

# The New England Journal of Medicine

Copyright © 2001 by the Massachusetts Medical Society

VOLUME 344

FEBRUARY 15, 2001

NUMBER 7



## THE CONTINUING VALUE OF THE APGAR SCORE FOR THE ASSESSMENT OF NEWBORN INFANTS

BRIAN M. CASEY, M.D., DONALD D. MCINTIRE, PH.D., AND KENNETH J. LEVENO, M.D.

### ABSTRACT

**Background** The 10-point Apgar score has been used to assess the condition and prognosis of newborn infants throughout the world for almost 50 years. Some investigators have proposed that measurement of pH in umbilical-artery blood is a more objective method of assessing newborn infants.

**Methods** We carried out a retrospective cohort analysis of 151,891 live-born singleton infants without malformations who were delivered at 26 weeks of gestation or later at an inner-city public hospital between January 1988 and December 1998. Paired Apgar scores and umbilical-artery blood pH values were determined for 145,627 infants to assess which test best predicted neonatal death during the first 28 days after birth.

**Results** For 13,399 infants born before term (at 26 to 36 weeks of gestation), the neonatal mortality rate was 315 per 1000 for infants with five-minute Apgar scores of 0 to 3, as compared with 5 per 1000 for infants with five-minute Apgar scores of 7 to 10. For 132,228 infants born at term (37 weeks of gestation or later), the mortality rate was 244 per 1000 for infants with five-minute Apgar scores of 0 to 3, as compared with 0.2 per 1000 for infants with five-minute Apgar scores of 7 to 10. The risk of neonatal death in term infants with five-minute Apgar scores of 0 to 3 (relative risk, 1460; 95 percent confidence interval, 835 to 2555) was eight times the risk in term infants with umbilical-artery blood pH values of 7.0 or less (relative risk, 180; 95 percent confidence interval, 97 to 334).

**Conclusions** The Apgar scoring system remains as relevant for the prediction of neonatal survival today as it was almost 50 years ago. (N Engl J Med 2001; 344:467-71.)

Copyright © 2001 Massachusetts Medical Society.

**I**N 1952, Virginia Apgar proposed the Apgar score as a means of evaluating the physical condition of infants shortly after delivery.<sup>1</sup> This scoring system, which encouraged delivery-room personnel to pay close attention to the newborn, was rapidly adopted in delivery rooms throughout the United States and elsewhere. Indeed, it has been said

that “every baby born in a modern hospital anywhere in the world is looked at first through the eyes of Virginia Apgar.”<sup>2</sup> Each of five easily identifiable characteristics — heart rate, respiratory effort, muscle tone, reflex irritability, and color — is assessed and assigned a value of 0 to 2. The total score is the sum of the five components, and a score of 7 or higher indicates that the baby’s condition is good to excellent. The Apgar score is determined at one and five minutes after delivery and is therefore a rapid way to evaluate the physical condition of newborn infants. Of the two scores, the five-minute score has come to be regarded as the better predictor of survival in infancy.<sup>3</sup>

The value of the Apgar score has become controversial because of attempts to use it as a predictor of the neurologic development of the infant, a use for which it was never intended.<sup>4</sup> For example, the use of the Apgar score to identify birth asphyxia is a misapplication, since conditions such as congenital anomalies, preterm birth, and administration of drugs to the mother can result in low scores that are not reflective of asphyxia.<sup>5</sup> Subsequently, measurement of umbilical-artery blood pH has been widely adopted in the United States as an adjunct to the Apgar score for assessing the condition of newborn infants.<sup>6-8</sup> The Committee on Obstetric Practice of the American College of Obstetricians and Gynecologists, in concert with the American Academy of Pediatrics, has issued a publication entitled “Use and Abuse of the Apgar Score” in an effort to emphasize the limitations of the Apgar system in identifying birth asphyxia and predicting neurologic outcome.<sup>9,10</sup>

We undertook this study to examine whether the original intent of the Apgar system, to predict survival during the neonatal period,<sup>11</sup> remains pertinent almost 50 years after the introduction of the system into American obstetrical and pediatric practice.

From the Department of Obstetrics and Gynecology, University of Texas Southwestern Medical Center, Dallas. Address reprint requests to Dr. Casey at the Department of Obstetrics and Gynecology, University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd., Dallas, TX 75390, or at brian.casey@utsouthwestern.edu.

