

Copyright © 2009 S. Karger AG, Basel

## Introduction

Illness severity scores were designed to estimate the mortality risk based on data obtained shortly after admission to an intensive care unit (ICU) [1]. They were created to avoid attributing to suboptimal care what should be attributed to each individual's endogenous risk of dying in the ICU. The goal was to adjust each ICU's mortality rate using the illness severity scores of those admitted to that unit. Then, the mortality rates of the ICUs could be compared along with their medical care practices in order to find ways to reduce the risk of death among ICU admissions. This has been successful for studies of adults [2] and newborns [3, 4].

Severity of illness among very preterm newborns is not only associated with mortality, but also with neonatal interventions [5] and neonatal morbidities, such as intraventricular hemorrhage [6], chronic lung disease [6] and retinopathy of prematurity [7]. This should not come as a surprise since some early neonatal characteristics of scores such as the Score for Neonatal Acute Physiology (SNAP) overlap with characteristics/correlates of neonatal morbidity [8].

We know of only one study that evaluated SNAP scores and developmental outcomes [9]. It calculated the SNAP every day for 96 infants born before the 31st week of gestation. Infants in the highest quartile of cumulative daily SNAP scores did not do as well as their peers on multiple measures of function at 2–3 years of age, including the Bayley Scales of Infant Development, 2nd Edition, and the Receptive-Expressive Emergent Language Scale.

Since that study, SNAP has been simplified, resulting in two modifications (SNAP-II and SNAPPE-II) [10]. We wanted to see if the illness severity measures collected only during the first 12 postnatal hours in a sample of extremely premature infants predicted neonatal cerebral abnormalities identified on cranial ultrasound images and later neurodevelopmental dysfunctions.

## Methods

The ELGAN (Extremely Low Gestational Age Newborns) Study was designed to identify characteristics and exposures that increase the risk of structural and functional neurologic disorders in ELGANs. From 2002 to 2004, women who delivered before 28 weeks gestation at 14 participating institutions in 11 cities in 5 states were asked to enroll in the study. The enrollment and consent processes were approved by the individual institutional review boards.

Mothers were approached for consent either upon antenatal admission or shortly after delivery, depending on clinical circum-

**Table 1.** Sample description

	Available	
	no	yes
Patients		1,506
SNAP-II and SNAPPE-II calculated	71	1,435
Had an ultrasound	36	1,399
Eligible for the 24-month assessment	250	1,149
Had a neurological exam at 24 months (CP diagnosis)	135	1,014
Had GMFCS at 24 months	137	1,012
Had Bayley MDI and PDI at 24 months	174	975
Had M-CHAT at 24 months	107	1,042
M-CHAT and GMFCS <1 at 24 months	241	908
M-CHAT and GMFCS <1 at 24 months and no ear or eye issues	264	885
HC measured at 24 months and at birth	185	964

stances and institutional preference. One thousand two hundred and forty-nine mothers of 1,506 infants consented. Approximately 260 women were either missed or did not consent to participate. We excluded an additional 71 children because they were not eligible for SNAP-II scoring or did not have all the components for calculating SNAP-II [10], and an additional 36 infants who did not have a single protocol ultrasound scan were also excluded (table 1). The remaining 1,399 inborn infants constituted the sample for this study.

After excluding all deaths, 1,149 infants were eligible for a neurodevelopmental assessment at approximately 24 months post-term equivalent. Not all children evaluated had every component of the assessment protocol. All children who had each assessment were included in the sample for that assessment. Thus, the denominator varies from 964 for head circumference outcomes to 1,042 for the Modified-Checklist for Autism in Toddlers (M-CHAT).

### *Newborn Variables*

Data were collected by research nurses trained specifically for the ELGAN study. The gestational age estimates were based on a hierarchy of the quality of available information. Most desirable were estimates based on the dates of embryo retrieval or intrauterine insemination or fetal ultrasound before the 14th week (62%). When these were not available, reliance was placed sequentially on a fetal ultrasound at 14 or more weeks (29%), last menstrual period (LMP) without fetal ultrasound (7%) and gestational age recorded in the log of the neonatal ICU (1%).

We collected all the physiology, laboratory and therapy data for the first 12 h needed to calculate a SNAP-II score [10]. SNAP-II includes points for the lowest mean blood pressure, lowest temperature, lowest pH, respiratory dysfunction (the lowest of PaO<sub>2</sub>/FiO<sub>2</sub> ratios at 3 points), low urine output and seizures. In addition to these, SNAPPE-II includes points for low birth weight, low 5-minute Apgar score and being small for gestational age. Moribund infants were excluded from SNAP data collection.

We arbitrarily selected a SNAP-II value of 30 or more as high, which identified 28% of our sample, and a SNAPPE-II above a cut-off of 45, which was the score for 33% of the newborns. We also identified cut-offs for each week of gestational age that defined the top quartile and top decile.

The Vermont Oxford Network (VON) SNAP Pilot Project provided the median and standard deviation of SNAP-II and SNAPPE-II for each week of gestation [4]. This allowed us to create a Z score SNAP-II and SNAPPE-II Z score for each newborn in our study.

The SNAP-II Z score is the difference between the observed SNAP-II and the mean SNAP-II for the same gestational age in the VON sample divided by the standard deviation of the SNAP-II at that gestational age. Because they incorporate the standard deviation, Z scores provide information about SNAP-II variability at each gestational age. They are expressed as units of standard deviations from the gestational age-specific mean and follow a Gaussian distribution with a mean of 0 and variance of 1.

#### *Ultrasound Protocol Scans*

Routine scans were performed by technicians at all of the hospitals using digitized high-frequency transducers (7.5 and 10 MHz). Ultrasound studies always included the 6 standard quasi-coronal views and 5 sagittal views using the anterior fontanel as the sonographic window [11].

The 3 sets of protocol scans were defined by the postnatal day on which they were obtained. Protocol 1 scans were obtained between the 1st and 4th day (n = 1,075); protocol 2 scans were obtained between the 5th and 14th day (n = 1,247) and protocol 3 scans were obtained between the 15th day and the 40th week (n = 1,215).

After creation of a manual and data collection form, observer variability minimization efforts included conference calls discussing aspects of images prone to different interpretations [12]. Templates of multiple levels of ventriculomegaly were included in the manual.

