Discrepancies in Birth Weights Between Hospital Records and Health Department Data for Low Birth Weight Infants in New York City

ANGELO FERRARA, MD, PhD
YUCEL S. ATAKENT, MD, MS
BARRY LEVINSON, MD

Dr. Ferrara is Professor and Dr. Atakent is Clinical Associate Professor in the Department of Pediatrics, New York University Medical Center. Dr. Levinson is a resident in internal medicine at Parkland Memorial Hospital, Dallas, TX.

This study was supported in part by the National Center for Health Services Research, Office of the Assistant Secretary for Health, under grant HS-03832.

Tear sheet requests to Dr. Angelo Ferrara, Department of Pediatrics, New York University Medical Center, 550 Fifth Ave., New York, NY 10016.

Synopsis

The study is one of the first to compare corresponding birth weights documented on New York City Health Department Vital Statistics (HDVS) birth tapes and the neonatal medical records of the hospital of birth. Only those infants with birth weights of 2,500 grams (g) or less were studied. Analyses were made of the scope, magnitude of error, and direction of the discrepancies observed.

Concordance was considered present if the discrepancy in birth weight was 30 g or less. HDVS birth tapes and the hospital charts of 3,864 neonates were reviewed. The study population came from 48 of 53 hospitals in the metropolitan area. Hospitals were divided into three categories by the level of care offered. Each level of care was subdivided into groups by type of hospital ownership, that is, proprietary, voluntary, and municipal.

Concordance was 87 percent overall and ranged from 67 to 96 percent among the study hospitals. More discrepancies were found for levels II and III hospitals than in level I hospitals, those with less sophisticated resources. Municipal hospitals had more discrepancies in birth weights than voluntary hospitals. Infants who had been transported from the birth facility to another facility had significantly higher concordance rates than the nontransported infants, after adjusting for levels of care, type of ownership of the hospital, and birth weight categories. Increased concordance rates were shown to be associated with increased birth weights.

ACCURATE RECORDS of infants' weights at birth and the continued accuracy of those records in subsequent statistical compilations are essential to effective studies of maternal and child health and perinatal epidemiology in particular.

Vital statistics of birth weight categories are standardized routinely with reference to neonatal and infant morbidity and mortality. Infant mortality statistics for developed nations are heavily influenced by the distributions of birth weights and birth weight-specific mortality (*1*).

Research using data for such cities as Baltimore, New York, Denver, and Portland has shown that neonatal mortality rates decrease with increased birth weights (*2–4*). In addition to its correlation with mortality, birth weight may correlate with morbidity. Those studying trends or making comparative analyses of morbidity or mortality statistics for infants need to reconcile differences in birth weights between and within populations.

Vital statistics are reliable and valid only to the extent of their completeness and accuracy. Ferrara and coworkers collected study population data from source documents to determine the effectiveness of the practice of transporting high risk infants from the birth facility to specialized care centers (*5*). They noted discrepancies between birth weights reported on magnetic tape from the Department of Health of the City of New York, and the birth weights noted on neonatal medical records.

The objective of the study reported here was to examine the scope and magnitude of discrepancies of more than 30 grams (g) in the birth weights of low birth weight infants as recorded on Health Department Vital Statistics (HDVS) birth tapes and

on the corresponding neonatal hospital records in New York City in 1979.

Methodology

Population. The 1979 live birth population of New York City was 106,021 (*6*), 99 percent of whom were born in 53 hospitals. Data were collected from 48 of the hospitals during 1981–82, as many as time allowed. The study population included all low birth weight infants (those weighing 2,500 g or less at birth) born in 1979 in the 48 hospitals. The base population sampled from the 48 hospitals was 8,465 newborns, representing 87 percent of all low birth weight babies born in New York City that year (*7*).

Of the total, a sample of 3,950, or 47 percent, low birth weight infants was selected from the HDVS tapes using a sampling process described subsequently. We eliminated from the study 18 for whom more than 1 birth weight was listed and 68 who were delivered at home. The resulting group included 3,864 neonates, 605 of whom had been transported, and 3,259 who had not. During 1979, 90 percent of the cases of transporting infants in the New York metropolitan area were handled by the New York City Infant Transport Service (ITS) operated by the New York University-Bellevue Medical Centers (*8*).