## METHODS

## Study Design

Selected obstetrical outcomes for all women who give birth at Parkland Hospital, Dallas, as well as the neonatal outcomes, are entered into a computerized data base. Nurses attending each delivery complete an obstetrical data sheet, and research nurses assess the data for consistency and completeness before they are stored electronically. Data on infants' outcomes are abstracted from discharge records. Parkland Hospital is a tax-supported institution serving Dallas County that has a level III neonatal intensive care unit<sup>12</sup> adjacent to the labor and delivery units. The obstetrics service is staffed by house officers and faculty members of the Department of Obstetrics and Gynecology at the University of Texas Southwestern Medical School, and the neonatology service is staffed by house officers and faculty members of the Department of Pediatrics.

Between January 1988 and December 1998, a total of 151,891 women delivered live-born singleton infants with gestational ages of 26 weeks or greater at Parkland Hospital. Infants less than 26 weeks of gestational age were excluded because their intrapartum treatment might have been influenced by concern about viability. Infants with major malformations identified by the time of discharge or at autopsy were also excluded. In all live-born infants, umbilical-artery blood samples were routinely drawn from a doubly clamped segment of the umbilical cord into heparinized 3-ml syringes and placed in ice for transport to the hospital laboratory for the measurement of blood gases and pH. The results were linked to the perinatal data base. Similarly, Apgar scores at one and five minutes were routinely assigned to all live-born infants, either by the nurses attending normal deliveries at term or by third-year pediatric house officers attending high-risk deliveries. The latter included all deliveries at less than 36 weeks of gestation and term deliveries with risk factors such as cesarean or forceps delivery, maternal fever during labor, and meconium-stained amniotic fluid. In addition, a resuscitation team consisting of a specially trained nurse from the neonatal intensive care unit, a respiratory therapist, and a fellow in neonatology was immediately available.

The outcome of interest was neonatal death during the first 28 days after birth. Preterm infants were defined as those born between 26 and 36 weeks of gestation, and term infants were defined as those born at or after 37 weeks of gestation. The gestational age assigned to each infant was based on the obstetrical estimate that was used to manage the care of the mother during the intrapartum period.

## Statistical Analysis

All analyses were performed with SAS software (version 8.0, SAS Institute, Cary, N.C.). The Mantel-Haenszel chi-square test for trend and the chi-square goodness-of-fit test were used for categorical data. Relative risks and 95 percent confidence intervals were calculated by the Mantel-Haenszel method. Continuous data are presented as means  $\pm$  SD. All P values were derived from two-sided tests.

## RESULTS

Paired Apgar scores and umbilical-artery blood gas values (pH, partial pressure of carbon dioxide, partial pressure of oxygen, bicarbonate concentration, and base deficit) were available for 145,627 infants, of whom 13,399 (9 percent) were delivered before term and 132,228 (91 percent) were delivered at term. Of the infants whose outcomes are described in this report, 58 percent (84,122) were Hispanic, 26 percent (38,255) were black, 13 percent (19,174) were white, and 3 percent (4076) were of other racial or ethnic backgrounds. The mean maternal age was  $24 \pm 6$  years; 3 percent (3974) of the women were under the age of 16, and 4 percent (6498) were 35 years of age or

**TABLE 1.** INCIDENCE OF NEONATAL DEATH IN 13,399 SINGLETON INFANTS BORN BEFORE TERM (AT 26 TO 36 WEEKS OF GESTATION) IN RELATION TO APGAR SCORES AT FIVE MINUTES OF AGE.\*

| FIVE-MINUTE APGAR SCORE | NO. OF LIVE BIRTHS | NEONATAL DEATH             | RELATIVE RISK (95% CI) |
|-------------------------|--------------------|----------------------------|------------------------|
|                         |                    | no. (rate per 1000 births) |                        |
| 0-3                     | 92                 | 29 (315)                   | 59 (40-87)             |
| 4-6                     | 556                | 40 (72)                    | 13 (9-20)              |
| 7-10                    | 12,751             | 68 (5)                     | 1                      |