All ultrasound scans were read by two independent readers who were not provided clinical information. Each set of scans was 1st read by 1 study sonologist at the institution of the infant's birth. The images, usually as electronic images on a CD imbedded in the software eFilm Workstation™ (Merge Healthcare/Merge eMed, Milwaukee, Wisc., USA), were sent to a sonologist at another ELGAN study institution for a 2nd reading. The eFilm program allowed the 2nd reader to see what the 1st reader saw, and provided options to adjust and enhance the studies similar to the original reader, including the ability to zoom and alter gains.

When the two readers differed in their recognition of intraventricular hemorrhage, moderate/severe ventriculomegaly, echodense (hyperechoic) lesion and echolucent (hypoechoic) lesion, the films were sent to a 3rd (tie-breaking) reader who did not know what the first 2 readers had reported.

#### *24-Month Developmental Assessment*

Families were invited to bring their child for a developmental assessment close to the time when s/he would have a corrected age of 24 months. Fully 91% of children had this developmental assessment, which included a neurological examination and the Bayley Scales of Infant Development, 2nd Edition [13]. Of these children, 77% had their exam within the range of 23.5–27.9 months. All Bayley Scales assessments were age-adjusted as appropriate.

The parent or caregiver accompanying the child was asked to complete the M-CHAT [14].

#### *Head Circumference*

The head circumference was measured as the largest possible occipital-frontal circumference. Measurements were rounded to the closest 0.1 cm when taken at birth and examined at 24 months (corrected age). All head circumferences are presented as Z scores because newborns were assessed at different gestational ages at birth (23–27 weeks) and at different age-corrected approximations of 24 months (range: 16–44 months corrected age, with 68% assessed at 23–25 months corrected age). Z scores were based on standards in the CDC data sets [15].

#### *Cerebral Palsy*

Neurologic examiners used a manual, a data collection form and an instructional CD designed to minimize examiner variability, and demonstrated acceptably low variability [16]. The topographic diagnosis of cerebral palsy (CP; quadriplegia, diparesis or hemiparesis) was based on an algorithm created using these same sources of information [17]. Children allocated to 1 CP diagnosis differed from their peers with the 2 other CP diagnoses in their score on the Gross Motor Functional Classification Scale (GMFCS) [18], as well as in the frequency of microcephaly, cognitive impairment and M-CHAT positivity. Only 4% of the examiners indicated at the time of the examination that they had knowledge of the child's brain-imaging studies.

#### *GMFCS*

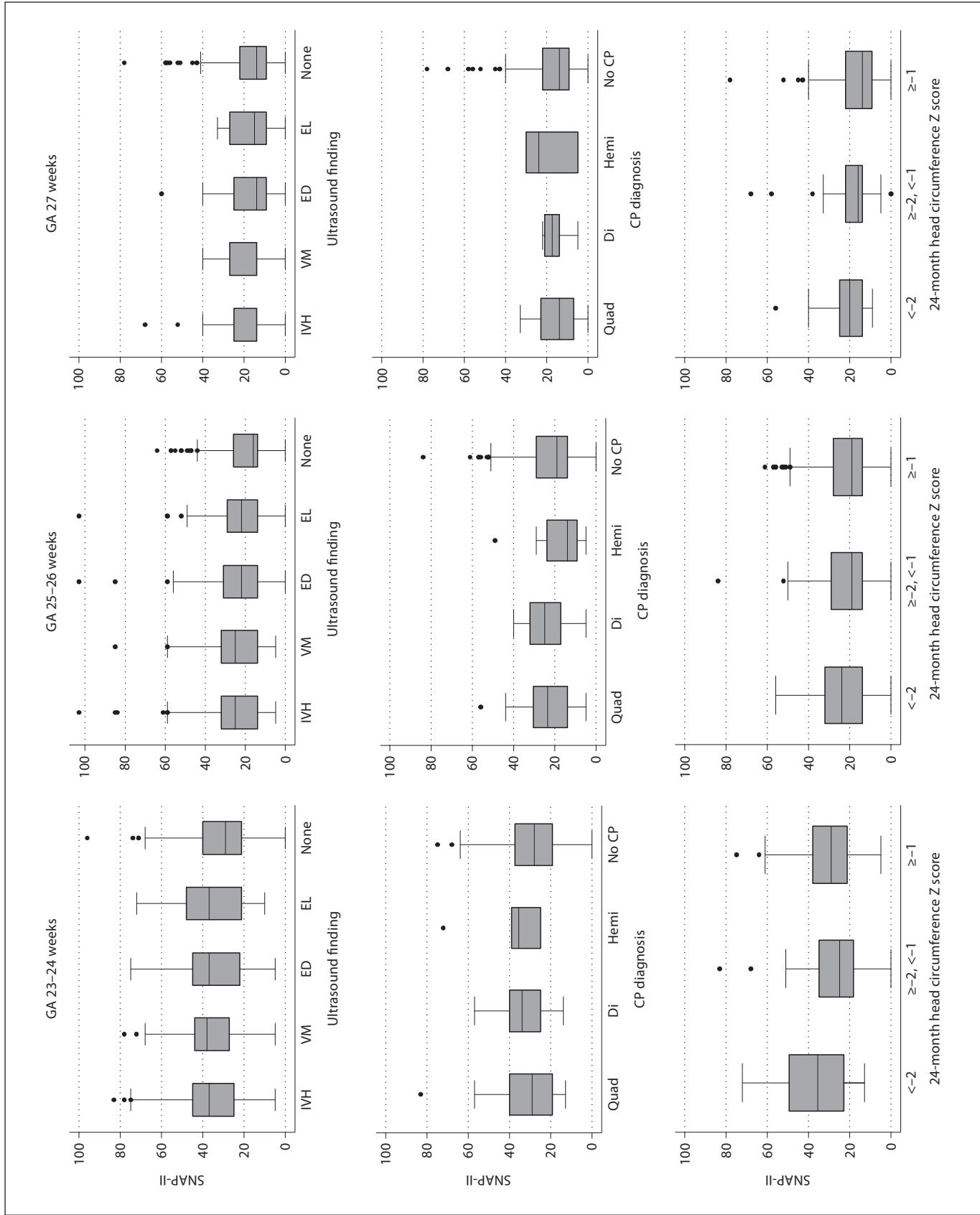
In addition to performing the neurological examination, examiners rated children on the GMFCS, separate from the neurological examination.