Sampling process and research design. The sampling procedure was that used in the transport efficacy study of 1980–82 (*5*). HDVS birth tapes listed both transported and nontransported newborns by hospital of birth, birth weight, sex, race (black, white, and other), and 5-minute Apgar score (a numerical expression of the condition of a newborn infant obtained by assessing 5 physiological and neurological indicators at 1 and 5 minutes after birth, assigning scores of 0, 1, or 2 for each indicator; total scores range up to 10 for the most favorable condition). Randomization was achieved by matching each transported neonate by the variables previously listed for a stratified random sample of nontransported neonates.

Chart review was carried out with hospital authorization. A list of the newborns whose records were to be reviewed was taken from the HDVS tapes and provided to the particular hospital. Data were collected from the charts on precoded abstract forms which were keypunched and stored for later retrieval and analysis. Record room abstracting was done by registered nurses who were trained in coding, abstracting, and editing the data.

'In 65 percent of the discrepancies found in this study, the HDVS tapes reflected a lower weight than the medical record. This finding was observed in virtually all ranges of birth weight at each category of magnitude of error.'

Birth weight was defined as the body weight in grams at birth recorded on the neonatal medical records of the hospital.

The hospitals of birth were categorized into three groups according to the level of care provided, as defined by the standards of the voluntary health agencies (*9*) and the New York City Health Department (*10*). The 18 level I hospitals were those with fewer deliveries and less sophisticated resources, which were able to manage only low risk obstetrical and neonatal problems. Their mean number of deliveries (± 1 SD) was 1,521, ± 328. The 17 level II hospitals were those with larger delivery services, able to adequately handle high risk obstetrical patients and stabilize sick neonates. Their mean number of deliveries (± 1 SD) was 2,258, ± 422. The 13 level III hospitals were those having large regional perinatal and neonatal centers able to serve all born-in and transported high risk obstetrical and neonatal patients. Their mean number of deliveries (± 1 SD) was 2,574, ± 689.

Hospitals also were categorized by three categories of ownership: privately owned hospitals operated for profit (proprietary); private, not for profit (voluntary); and operated by the Health and Hospitals Corporation of the City of New York, not for profit (municipal).

HDVS birth tapes were computerized records abstracted from birth certificates registered at the New York City Health Department.

A birth weight discrepancy was a disagreement of more than 30 g in birth weight between the hospital chart and the birth tape.

The correct birth weight, the "gold standard" for the purposes of this study, was that recorded on the neonatal medical records at the hospital of birth.

Data analysis. Statistical analysis of the data included the Yates' corrected chi-square analyses for 2 × 2 contingency tables, the chi-square analysis for a 2 × k contingency table, the multiple com-

Table 1. Distribution of study hospitals with more than 90 percent concordance between paired birth weight records, by level of care provided by the hospital and type of hospital ownership

Level of care provided by type of hospital ownership	Number of hospitals	Hospitals with 90 percent concordance		Percent concordance range among hospitals
		Number	Percent	
Level I:				
Proprietary	3	2	67	89–97
Voluntary	12	5	42	72–95
Municipal	3	1	33	87–91
Level II:				
Voluntary	13	4	31	72–96
Municipal	4	0	0	76–82
Level III:				
Voluntary	10	3	30	67–94
Municipal	3	0	0	75–80
Total	48	15	31	67–96

Table 2. Number and percent distribution of concordances of paired birth weights, by level of care, type of hospital ownership, and birth weight categories in grams

Level of care and type of ownership	500 g		501–1,000 g		1,001–1,500 g		1,501–2,000 g		2,001–2,500 g		More than 2,500[1] g		Total	
	Number	Percent	Number	Percent	Number	Percent	Number	Percent	Number	Percent	Number	Percent	Number	Percent
Level I, total	4	0	79	87	117	89	254	93	886	96	70	12	1410	90
Municipal	1	0	19	79	38	82	75	97	160	95	17	18	310	88
Voluntary	3	0	56	89	73	93	158	92	618	95	47	11	955	90
Proprietary	0	0	4	100	6	83	21	90	108	97	6	0	145	92
Level II, total	32	88	116	81	164	87	268	84	911	92	84	6	1575	85
Municipal	11	73	34	82	52	90	77	88	216	90	31	0	421	80
Voluntary	21	95	82	81	112	85	191	87	695	93	53	10	1154	87
Level III, total	16	81	74	78	146	84	135	81	476	91	32	22	879	85
Municipal	6	67	18	78	42	76	39	82	99	83	13	38	217	78
Voluntary	10	90	56	79	104	90	96	80	377	94	19	11	662	88
Total	52	79	269	82	427	87	657	87	2273	94	186	11	3864	87