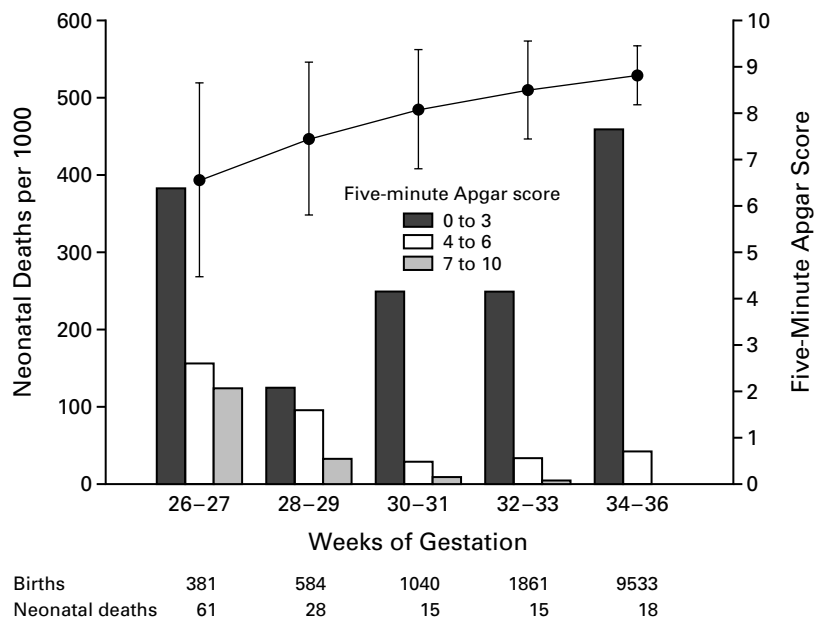
\*Infants with five-minute Apgar scores of 7 to 10 served as the reference group. CI denotes confidence interval.

older. A total of 56,391 women (39 percent) were delivering their first child.

The incidence and relative risk of neonatal death in preterm infants according to their Apgar scores at five minutes are shown in Table 1. The incidence of neonatal death was 315 per 1000 preterm infants with five-minute Apgar scores of 0 to 3 but only 5 per 1000 for those with five-minute Apgar scores of 7 or greater (relative risk, 59; 95 percent confidence interval, 40 to 87). The incidence of neonatal death in preterm infants with five-minute Apgar scores of 4 to 6 was 72 per 1000 (relative risk, 13; 95 percent confidence interval, 9 to 20). The one-minute Apgar score was less useful in predicting neonatal death than the five-minute score (data not shown).

The mean five-minute Apgar score was related to gestational age in preterm infants (Fig. 1). For example, the mean score was  $6.6 \pm 2.1$  in infants born at 26 to 27 weeks of gestation, as compared with  $8.7 \pm 0.8$  in infants born at 34 to 36 weeks. However, the incidence of neonatal death was highest for five-minute Apgar scores of 3 or less, regardless of gestational age. For example, at 26 to 27 weeks of gestation, the neonatal death rate was 385 per 1000 live-born infants with five-minute Apgar scores of 0 to 3, as compared with 147 per 1000 for those with scores of 4 to 6 ( $P=0.002$ ).

Similar analyses were performed for infants delivered at or after 37 weeks of gestation (Table 2). The incidence of neonatal death in term infants with five-minute Apgar scores of 0 to 3 was 244 per 1000, whereas the incidence in term infants with scores of 7 or more was 0.2 per 1000. The mean five-minute Apgar score for infants delivered at 37 to 38 weeks was  $8.9 \pm 0.4$  and was not significantly different for infants born at 39 to 40 weeks and those born at 41 to 42 weeks. The neonatal death rates were related to the five-minute Apgar scores in term infants, regardless of gestational age (Fig. 2). For example, at 39 to 40 weeks, the neonatal death rate was 180 per 1000 in infants with five-minute Apgar scores of 0 to 3, as compared with 12 per 1000 in infants with scores of



**Figure 1.** Mean ( $\pm$ SD) Five-Minute Apgar Scores in Preterm Infants According to Gestational Age (Curve) and Neonatal Death Rates for Infants with Five-Minute Apgar Scores of 0 to 3, 4 to 6, and 7 to 10 (Bars).

At 34 to 36 weeks of gestation, the neonatal death rate was 0.5 per 1000 for Apgar scores of 7 to 10.