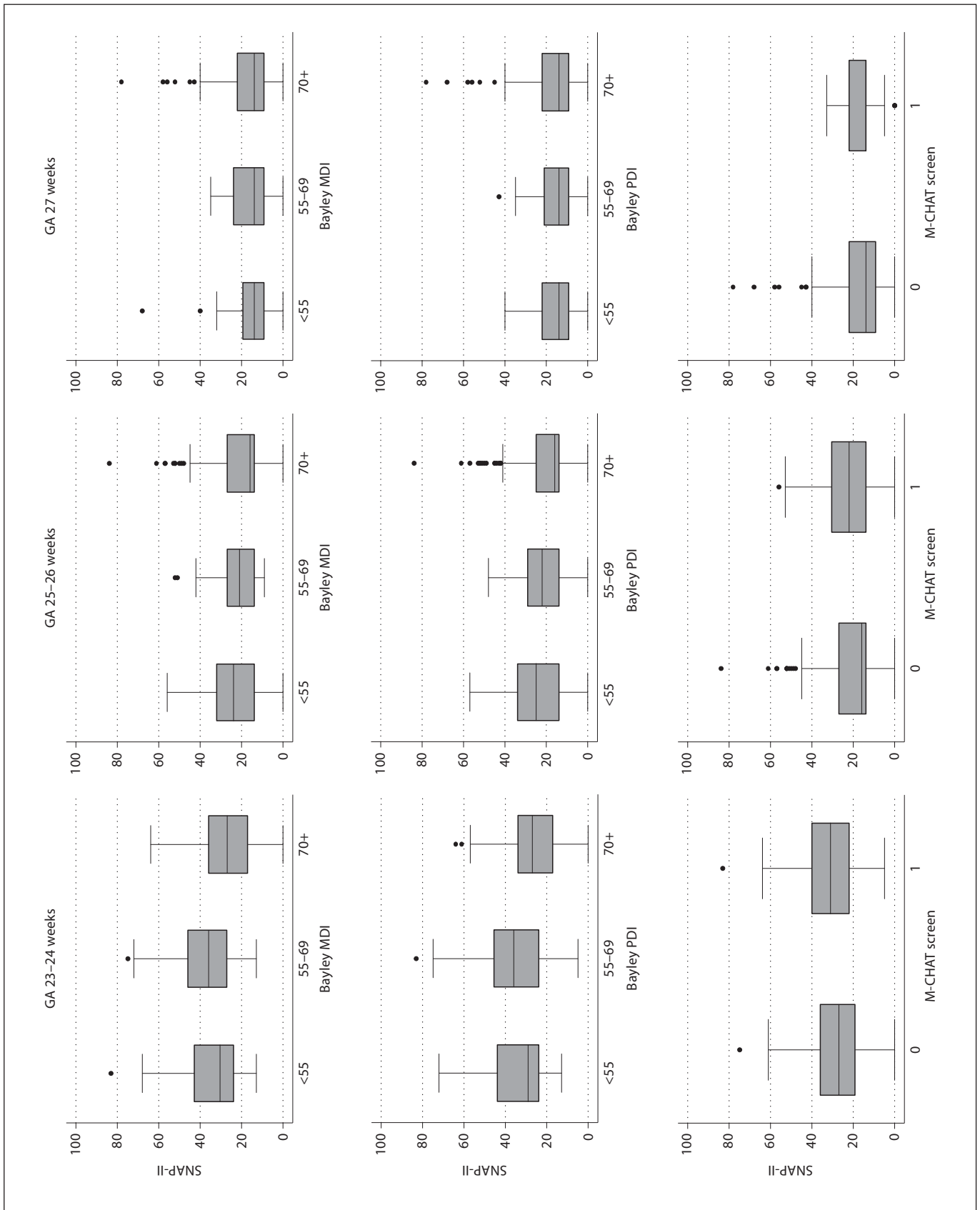
#### *Bayley Scales of Infant Development, 2nd Edition*

Certified examiners administered and scored the Bayley Scales of Infant Development, 2nd Edition [13]. Before testing, the examiners were told the child's age. After completion of testing, they were told the gestational age so that the unadjusted mental developmental index (MDI) and psychomotor developmental index (PDI) could be obtained. Only 2% of the examiners indicated at the time of the examination that they had more than a limited amount of information about the child.

(For figure see next pages.)

**Fig. 1.** Box and whiskers displays of the central tendency and dispersion of SNAP-II in gestational age groups. The central tendency is indicated by the line close to the middle of the box, which is the median, and by the top and bottom of each box, which indicate the 25th and 75th centiles. The dispersion of SNAP-II is indicated by the length of the vertical lines that emanate from the box, as well as by the black dots, which identify outliers. GA = Gestational age; IVH= intraventricular hemorrhage; VM = ventriculomegaly; ED = echodense lesion; EL = echolucent lesion; Quad = quadriplegia; Di = diparesis; Hemi = hemiparesis.





The child was classified as nontestable if her/his impairments prohibited standardized administration or more than 2 items were judged to be 'not applicable'. On the basis of their score on scale No. 5 of the Vineland Adaptive Behavior Scales, 26 of 33 children considered nontestable were assigned an MDI equivalent of <70 (n = 23) or 70+ (n = 3). On the basis of the motor scale (No. 4) of the Vineland Adaptive Behavior Scales, 32 of 38 children considered nontestable were assigned a PDI equivalent of <70 (n = 27) or 70+ (n = 5). The 7 (6) children not assigned an MDI (PDI) equivalent had no Vineland score available.

#### *M-CHAT*

The parent or caregiver accompanying the child was asked to complete the M-CHAT [14]. A child screened positive if 2 of 6 'critical' items were identified or any 3 of the 23 items. Because vision, hearing and motor limitations might account for some children screening positive for an autism spectrum disorder, we limited 1 set of analyses to children who were not blind in either eye, did not wear a hearing aid or receive services for the hearing impaired, and had a GMFCS <1, indicating they had no difficulty walking.

#### *Data Analysis*

SNAP-II and SNAPPE-II decline with increasing gestational age (fig. 1). We considered the possibility that a SNAP-II of 35 at 23 weeks conveys different information than a SNAP-II of 35 at 27 weeks, prompting us to classify newborns by their SNAP scores within gestational age strata. This is the equivalent to creating internal-based Z scores. Because we also wished to use an external standard, we created Z scores based on the VON data.

Each child was classified by 4 dichotomies of SNAP-II (an arbitrary cut-off at 30, the highest quartile for gestational age, the highest decile for gestational age and a Z score >1), and the equivalent 4 dichotomies of SNAPPE-II (with an arbitrary cut-off at 45). Doing so allowed us to evaluate which derivative of SNAP-II conveyed the most discriminating information.