[1]Infants listed on the HDVS tapes as weighing 2,500 g or less, but recorded on the neonatal medical records as weighing more than 2,500 g at birth.

parisons for proportions (*11*), the Mantel-Haenszel χ^2 statistic, and the linear regression for proportions (*12*). A slight modification of the Freeman and Tukey arcsine transformation for percentage data (*11, 13*) was used to perform the analysis of multiple comparisons for proportions. Throughout the study, an error with a P of 0.05 or less was considered statistically significant.

Results

HDVS tapes and the neonatal medical records were in agreement within 30 g (or ± 1 oz) of birth weight for 87 percent of the births. The frequency distribution of the study hospitals with more than 90 percent agreement in birth weight registers is shown in table 1. Proprietary hospitals provided only level I care.

Of the 48 hospitals in the study, 15 hospitals, or 31 percent, had at least 90 percent of the recorded paired birth weights within 30 g body weight agreement. Concordance ranged from 67 percent to 96 percent among the 48 study hospitals.

Hospital categories. When concordance rates were examined by the level of care provided, irrespective of ownership, concordance was 90 percent for level I and 85 percent for both level II and level III hospitals ($\chi^2_2 = 21.99$, $P < 0.005$) (table 2). Further analysis showed that the proportion of concordance at level I hospitals was different from either level II ($Q_{\infty,3} = 11.75$, $P < 0.001$) or level III ($Q_{\infty,3} = 9.88$, $P < 0.001$) hospitals at a statistically significant level.

Among level I hospitals, the proportions of concordance were 88 percent for municipal, 90 percent for voluntary, and 92 percent for propri-

etary hospitals. Differences in the percentages of concordant birth weight records in level I hospitals were not statistically significant ($\chi^2_2 = 1.29$, $P > 0.05$).

Analyses were carried out for both municipal and voluntary hospitals at all three levels of care. Among voluntary hospitals, the proportions of paired concordant birth weights were 90 percent for level I, 87 percent for level II, and 88 percent for level III hospitals. For municipal hospitals, the percentages of paired birth weights in accord were 88 percent in level I, 80 percent in level II, and 78 percent in level III hospitals. After controlling for the level of care provided by the hospitals, voluntary hospitals had a statistically significant higher portion of concordant birth weights within 30 g by Mantel-Haenzsel testing ($\chi^2 = 23.32$, $P < 0.001$) compared with the municipally operated hospitals.

The percentages of concordant birth weights among the three levels of care provided by voluntary hospitals were not sufficiently different to be statistically significant ($\chi^2_2 = 4.87$, $P > 0.05$). However, among the municipal hospitals, those providing level I care had a statistically significant higher proportion of birth weight records in accord compared with hospitals either in level II ($Q_{\infty,3} = 8.28$, $P < 0.001$), or level III ($Q_{\infty 3} = 8.56$, $P < 0.001$) categories. Municipal hospitals showed no significant difference in concordance rates between level II and III ($Q_{\infty,3} = 1.66$, $P > 0.20$).

Birth weight categories. The concordance of birth weights was analyzed for all hospitals by birth weight categories. The proportions of concordance were 79 percent for birth weights of < 500 g; 82 percent for 501–1,000 g; 87 percent for 1,001–1,500 g; 87 percent for 1,501–2,000 g; and 94 percent for 2,001–2,500 g (table 2). Differences among these percentages were statistically significant ($\chi^2_4 = 79.45$, $P < 0.001$). The proportions of concordance tended to increase steadily with increased birth weight, but the tendency was not linear (χ^2 slope = 72.95, $P < 0.001$; χ^2_3 linearity = 6.50, $P > 0.05$).

Transport categories. The transported infants in the sample represented 90 percent of all transported neonates. The severity of illness as measured by Apgar scores of those sampled was not significantly different ($P > 0.05$) from those not sampled (table 3). The sample of infants who were not transported was 42 percent of all nontransported neonates.