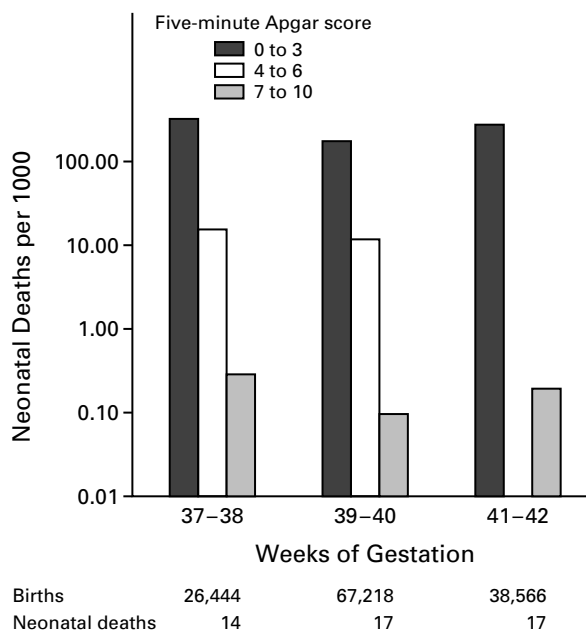
**TABLE 2.** INCIDENCE OF NEONATAL DEATH IN 132,228 SINGLETON INFANTS BORN AT TERM (37 WEEKS OF GESTATION OR LATER) IN RELATION TO APGAR SCORES AT FIVE MINUTES OF AGE.\*

| FIVE-MINUTE APGAR SCORE | NO. OF LIVE BIRTHS | NEONATAL DEATH             | RELATIVE RISK (95% CI) |
|-------------------------|--------------------|----------------------------|------------------------|
|                         |                    | no. (rate per 1000 births) |                        |
| 0-3                     | 86                 | 21 (244)                   | 1460 (835-2555)        |
| 4-6                     | 561                | 5 (9)                      | 53 (20-140)            |
| 7-10                    | 131,581            | 22 (0.2)                   | 1                      |

\*Infants with five-minute Apgar scores of 7 to 10 served as the reference group. CI denotes confidence interval.

4 to 6 ( $P < 0.001$ ) and 0.1 per 1000 in infants with scores of 7 to 10 ( $P < 0.001$ ).

The relative risks of neonatal death are shown in Table 3 for both preterm infants (137 deaths) and term infants (48 deaths) with five-minute Apgar scores of 0 to 3 or umbilical-artery blood pH values of 7.0 or less at birth. These cutoff values were selected for analysis because they are thresholds commonly used to assess the condition of newborn infants.<sup>3</sup> For neither preterm nor term infants did the severity of aci-



**Figure 2.** Neonatal Death Rates in Term Infants with Five-Minute Apgar Scores of 0 to 3, 4 to 6, and 7 to 10, According to Gestational Age.

There were no neonatal deaths among infants born at 41 to 42 weeks of gestation who had Apgar scores of 4 to 6. The scale for neonatal deaths is logarithmic.

**TABLE 3. RELATIVE RISK OF NEONATAL DEATH IN PRETERM AND TERM INFANTS WITH FIVE-MINUTE APGAR SCORES OF 0 TO 3 AND VARIOUS DEGREES OF UMBILICAL-ARTERY BLOOD ACIDEMIA.\***

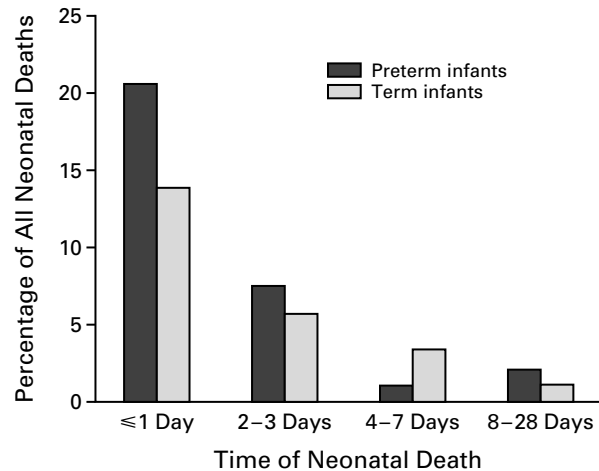
| CHARACTERISTIC                                      | PRETERM INFANTS        | TERM INFANTS     |
|---|------------------------|------------------|
|   | relative risk (95% CI) |                  |
| Five-minute Apgar score of 0–3                      | 59 (40–87)             | 1460 (835–2555)  |
| Umbilical-blood pH                                  |                        |                  |
| ≤7.0  | 20 (11–34)             | 180 (97–334)     |
| ≤6.9  | 22 (11–46)             | 708 (381–1320)   |
| ≤6.8  | 43 (21–91)             | 1407 (736–2689)  |
| Metabolic acidemia†                                 | 38 (17–85)             | 157 (64–385)     |
| Combined five-minute Apgar score of 0–3 and pH ≤7.0 | 102 (65–160)           | 3204 (1864–5508) |