We evaluated the following generalized null hypotheses:

- 1 High SNAP-II or SNAPPE-II scores do not predict ultrasound lesions of the brain;
- 2 High SNAP-II or SNAPPE-II scores do not predict neurodevelopmental dysfunctions;
- 3 High SNAP-II or SNAPPE-II scores do not predict a head circumference >2 standard deviations below the expected mean;
- 4 Distributions of SNAP-II and SNAPPE-II do not vary appreciably among children with and without ultrasound lesions of the brain.
- 5 Distributions of SNAP-II and SNAPPE-II do not vary appreciably among children with and without each neurodevelopmental dysfunction at approximately 24-month post-term equivalent.
- 6 Distributions of SNAP-II and SNAPPE-II do not vary appreciably among children with and without a head circumference >2 standard deviations below the expected mean at approximately 24-month post-term equivalent.

We evaluated the first 3 hypotheses with logistic regression models that adjusted for gestational age by both week of gestation and groups of weeks [19–23]. Each adjustment provided almost identical results. We prefer the groups of weeks because it mini-

mizes the degrees of freedom. To account for the possibility that infants born at a particular hospital are more like each other than like infants born at other hospitals, a hospital cluster term was included in all models [24].

## **Results**

Of the 1,149 infants eligible for the 24-month follow-up, 91 did not have any examination performed. They differed significantly from the 1,058 who had at least one 24-month examination with regard to sociodemographic variables. Mothers of children not seen were less likely to have a college education, more likely to be single and more likely to have public insurance. The children did not differ, however, on their SNAP-II or SNAPPE-II scores nor in their rates of ultrasound abnormalities.

#### *Inter-Relationships of SNAP-II and SNAPPE-II Cut-Offs*

Since all SNAP-II items are included in the SNAPPE-II score, it is not surprising that the 4 dichotomies of SNAP-II and SNAPPE-II are highly related to one another (table 2). For example, all children with a SNAP-II Z score >1 had a SNAP-II >30 and a SNAP-II in the highest quartile for gestational age, while 90% had a SNAP-II in the highest decile for gestational age. Similar agreements were seen for SNAPPE-II. All 106 children who had a SNAPPE-II Z score >1 had a SNAPPE-II >45, a SNAPPE-II in the highest quartile for gestational age and a SNAPPE-II in the highest decile for gestational age. As would be expected, 100% of infants whose SNAP-II was in the highest decile for gestational age also had a SNAP-II in the highest quartile for gestational age. The same was seen for SNAPPE-II.

On the other hand, the 4 dichotomies of SNAP-II and SNAPPE-II are not entirely concordant. For example, of the 377 children who had a SNAP-II >30, only 73% were in the top quartile of SNAP-II for each week of gestation, and only 25% had a SNAPPE-II Z score >1. Similarly, only 73% of the 450 infants whose SNAPPE-II was higher than 45 had a SNAP-II >30.

#### *Ultrasound Lesions*

At the younger gestational ages, infants who had each of the 4 ultrasound lesions tended to have higher SNAP scores than infants who had none of the lesions (fig. 1). A SNAP-II  $\geq 30$  predicted intraventricular hemorrhage, ventriculomegaly and an echodense lesion, but did not predict an echolucent lesion with statistical significance (table 3). By and large, a SNAP-II in the high-

**Table 2.** SNAP-II and SNAPPE-II characteristics of children in the study

SNAP-II classification	SNAP-II				SNAPPE-II				
	>30	top quartile	top decile	Z >1	>45	top quartile	top decile	Z >1	
<i>SNAP-II</i>									
Score >30	Y	73	39	40	87	61	31	25	
	N	9	0	0	12	11	3	2	
Highest quartile for GA	Y	77		41	42	68	73	36	29
	N	10		0	0	21	8	2	1
Highest decile for GA	Y	100	100		91	83	94	73	60
	N	19	18		1	27	17	3	2
Z score >1	Y	100	100	90		81	93	70	60
	N	19	18	1		27	17	4	2
<i>SNAPPE-II</i>									
Score >45	Y	73	55	28	27		62	31	26
	N	6	13	3	3		7	1	0
Highest quartile for GA	Y	68	78	41	41	81		44	34
	N	15	10	1	1	17		0	0
Highest decile for GA	Y	78	86	73	70	93	100		77
	N	22	19	3	4	26	16		0
Z score >1	Y	81	92	77	79	100	100	100	
	N	23	21	5	5	27	18	3	
Total number of children		377	360	140	143	450	338	140	106
All values represent percentages except where otherwise stated.									

**Table 3.** Point estimate (and 95% confidence intervals) of the odds ratio of each ultrasound lesion associated with each of the 3 different measures of high SNAP-II and SNAPPE-II

SNAP classification	Cranial ultrasound lesion			
	IVH	VM	ED	EL
<i>SNAP-II</i>				
Score >30	2.0 (1.5, 2.7)	2.0 (1.4, 2.9)	1.5 (1.05, 2.1)	1.2 (0.8, 1.9)
Highest quartile for GA	1.8 (1.4, 2.4)	1.8 (1.2, 2.5)	1.2 (0.9, 1.7)	1.1 (0.7, 1.7)
Highest decile for GA	2.5 (1.7, 3.7)	2.3 (1.5, 3.6)	1.8 (1.1, 2.7)	1.6 (0.9, 2.8)
Z score >1	2.1 (1.4, 3.1)	2.3 (1.4, 3.5)	1.5 (0.95, 2.3)	1.8 (1.1, 3.1)
<i>SNAPPE-II</i>				
Score >45	1.5 (1.2, 2.1)	1.8 (1.2, 2.6)	1.3 (0.9, 1.8)	1.0 (0.6, 1.6)
Highest quartile for GA	1.6 (1.2, 2.1)	1.6 (1.1, 2.2)	1.4 (0.98, 1.9)	1.1 (0.7, 1.7)
Highest decile for GA	1.5 (0.99, 2.2)	1.5 (0.9, 2.5)	1.4 (0.9, 2.2)	1.4 (0.8, 2.5)
Z score >1	1.5 (0.98, 2.4)	1.5 (0.8, 2.6)	1.3 (0.8, 2.1)	1.3 (0.7, 2.5)

The referent group for each set of analyses consists of all newborns whose SNAP-II or SNAPPE-II was low. All models are adjusted for gestational age (23–24, 25–26 and 27 weeks) and include a ‘group’ term for hospital. GA = Gestational age; IVH = intraventricular hemorrhage; VM = ventriculomegaly; ED = echodense lesion; EL = echolucent lesion.