The percent distribution of Apgar scores for neonates not transported is shown in table 4. The nontransported infants in the sample had a higher percentage of Apgar scores of 6 or less (15 percent) than the nonsampled (8 percent). However, 85 percent of those sampled had scores of 7 or more, similar to the universe population.

Table 3. Percent distribution of 5-minute Apgar scores of 605 transported low birth weight neonates matched with transported group not in sample

Infants transported	*Apgar 3 or less*	*Apgar 4–6*	*Apgar 7–10*
Sampled	6	15	79
Not in sample	6	11	83

Table 4. Percent distribution of 5-minute Apgar scores of 259 nontransported low birth weight neonates matched with nontransported group not in sample

Infants not transported	*Apgar 3 or less*	*Apgar 4–6*	*Apgar 7–10*
Sampled	7	8	85
Not in sample	4	4	92

The concordance rate for transported neonates was 92 percent, compared with 86 percent for the nontransported, a statistically significant difference ($\chi^2 = 15.98$, $P < 0.001$) (table 5).

Those transported had consistently higher concordance rates than the nontransported at all three institutional levels of care. The concordance rates of the transported population were 93 percent for level I, 90 percent for level II, and 86 percent for level III hospitals. For the nontransported population, the respective rates were 88, 84, and 85 percent. These concordance differences by levels of care are notably higher for the transported infants, using Mantel-Haenszel testing ($\chi^2 = 12.77$, $P < 0.001$). The transported infants had a statistically significant higher concordance rate than the nontransported in all birth weight categories from 501 g or more through 2,500 g, using Mantel-Haenszel testing ($\chi^2 = 9.40$, $P < 0.01$).

Linear regressions in proportions were performed for nontransported infants, and showed χ^2 slope = 63.35, $P < 0.001$, with significant deviation from linearity (χ^2_3 linearity = 10.01, $P < 0.025$). For the transported group, there were no significant differences in concordance rates by birth weight categories ($\chi^2_3 = 3.98$, $P < 0.05$).

Considering only the 498 nonconcordant birth weights for all hospitals in the sample, there was a distribution of the magnitude of error as follows: 31–100 g, 36 percent; 101–250 g, 20 percent;

Table 5. Number and percent distribution of concordance of paired birth weights, by hospital level of care and birth weight categories, for transported and nontransported low birth weight infants

Care category	Less than 500 g		501–1,000 g		1,001–1,500 g		1,501–2,000 g		2,001–2,500 g		More than 2,500 g[1]		Totals		Combined total
	T	NT[2]	T	NT	T	NT	T	NT	T	NT	T	NT	T	NT	
Level I:															
Number	0	4	47	32	100	17	162	92	137	749	1	69	447	963	1410
Percent	0	0	92	81	89	88	92	96	96	95	100	10	93	88	90
Level II:															
Number	0	32	27	89	24	140	27	241	45	866	0	84	123	452	1575
Percent	0	88	93	78	88	86	93	84	89	92	0	6	90	84	85
Level III:															
Number	0	16	7	67	6	140	10	125	12	464	0	32	35	844	879
Percent	0	81	100	76	83	86	80	81	83	92	0	22	86	85	85
Total:															
Number	0	52	81	188	130	297	199	458	194	2079	1	185	605	3259	3864
Percent	0	79	93	78	88	87	92	85	94	93	100	10	92	86	87

[1]Infants listed on HDVS tapes as weighing 2,500 g or less at birth, but as more than 2,500 g on neonatal medical records.

[2]*T* is transported infants; *NT* is nontransported infants.

251–500 g, 20 percent; and > 500 g, 24 percent ($\chi^2_3 = 18.94$, $P < 0.005$). Most of the errors were of small magnitude, such as 31–100 g ($\chi^2 = 62.66$, $P < 0.001$).

The 498 discrepancies were categorized by transport status and by magnitude of error (table 6). The distribution of error for the transported group shows a greater incidence in the 31–100 g and the 101–250 g groups of error than the nontransported sample. There were proportionally fewer discrepancies of more than 250 g among the transported than among the nontransported infants ($\chi^2_3 = 9.04$, $P < 0.05$). The majority of discrepancies (323 out of 498, or 65 percent) occurred in the direction of birth weights on HDVS tapes being less than the neonatal medical record ($\chi^2 = 34.74$, $P < 0.001$). Moreover, the trend was apparent in virtually all birth weight categories for each of the magnitude of error groups.