\*Infants with five-minute Apgar scores of 7 to 10 served as the reference group. CI denotes confidence interval.

†Metabolic acidemia was defined by an umbilical-artery blood pH of no more than 7.0 with a partial pressure of carbon dioxide of no more than 76.3 mm Hg, a bicarbonate concentration of no more than 17.7 mmol per liter, and a base excess of no more than –10.3. These values are 2 SD from the means for all infants.

demia appreciably modify the relative risk of neonatal death associated with five-minute Apgar scores of 3 or less. A combination of five-minute Apgar scores of 0 to 3 and pH values of 7.0 or less did, however, increase the risk of neonatal death in both preterm and term infants. This increased risk persisted when infants with five-minute Apgar scores of 0 to 3, pH values of 7.0 or less, or both were compared with infants with five-minute Apgar scores of 7 to 10 and pH values of more than 7.0.

The timing of neonatal death in preterm and term infants with five-minute Apgar scores of 0 to 3 is shown in Figure 3. The poor condition at birth that was reflected in very low five-minute scores (0 to 3) was significantly associated with early neonatal death (at no more than one day of age) in both preterm and term infants ( $P < 0.001$  and  $P = 0.009$  for trend, respectively). The neonatal deaths were grouped according to their apparent cause: those attributed to complications of prematurity (94 infants), sepsis (37 infants), hypoxic–ischemic encephalopathy (24 infants), and unknown causes (30 infants). Hypoxic–ischemic encephalopathy was defined by the presence of seizures or brain death determined electroencephalographically. The relative risk of neonatal death according to the cause was analyzed in infants with five-minute Apgar scores of 0 to 3, as compared with infants with scores of 7 to 10. Apgar scores of 0 to 3 at five minutes were not significantly associated with the cause of death in preterm infants (data not shown). Deaths due to hypoxic–ischemic encephalopathy, however, were significantly more likely in term infants with five-minute scores of 0 to 3 (relative risk, 13; 95 percent confidence interval, 3 to 58).



**Figure 3.** Timing of Neonatal Deaths among Preterm and Term Infants with Five-Minute Apgar Scores of 0 to 3.

## DISCUSSION

In 1952, Apgar reported that neonatal survival through 28 days of age was related to the condition of the infant in the delivery room.<sup>1,3</sup> Our analysis of the relation of five-minute Apgar scores to neonatal survival indicates that the Apgar score is just as meaningful today as it was almost 50 years ago. In both preterm and term infants, survival increased as Apgar scores increased. Among both preterm and term infants, those with five-minute Apgar scores of 0 to 3 had the highest risk of neonatal death. In term infants, the risk of neonatal death was 0.2 per 1000 for those with Apgar scores of 7 to 10, as compared with 244 per 1000 for those with scores of 0 to 3 at birth. This finding parallels the original results reported by Apgar.<sup>1,3</sup> Although preterm infants had low five-minute Apgar scores that reflected gestational age, very low scores (0 to 3) were still associated with an increased risk of neonatal death. Moreover, these very low five-minute Apgar scores were related to the time of neonatal death in both preterm and term infants; specifically, neonatal deaths in infants with low scores occurred soon after birth.

The five-minute Apgar score was a better predictor of neonatal outcome than was measurement of umbilical-artery blood pH, even for newborn infants with severe acidemia. However, the combination of five-minute Apgar scores of 0 to 3 and umbilical-artery blood pH values of 7.0 or less increased the relative risk of death in both preterm and term infants.