**Table 4.** Point estimate (and 95% confidence intervals) of the odds ratio of each CP diagnosis associated with each of the 3 different measures of high SNAP-II and the 3 different measures of high SNAPPE-II

SNAP classification	CP			GMFCS $\geq 2$
	Quadripareisis	Diparesis	Hemiparesis	
<i>SNAP-II</i>				
Score $>30$	1.1 (0.6, 2.0)	1.6 (0.7, 3.5)	1.4 (0.5, 3.9)	0.9 (0.5, 1.7)
Highest quartile for GA	1.3 (0.7, 2.4)	1.0 (0.5, 2.3)	0.9 (0.3, 2.7)	1.3 (0.7, 2.4)
Highest decile for GA	1.7 (0.8, 3.8)	0.6 (0.1, 2.8)	2.3 (0.6, 8.5)	1.7 (0.8, 3.9)
Z score $>1$	1.6 (0.7, 3.5)	– <sup>a</sup>	2.4 (0.7, 8.5)	1.7 (0.8, 3.8)
<i>SNAPPE-II</i>				
Score $>45$	1.5 (0.8, 2.7)	1.6 (0.7, 3.7)	2.7 (0.98, 7.4)	1.8 (0.96, 3.3)
Highest quartile for GA	1.4 (0.8, 2.6)	0.7 (0.3, 1.7)	1.5 (0.5, 4.2)	1.6 (0.8, 3.)
Highest decile for GA	1.5 (0.6, 3.5)	0.3 (0.03, 2.0)	2.6 (0.7, 9.7)	2.1 (0.9, 4.7)
Z score $>1$	0.9 (0.3, 3.0)	– <sup>a</sup>	2.6 (0.6, 12)	1.0 (0.3, 3.4)

The referent group for each set of analyses consists of all newborns whose SNAP-II or SNAPPE-II was lower. All models are adjusted for gestational age (GA; 23–24, 25–26 and 27 weeks) and include a ‘group’ term for hospital.

<sup>a</sup> Zero cell (no diplegias with Z score  $<-1$ ).

**Table 5.** Point estimate (and 95% confidence intervals) of the odds ratio of each category of a low Bayley Scale associated with each of the 3 different measures of high SNAP-II and the 3 different measures of high SNAPPE-II

SNAP classification	MDI		PDI	
	$<55$	55–69	$<55$	55–69
<i>SNAP-II</i>				
Score $>30$	1.4 (0.9, 2.2)	1.4 (0.8, 2.2)	1.7 (1.1, 2.5)	1.6 (1.01, 2.5)
Highest quartile for GA	1.3 (0.9, 2.0)	1.5 (0.96, 2.4)	1.9 (1.3, 2.8)	1.4 (0.9, 2.1)
Highest decile for GA	2.0 (1.1, 3.5)	1.1 (0.5, 2.2)	1.8 (1.1, 3.2)	1.4 (0.8, 2.7)
Z score $>1$	1.7 (0.96, 3.1)	1.1 (0.5, 2.2)	1.7 (0.98, 3.0)	1.2 (0.6, 2.2)
<i>SNAPPE-II</i>				
Score $>45$	2.1 (1.4, 3.3)	1.7 (1.02, 2.7)	2.1 (1.4, 3.1)	2.2 (1.4, 3.4)
Highest quartile for GA	1.8 (1.2, 2.8)	1.4 (0.9, 2.3)	1.8 (1.2, 2.7)	1.6 (1.01, 2.5)
Highest decile for GA	1.6 (0.9, 2.9)	1.2 (0.6, 2.4)	1.9 (1.03, 3.4)	1.8 (0.98, 3.4)
Z score $>1$	1.4 (0.7, 3.1)	1.5 (0.7, 3.3)	1.4 (0.7, 3.0)	2.0 (1.00, 4.0)

The referent group for each set of analyses consists of all newborns whose SNAP-II or SNAPPE-II was lower. All models are adjusted for gestational age (GA; 23–24, 25–26 and 27 weeks) and include a ‘group’ term for hospital.

est decile for gestational age at birth did a better job than a SNAP-II  $\geq 30$  and a SNAP-II in the highest quartile for gestational age at birth predicting the occurrence of all 4 lesions. A SNAP-II Z score  $>1$  did as well as a SNAP-II in the highest decile for gestational age at birth. SNAPPE-II also provided discriminating infor-

mation, but none of the 4 cut-offs did as well as their SNAP-II equivalents, with the exception of a SNAPPE-II in the highest quartile for gestational age predicting echodense lesions slightly better than its SNAP-II equivalent. Intraventricular hemorrhage and ventriculomegaly tended to be slightly better predicted by high illness



**Table 6.** Point estimate (and 95% confidence intervals) of the odds ratio of screening positive for autism spectrum disorder on the M-CHAT associated with each of the 3 different measures of high SNAP-II and the 3 different measures of high SNAPPE-II

SNAP classification	M-CHAT+	
	total sample	restricted sample <sup>1</sup>
<i>SNAP-II</i>		
Score >30	1.4 (0.9, 2.0)	1.4 (0.9, 2.2)
Highest quartile for GA	1.7 (1.2, 2.4)	1.7 (1.2, 2.6)
Highest decile for GA	1.4 (0.8, 2.3)	1.3 (0.7, 2.4)
Z score >1	1.5 (0.9, 2.4)	1.4 (0.8, 2.5)
<i>SNAPPE-II</i>		
Score >45	1.6 (1.1, 2.3)	1.4 (0.9, 2.2)
Highest quartile for GA	1.8 (1.3, 2.5)	1.6 (1.1, 2.4)
Highest decile for GA	1.4 (0.8, 2.4)	1.0 (0.5, 1.9)
Z score >1	1.3 (0.7, 2.3)	1.0 (0.5, 2.2)

The referent group for each set of analyses consists of all newborns whose SNAP-II or SNAPPE-II was lower. All models are adjusted for gestational age (GA; 23–24, 25–26 and 27 weeks) and include a ‘group’ term for hospital.

<sup>1</sup> Excludes children who could not walk independently, were blind in one eye, used a hearing aid or received special services for the hearing impaired.

severity scores than were echodense and echolucent parenchymal lesions.