Discussion

The study was one of the first comparisons of birth weights on HDVS tapes and hospital charts in New York City. Depending on the New York City hospital, birth weights are recorded either in grams or pounds and ounces. On HDVS birth tapes, body weights are recorded using metric measurements. To minimize the effects of rounding error that can occur in conversions to metric measurement, only differences in birth weights greater than 30 g between HDVS tapes and the neonatal medical records were considered in this study. Although the impact of discrepancies can be reduced by tabulating data in groups (*14*), rounding errors in conversions appear to involve changes of ± 1 oz. This type of error probably does not limit the usefulness of the data (*15*).

The overall birth weight concordance rate of 87 percent between the HDVS tapes and hospital charts for the study population in New York City is low in comparison with New York State, excluding New York City. For New York State, excluding New York City, the concordance rate within ± 1 oz. of recorded birth weights between birth certificates and hospital charts for 1972 was reported to be 91.6 percent (*16*). However, the data for New York City were abstracted for low birth weight infants, whereas for the State, all birth weight categories were analyzed. In the National Natality Survey (*14*), conducted in 1972, a comparison of birth weights was made between responses to a questionnaire sent to hospitals and mothers, and information on birth certificates. The comparison showed that concordances were lower for low birth weight infants compared with newborns at higher birth weights. The greater proportion of discrepancies among low birth weight infants was not explained.

Most of the discrepancies found in this study occurred for births at the municipally operated hospitals in care levels II and III. Only one municipally operated hospital, which performed relatively better even at level I, had a concordance

rate of more than 90 percent. Moreover, there was a linear trend of low concordance with decreased birth weight.

In 65 percent of the discrepancies found in this study, the HDVS tapes reflected a lower weight than the medical record. This finding was observed in virtually all ranges of birth weight at each category of magnitude of error. If a population survey had been conducted with birth weight as a variable, the analysis of HDVS birth tapes alone would have led to a false conclusion of a higher incidence of neonates with lower birth weights. Such an outcome would be inevitable if the tapes were the source for identifying and linking variables with their respective medical records.

A plausible explanation of the results of our study would be that the HDVS tapes cannot account for the differences in concordance. The explanation assumes that possible errors by the Bureau of Vital Statistics in transcribing and keypunching birthweight records would be randomized and distributed evenly among all categories of hospital levels of care and types of ownership, which is not the case.

Apparently the discrepancies occur at the hospitals of birth. Several sources of error may be postulated, such as errors in transcribing numbers, misreading or reversing digits; different persons handling the data; and errors in conversion between systems of measurement. Presumably, errors are being recorded on hospital charts and birth certificates by health care professionals, such as physicians, nurse midwives, and, possibly, nurses. Considering these factors, why are the records of level I hospitals, those with less sophisticated resources, in better agreement than those at other levels of care? Among possible explanations are the following:

1. Level I hospitals do not have graduate medical education programs. Because deliveries are supervised predominantly by private physicians, who are already trained, fewer errors are likely to be recorded.

2. Level I hospitals have fewer deliveries, and staff members have more time to record data more accurately on charts and birth certificates.

3. Level I hospitals may be expected to have smaller predictable risk populations. Perceived high risk mothers would tend to be delivered at level II or III facilities. Since more errors were made in the very low birth weight categories, less than 1,500 g, than in the heavier birth weight categories, fewer errors would be expected at level I hospitals.

Table 6. Distribution of 498 birth weight discrepancies in sample hospitals in New York City in 1979, by magnitudes of error and infant transport status

Error in grams	*Transported*	*Nontransported*
31–100	18	163
101–250	13	84
251–500	12	88
>500	4	116
Total	47	451

'Several sources of error may be postulated, such as errors in transcribing numbers, misreading or reversing digits; different persons handling the data; and errors in conversion . . .'

Why would very low birth weight infants show at all hospital levels of care a greater rate of discrepancy than the heavier infants? A possible explanation is that the very low birth weight infant is usually quite sick and in need of stabilization, conditions of more importance than accurate weight recording at that time.