The limitations of our analysis include the possibility that differences in the treatment of immature infants might influence the risk of neonatal death. For example, infants with very low Apgar scores who were

delivered at 26 weeks of gestation might have been treated less aggressively than infants delivered at 26 weeks who had higher scores. However, the treatment approach used during the study was to provide all necessary support for infants delivered at 26 weeks of gestation or later.<sup>14</sup> (Infants born before 26 weeks of gestation were not included in the study.)

Another limitation of this study is the small number of infants with very low five-minute Apgar scores. During a period of 11 years, five-minute Apgar scores of 0 to 3 were assigned to only 178 of the 145,627 infants born at our hospital who were included in our study. Nonetheless, our results suggest that very low five-minute scores, although rare, continue to be strongly predictive of early neonatal death. Finally, a potential ascertainment bias is caused by the exclusion of 6264 infants for whom umbilical-artery blood gas results were not available. Indeed, the incidence of neonatal death in this group of infants was 4.5 per 1000, as compared with 1.2 per 1000 ( $P=0.002$ ) in infants for whom blood gas analyses were performed. However, an Apgar score of 0 to 3 at five minutes remained a significant predictor of neonatal death in both preterm and term infants, whether or not umbilical-blood gas analyses were available.

We cannot dispute the contemporary viewpoint that use of the Apgar score for the prediction of long-term neurologic outcome is inappropriate.<sup>4,9,10</sup> However, the poor performance of the Apgar system as a predictor of neurologic development, a purpose for which it was never intended, should not undermine the continuing value of assigning Apgar scores to newborn infants. In our view, it should not be surprising that features of vital activity such as heart rate, respiration,

and neuromuscular function reflect the prognosis for neonatal survival even in premature infants. We conclude that the Apgar system continues to be relevant to the prediction of neonatal outcome after almost half a century.

## REFERENCES

1. Apgar V. Proposal for new method of evaluation of newborn infant. *Anesth Analg* 1953;32:260-7.
2. Skolnick AA. Apgar quartet plays perinatologist's instruments. *JAMA* 1996;276:1939-40. [Erratum, *JAMA* 1997;277:1762.]
3. Drage JS, Kennedy C, Schwarz BK. The Apgar score as an index of neonatal mortality: a report from the Collaborative Study of Cerebral Palsy. *Obstet Gynecol* 1964;24:222-30.
4. Nelson KB, Ellenberg JH. Apgar scores as predictors of chronic neurologic disability. *Pediatrics* 1981;68:36-44.
5. Gilstrap LC III, Hauth JC, Hankins GDV, Beck AW. Second-stage fetal heart rate abnormalities and type of neonatal acidemia. *Obstet Gynecol* 1987;70:191-5.
6. Gilstrap LC III, Leveno KJ, Burris J, Williams ML, Little BB. Diagnosis of birth asphyxia on the basis of fetal pH, Apgar score, and newborn cerebral dysfunction. *Am J Obstet Gynecol* 1989;161:825-30.
7. Winkler CL, Hauth JC, Tucker JM, Owen J, Brumfield CG. Neonatal complications at term as related to the degree of umbilical artery acidemia. *Am J Obstet Gynecol* 1991;164:637-41.
8. Goodwin TM, Belai I, Hernandez P, Durand M, Paul RH. Asphyxial complications in the term newborn with severe umbilical acidemia. *Am J Obstet Gynecol* 1992;167:1506-12.
9. ACOG Committee on Obstetric Practice, American Academy of Pediatrics Committee on Fetus and Newborn. Use and abuse of the Apgar score. In: *Compendium of selected publications*. No. 174. Washington, D.C.: American College of Obstetricians and Gynecologists, July 1996.
10. Use and abuse of the Apgar score. *Pediatrics* 1996;98:141-2.
11. Apgar V, Holaday DA, James LS, Weisbrot IM, Berrien C. Evaluation of the newborn infant — second report. *JAMA* 1958;168:1985-8.
12. Guidelines for perinatal care. 4th ed. Elk Grove Village, Ill.: American Academy of Pediatrics, August 1997:4-6.
13. Apgar V, James LS. Further observations on the newborn scoring system. *Am J Dis Child* 1962;104:419-28.
14. Cunningham FG, MacDonald PC, Gant NF. *Williams obstetrics*. 18th ed. East Norwalk, Conn.: Appleton & Lange, 1989:748.

Copyright © 2001 Massachusetts Medical Society.