#### *CP Diagnoses*

Children who were given a CP diagnosis at a 2-year post-term equivalent had SNAP scores on postnatal day 1 that did not differ appreciably from the scores of infants not given a CP diagnosis (fig. 1). The approximately doubling or even tripling of risk of hemiparetic CP seen with some SNAP-II and SNAPPE-II cut-offs did not achieve statistical significance. Neither was any of the SNAP-II cut-offs, nor SNAPPE-II cut-offs, significantly associated with an increased risk of any of the other 2 CP diagnoses (table 4) or a GMFCS  $\geq 2$ , an indicator that the child is not able to walk even when his/her hand is held (data not shown). Nevertheless, the number of relative risk estimates in table 4 with an effect size  $>2.0$  and a confidence interval that almost excludes 1.0 deserves recognition.

#### *Bayley Scales of Infant Development*

At younger gestational ages, infants who later had low Bayley Scales tended to have higher SNAP scores on postnatal day 1 than did infants whose Bayley was above 70

(fig. 1). Only a SNAP-II in the highest decile for gestational age was associated with significantly increased risk of a MDI  $<55$ , while almost all 4 SNAP-II cut-offs predicted a PDI  $<55$  (table 5). None of the SNAP-II dichotomies significantly predicted an MDI in the 55–69 range (between 2 and 3 standard deviations below the expected mean) and only a SNAP-II  $>30$  was associated with a significantly increased risk for a PDI in this range. The 4 SNAPPE-II dichotomies predicted low MDI and PDI scores better when their definition did not invoke an adjustment for gestational age.

#### *M-CHAT*

The odds that a child would be screened as positive for an autism spectrum disorder on the M-CHAT was increased if she had a SNAP-II or SNAPPE-II in the highest quartile for gestational age at birth or a SNAPPE-II  $>45$  (table 6). Surprisingly, the odds of screening positive on the M-CHAT were lower when the children were divided at the highest decile for gestational age of both SNAP-II and SNAPPE-II than at the highest quartile.

Motor, as well as vision and hearing, limitations can result in an infant inappropriately classified as at risk for autism spectrum disorder [25]. To avoid such confusion, we restricted 1 set of analyses to children who had no difficulty walking, were not classified as blind in either eye, did not have a hearing aid and were not receiving special services for the hearing impaired. In this sample with 157 fewer children, the point estimates of the odds ratios were somewhat smaller, with slightly wider confidence intervals. Still, a SNAP-II or SNAPPE-II in the highest quartile for gestational age at birth significantly predicted a positive autism screen.

#### *Small Head Circumference at 24 Months*

Although a head circumference Z score of  $<-2$  defines the lowest 2.5% in a normal sample, 10% of children in the ELGAN sample had such an extremely small head circumference at 24 months. At all gestational ages, these children had higher SNAP scores than children whose 24-month head circumference was in the normal range (fig. 1).

A head circumference Z score of  $\geq -2$  and  $<-1$  normally defines the next lowest 14%, but in our sample it included 18%. The distribution of their SNAP-II scores in this group did not differ from those of children whose 24-month head circumference was larger (fig. 1).

A SNAP-II in the highest quartile for gestational age at birth, a SNAPPE-II  $>45$ , a SNAPPE-II in the highest quartile or decile for gestational age and a SNAPPE-II Z score  $>1$  predicted appreciably increased odds of a 24-

**Table 7.** Point estimate (and 95% confidence intervals) of the odds ratio of having a 24-month head circumference Z score <-2 associated with each of the 4 different measures of high SNAP-II and the 4 different measures of high SNAPPE-II

SNAP-II classification	Adjusted for microcephaly at birth	
	no	yes
<i>SNAP-II</i>		
Score >30	1.5 (0.9, 2.4)	1.5 (0.9, 2.4)
Highest quartile for GA	1.8 (1.2, 2.9)	1.8 (1.1, 2.9)
Highest decile for GA	1.9 (0.99, 3.5)	1.9 (0.98, 3.6)
Z score >1	1.8 (0.96, 3.3)	1.7 (0.9, 3.2)
<i>SNAPPE-II</i>		
Score >45	2.1 (1.5, 3.7)	1.8 (1.1, 2.9)
Highest quartile for GA	3.1 (2.0, 4.8)	2.3 (1.4, 3.7)
Highest decile for GA	2.4 (1.3, 4.4)	1.5 (0.8, 3.0)
Z score >1	2.4 (1.2, 4.9)	1.4 (0.7, 3.1)

The referent group for each set of analyses consists of all newborns whose SNAP-II or SNAPPE-II was lower. All models are adjusted for gestational age (GA; 23–24, 25–26 and 27 weeks) and include a ‘group’ term for hospital. Models for the right column are also adjusted for microcephaly at birth (birth head circumference Z score of <-2, ≥-2 but <-1, ≥-1).

month head circumference Z score of <-2 (table 7). Adjusting for a small head circumference at birth did not change these odds ratios appreciably. However, the same adjustment reduced the magnitude of the odds of a 24-month head circumference Z score <-2 associated with all measures of SNAPPE-II, with only 2 retaining nominal statistical significance (>45 and highest quartile for gestational age).

## Discussion

The major finding of our study is that SNAP-II and SNAPPE-II convey information about the risks of some cerebral ultrasound lesions, low Bayley scores (both MDI and PDI), screening positive for an autism spectrum disorder and a small head circumference at the 24-month post-term equivalent, but not CP. This risk information conveyed by an early illness severity score supplements that conveyed by gestational age.

### *What Is the Nature of the Information Conveyed by SNAP Scores?*

We consider 2 possibilities. First, SNAP scores might be in the causal chain of events leading to brain damage.

SNAP-II and SNAPPE-II are indicators of physiologic instability and illness severity in the first 12 postnatal hours. Some might therefore view the association between SNAP and brain lesions, dysfunctions and a small head circumference as indicators of the physiologic instability that reduces blood flow to the brain, which leads to energy failure, which, in turn, results in brain damage. If this view is correct, then the SNAP scores are in the causal chain. To date, however, convincing epidemiologic evidence has not been found to support the view that hypotension and related physiologic disturbances increase the risk of cerebral white matter damage [26, 27].

Alternatively, SNAP scores might be surrogates for items in the causal chain. SNAP scores decline with increasing gestational age at birth (fig. 1). Could it be that the SNAP scores describe physiological characteristics of immaturity better than our temporal measure of gestational age at birth does? Although we feel that the SNAP-II and SNAPPE-II scores reflect immaturity, a component of these scores should be seen as reflecting medical and nursing care characteristics (e.g., efforts to maintain temperature and blood pressure).