But the low proportion of concordance at municipal hospitals needs explanation. The medically indigent patient in the municipal setting is more at risk because of the higher incidence of low birth weight. Hence, the care provider attending the delivery may be more concerned with immediate patient care than with recording birth weights accurately. The provider is likely to be a house officer in training, whose priorities may be directed more toward patient care than record keeping.

An unexpected finding of this study was a distinct difference between the transported (more concordance) and nontransported subpopulations. Several reasons can be postulated for the differences:

1. Nearly 95 percent of all transports were from level I and level II hospitals. Physicians and nurses are aware soon after the birth, perhaps in the delivery room, that the neonate needs to be moved within the first few hours of life. The neonatal transportation process in New York City requires that standard forms first be completed, including maternal consent. The procedures require that birth variables, especially birth weight, be recorded accurately on the transport records.

Summary of study population, New York City, 1979

106,021	Live births
8,465	Base population: low birth weight infants born in 48 hospitals
3,950	Selected from HDVS tapes
3,864	Study sample after eliminating infants delivered at home or with record duplication
3,259	Sample, nontransported infants
605	Sample, transported infants

2. Neonates being transported are usually kept in a special care unit of the referring hospital where more accurate records and record keeping, and presumably better birth weight recording, is required.

These explanations are conjecture, and further investigation is needed to determine the causes of discrepancies between the birth weights recorded on hospital charts and on birth certificates.

References

1. Lee, K. S., Paneth, N., Gartner, L. M., and Pearlman, M.: The very low-birth-weight rate: principal of neonatal mortality in industrialized populations. J Pediatr 97: 759–764 (1980).
2. Sinclair, J. C., and Tudehope, D. I.: Birth weight, gestational age, and neonatal risk. *In* Behrman's neonatal-perinatal medicine, diseases of the fetus and infant, edited by A. A. Fanaroff, R. J. Martin, and I. R. Merkatz, The C. V. Mosby Co., St. Louis, MO, pp. 196–205 (1983).
3. Horwood, S. P., Boyle, M. D., Torrance, G. W., and Sinclair, J. C.: Mortality and morbidity of 500 to 1,499 gram birth weight infants live-born to residents of a defined geographic region before and after neonatal intensive care. Pediatrics 69: 613–620 (1982).
4. Kitchen, W. H., and Guadry, E.: Prediction of survival of very low birth weight infants. Aust Paediatr J. 10: 281–285 (1974).
5. Ferrara, A., et al.: Matched case control population of 2,500 g or less to detect efficacy of neonatal transport; abstract 424. Pediatr Res 16: 149A (1982).
6. Bureau of Health Statistics and Analysis: Summary of vital statistics, 1979. New York City Department of Health, New York, NY, 1980.
7. National Center for Health Services Research: Outcome assessment, EMS Transport Service of neonates of 2,500 g or less; executive summary and final report, grant 5-R18-HS 03832. DHHS, Public Health Service, Washington, DC, 1982.
8. New York City Infant Transport Service. Ninth annual report, Bellevue Hospital, New York University Medical Center, New York, NY, 1979.
9. National Foundation-March of Dimes, Committee on Perinatal Health: Toward improving the outcome of pregnancy: recommendations for the regional development of maternal and perinatal health services. National Foundation-March of Dimes, White Plains, NY, 1977.
10. City Health Information: Survival of low birthweight infants. New York City Department of Health 1: 23, New York, NY, 1982.
11. Zar, J. H.: Biostatistical analysis, Ed. 2. Prentice-Hall, Inc., Englewood Cliffs, NJ, 1984.
12. Fleiss, J. L.: Statistical methods for rates and proportions, Ed. 2. John Wiley and Sons, Inc., New York, NY, 1981.
13. Freeman, M. F., and Tukey, J. W.: Transformations related to the angular and the square root. Ann Math Statist 21: 607–611 (1950).
14. Querec, L. J.: Comparability of reporting between the birth certificate and the national natality survey. Vital Health Stat [2], No. 83, DHEW Publication (PHS) 80–1357. Public Health Service, National Center for Health Statistics, Washington, DC, 1980.
15. David, R. J.: The quality and completeness of birth weight and gestational age data in computerized birth files. Am J Public Health 70: 964–973 (1980).
16. Carucci, P. M.: Reliability of statistical and medical information reported on birth and death certificates. New York State Department of Health, Monograph 15, Albany, NY, 1979.