Consider that all babies at 25 weeks are not equivalent in their vulnerability to adversities. Some are more developmentally advanced than others despite their similar gestational age. Gestational age at birth treats all 25-weekers as identical [19]. SNAP scores can provide additional information about physiologic maturation, functioning as a marker for many processes that are developmentally regulated, including correlates of brain maturation and vulnerability, and the ability to synthesize growth factors and other proteins that protect. Thus, SNAP scores might contribute unique information about maturity that supplements gestational age at birth. In this interpretation, SNAP scores are not in the causal chain. Rather, SNAP is a marker for other aspects of developmental vulnerability.

### *Continuum of Casualty*

The concept of ‘the continuum of reproductive casualty’ postulates that what leads to death can also lead to less severe adversities [20, 21]. The forms of brain damage that are the focus of this paper might be viewed as located somewhere along the spectrum between death and the mildest of dysfunctions. Our finding that a predictor of death also predicts brain damage and sequelae can be viewed as support for ‘the continuum of fetal casualty’. However, we want to minimize any misunderstanding our findings might provoke. A too ready acceptance of this concept has the potential to minimize the search for

alternative explanations. Yes, death and brain damage can share some antecedents. On the other hand, what causes death is often qualitatively, not merely quantitatively different from what causes brain damage [22].

#### *Why Do the Highest Deciles Not Have Higher Odds Ratios than the Highest Quartiles?*

The higher the SNAP, the greater the physiologic instability. Thus, if this instability contributes to the outcome, then the highest decile of SNAP-II for gestational age at birth should have higher risk ratios than the highest quartile for gestational age at birth. This is not what we saw consistently. One possible explanation for this phenomenon assumes that some of the SNAP-II scores are misclassified. Another explanation assumes that not all the risk information is carried by extreme values, but also by values below. A 3rd explanation is that undetermined confounders or effect modifiers might distort our estimates of the associations in question. Such an explanation may account for the inability of either score to predict the occurrence of CP.

#### *Why Does SNAPPE-II Sometimes Appear to Be a Slightly Better Predictor of Neurodevelopmental Dysfunctions/Characteristics than SNAP-II?*

SNAPPE-II, unlike SNAP-II, has points for low birth weight and fetal growth restriction [10]. Fetal growth restriction increases the risk of congenital microcephaly [23] and postnatal microcephaly [28]. In addition, congenital microcephaly increases the risk of postnatal microcephaly [29]. Consequently, fetal growth restriction and its correlates might explain some of the improved prediction for microcephaly seen with SNAPPE-II. Support for this view is found in the observation that the SNAPPE-II odds ratios fell more than those of SNAP-II when adjustment was made for the head circumference Z score at birth, a fetal growth restriction correlate. It could be argued that SNAPPE-II is less sensitive to physiologic instability because it incorporates growth measures.

#### *SNAP-II Z Scores*

The VON generously provided the means, medians and standard deviations of SNAP-II and SNAPPE-II in the VON data for each week of gestation of the subjects in our study. This external standard allowed us to generalize our findings.

Although similar to our data, the VON data have important differences. If the sample is normally distributed, an internal Z score of  $>1$  should identify about 16% of the sample. The VON SNAP-II Z score  $>1$  identified 10% of

the subjects in our study, while a VON SNAPPE-II Z score  $>1$  identified only 7% of our subjects. This finding suggests that our cohort was slightly less sick and physiologically unstable than the overall VON data. At least in part, this may be attributable to the fact that VON includes outborn infants, while all ELGAN Study subjects were inborn.

#### *Differences with Different Dichotomies of SNAP-II and SNAPPE-II*

We have evaluated 4 different cut-offs for SNAP-II and 4 different cut-offs for SNAPPE-II. We did so for 2 reasons. First, no matter how well guided, an arbitrary cut-off should be seen as just that – arbitrary. Second, a cut-off at 1 gestational age might not be as appropriate at another gestational age. In our sample, the median SNAP-II at 27 weeks is about half the median SNAP-II at 23 weeks. Consequently, a ‘one size’ SNAP-II does not fit all. This perception led us to allow the data to tell us what defines the upper quartile and decile at each gestational age.

Our study has multiple important strengths. First, we included a large number of infants, making it unlikely that we have missed important associations due to lack of statistical power. Second, the rather narrow gestational age range reduces the likelihood that our results are confounded by characteristics that are closely related to gestational age. Third, we used multiple cut-offs for a ‘high score’ including 1 based on an external standard to explore the relative sensitivity of multiple derivatives of SNAP-II and SNAPPE-II. Fourth, the prospective character of our project, the use of data collection methods tested for reliability [12, 16, 17] and the use of multivariable data analysis to adjust for potential confounders maximize the validity of our observations. Fifth, outcomes were assessed by examiners who were not aware of the medical histories of the infants. Sixth, attrition in the first 2 years was modest, with information about the 2-year assessment available from 83% (for head circumference) to 91% (for M-CHAT) of all infants.

The major weakness of our study is a weakness of all observational studies, i.e. the inability to distinguish between causation and association as explanations for what we found.

We conclude that among the most immature newborns, SNAP-II and SNAPPE-II carry information about the risks of cranial ultrasound lesions when the infant is in the ICU and neurodevelopmental disabilities when the child is older. This information appears to supplement

information conveyed by gestational age, but we do not know if it reflects the effects of physiologic disturbances in the causal chain to brain damage or reflects its correlation with brain vulnerability. Neonatal severity of illness scores should be further evaluated in cohorts of preterm infants. However, we urge that SNAP-II and SNAPPE-II scores not be used to predict adverse neurologic outcomes in individual patients.

## Acknowledgements

This study was supported by a cooperative agreement with the National Institute of Neurological Diseases and Stroke (5U01NS040069-05), a program project grant from the National Institute of Child Health and Human Development (NIH-P30-HD-18655), the Susan Saltonstall Fund and the Richard Saltonstall Charitable Foundation. The authors gratefully acknowledge the contributions of our subjects and their families, as well as those of our colleagues.

## References

- Schusterschitz N, Joannidis M: Predictive capacity of severity scoring systems in the ICU. *Contrib Nephrol* 2007;156:92-100.
- Afessa B, Gajic O, Keegan MT: Severity of illness and organ failure assessment in adult intensive care units. *Crit Care Clin* 2007;23:639-658.
- Pollack MM, Koch MA, Bartel DA, Rapoport I, Dhanireddy R, El-Mohandes AA, Harkavy K, Subramanian KN: A comparison of neonatal mortality risk prediction models in very low birth weight infants. *Pediatrics* 2000;105:1051-1057.
- Zupancic JA, Richardson DK, Horbar JD, Carpenter JH, Lee SK, Escobar GJ: Revalidation of the score for neonatal acute physiology in the Vermont Oxford Network. *Pediatrics* 2007;119:e156-e163.
- Kling PJ, Sullivan TM, Leftwich ME, Roe DJ: Score for neonatal acute physiology and phlebotomy blood loss predict erythrocyte transfusions in premature infants. *Arch Pediatr Adolesc Med* 1997;151:27-31.
- Escobar GJ, Shaheen SM, Breed EM, Botas C, Greene JD, Yoshida CK, Zupancic J, Newman TB: Richardson score predicts short-term adverse respiratory outcomes in newborns  $\geq 34$  weeks gestation. *J Pediatr* 2004;145:754-760.
- Hagadorn JI, Richardson DK, Schmid CH, Cole CH: Cumulative illness severity and progression from moderate to severe retinopathy of prematurity. *J Perinatol* 2007;27:502-509.
- Richardson DK, Phibbs CS, Gray JE, McCormick MC, Workman-Daniels K, Goldmann DA: Birth weight and illness severity: independent predictors of neonatal mortality. *Pediatrics* 1993;91:969-975.
- Mattia FR, deRegnier RA: Chronic physiologic instability is associated with neurodevelopmental morbidity at one and two years in extremely premature infants. *Pediatrics* 1998;102:E35.
- Richardson DK, Corcoran JD, Escobar GJ, Lee SK: SNAP-II, SNAPPE-II: simplified newborn illness severity and mortality risk scores. *J Pediatr* 2001;138:92-100.
- Teele R, Share J: *Ultrasonography of Infants and Children*. Philadelphia, Saunders, 1991.
- Kuban K, Adler I, Allred EN, Batton D, Bezinque S, Betz BW, Cavenagh E, Durfee S, Ecklund K, Feinstein K, Fordham LA, Hampf F, Junewick J, Lorenzo R, McCauley R, Miller C, Seibert J, Specter B, Wellman J, Westra S, Leviton A: Observer variability assessing US scans of the preterm brain: the ELGAN study. *Pediatr Radiol* 2007;37:1201-1208.
- Bayley N: Bayley Scales Of Infant Development-II. San Antonio, Psychological Corp., 1993.
- Robins DL, Fein D, Barton ML, Green JA: The Modified Checklist for Autism in Toddlers: an initial study investigating the early detection of autism and pervasive developmental disorders. *J Autism Dev Disord* 2001;31:131-144.
- CDC: CDC Growth Charts: United States 2007. Atlanta, CDC, 2007.
- Kuban KC, O'Shea M, Allred E, Leviton A, Gilmore H, DuPlessis A, Krishnamoorthy K, Hahn C, Soul J, O'Connor SE, Miller K, Church PT, Keller C, Bream R, Adair R, Miller A, Romano E, Bassan H, Kerkering K, Engelke S, Marshall D, Milowic K, Wereszczak J, Hubbard C, Washburn L, Dillard R, Heller C, Burdo-Hartman W, Fagerman L, Sutton D, Karna P, Olomu N, Caldarelli L, Oca M, Lohr K, Scheiner A: Video and CD-ROM as a training tool for performing neurologic examinations of 1-year-old children in a multicenter epidemiologic study. *J Child Neurol* 2005;20:829-831.
- Kuban KC, Allred EN, O'Shea TM, Paneth N, Pagano M, Leviton A, ELGAN Study Cerebral Palsy-Algorithm Group: An algorithm for diagnosing and classifying cerebral palsy in young children. *J Pediatr* 2008;153:451-452.
- Palisano RJ, Hanna SE, Rosenbaum PL, Russell DJ, Walter SD, Wood EP, Raina PS, Galuppi BE: Validation of a model of gross motor function for children with cerebral palsy. *Phys Ther* 2000;80:974-985.
- Leviton A, Blair E, Dammann O, Allred E: The wealth of information conveyed by gestational age. *J Pediatr* 2005;146:123-127.
- Lilienfeld AM, Parkhurst E: A study of the association of factors of pregnancy and parturition with the development of cerebral palsy; a preliminary report. *Am J Hyg* 1951;53:262-282.
- Pasamanick B, Lilienfeld AM: Association of maternal and fetal factors with development of mental deficiency. I. Abnormalities in the prenatal and perinatal periods. *J Am Med Assoc* 1955;159:155-160.
- Ounsted M: Causes, continua and other concepts. I - the 'continuum of reproductive casualty'. *Paediatr Perinat Epidemiol* 1987;1:4-7.
- Leviton A, Holmes LB, Allred EN, Vargas J: Methodologic issues in epidemiologic studies of congenital microcephaly. *Early Hum Dev* 2002;69:91-105.
- Begg MD, Parides MK: Separation of individual-level and cluster-level covariate effects in regression analysis of correlated data. *Stat Med* 2003;22:2591-2602.
- Kuban KC, O'Shea TM, Allred EN, Tager-Flusberg H, Goldstein DJ, Leviton A: Prevalence and correlates of screening positive on the Modified-Checklist for Autism in Toddlers (M-CHAT) in extremely low gestational age newborns. *J Pediatrics* 2009;154:535-540.
- Dammann O, Allred EN, Kuban KC, Van Marter LJ, Pagano M, Sanoeka U, Leviton A: Systemic hypotension and white-matter damage in preterm infants. *Dev Med Child Neurol* 2002;44:82-90.
- Limperopoulos C, Bassan H, Kalish LA, Ringer SA, Eichenwald EC, Walter G, Moore M, Vanasse M, DiSalvo DN, Soul JS, Volpe JJ, du Plessis AJ: Current definitions of hypotension do not predict abnormal cranial ultrasound findings in preterm infants. *Pediatrics* 2007;120:966-977.
- Amin H, Singhal N, Sauve RS: Impact of intrauterine growth restriction on neurodevelopmental and growth outcomes in very low birthweight infants. *Acta Paediatr* 1997;86:306-314.
- Gross SJ, Kosmetatos N, Grimes CT, Williams ML: Newborn head size and neurological status: predictors of growth and development of low birth weight infants. *Am J Dis Child* 1978;132:753-756.