A sample of over 20,000 twin sets was collected and twinship data were developed from a unit record processing system. Using the critical variables of sex, survivorship, and certain gross birth anomalies, the authors discuss the development of information from records and its analysis.

STATISTICAL STUDY OF TWINS—METHODS

Alan Gittelsohn, Ph.D., and Samuel Milham, M.D., M.P.H.

THE PRESENT inquiry into the incidence of twinning is an outgrowth of a general investigation of live births and fetal deaths recorded in Upstate New York during the 1950 decade.[1] Through a routine procedure in the department, live birth records are currently updated with matching infant mortality data and, hence, the combined birth-death information is available for analysis. We are thus dealing with a well defined subset comprising about 1 per cent of all maternities or 2 per cent of all births. Except on a selective basis, no attempt has been made to link maternities together into sibships simply because the vital record mechanism has not been designed for such a purpose. Over the decade, the same mother may have had as many as ten births recorded. We recognize that certain types of dependency may thereby be overlooked and attempt to bear the point in mind when drawing inferences. Specifically, the increased familial incidence of dizygotic twinning has been the subject of commentaries since Weinberg's time.[2] The familial aggregation of anencephalus and spina bifida cases has been noted by one of us.[3] Certainly, all traits with high heritability or with high environmental contagion will be clustered in families. In short, the vital registration system, as presently constituted, does not easily lend itself to the development of the family grouping which, in itself, may be of greater biological significance than all the demographic factors, such as maternal age, parity, sex, race, and so forth, by which vital statistics reports are traditionally analyzed.

Taken as isolated events or in sibships, the phenomena of multiple births have long occupied the attention of biologists and medical workers within at least two major contexts. On the one hand, the two distinct twinning processes of polyovulation and of polyembryony provide a naturally occurring mechanism which simulates controlled experimentation. To this end, the twin method has been applied to the problem of evaluating the relative roles of heredity and environment as etiological factors in disease. On the other hand, multiple birth presents a series of medical hazards which, relative to single birth, are manifested in decreased viability of offspring and in increased frequency of complications of pregnancy, labor, and delivery. For the moment, we shall be concerned with the problems of developing twin information through vital records which may have a bearing within either context.

Twin Classification

The initial problem is one of definition since the twinning condition pre-

sents a wide spectrum. The majority of twin pregnancies result in two fetuses of similar size at delivery. Rarely, one fetus may die during the course of development while its mate develops normally. Under this circumstance, great differences in configuration may be seen at delivery. The one normal member may be contrasted with a degenerated embryo, a partial fetus or a fetus papyraceous, depending on the stage at which intrauterine death occurred. Conjoined twins further complicate the picture. These may be two fully developed individuals with only a superficial connection, or one member may be a small mass of tissue attached to or included in a fairly normal mate. A somewhat arbitrary convention must be adopted when classifying the plurality of parasitic twins or of twins joined at the chest with two heads, two arms, and two legs (dicephalus dipus dibrachius). Certainly, both conditions represent aspects of incomplete monozygotic twinning. To this end, the decision as to the number of individuals involved was left with the attending physician. Twin pairs with a single birth certificate were excluded from consideration. The underenumeration error thereby committed is small since such teratology is quite rare.

Accuracy and Completeness of Birth Records

The great virtue of the birth and death reporting system as an information source lies in its inclusiveness from a population standpoint. For present purposes, the population under study consists of all deliveries of over 20 weeks gestation with the sole selection criteria of occurrence and notification within the jurisdiction of Upstate New York. Hospital selection bias may thereby be avoided. The major defect in the system lies in the lack of uniformity in the basic observations. While a given formatted record is required by law within a particular state, this in only a small way insures comparability and accuracy of the raw data as recorded by the many parties involved—mothers, nurses, medical students, interns, residents, obstetricians, and secretaries. Beyond the simple fact of birth or death, the number of characteristics per case is small and the medical items of variable quality and completeness. A cursory investigation of the manner in which certificates are filled out in different institutions reveals striking variations in procedure, definition, and care of preparation.

A number of substudies of the quality of the medical certification portion of the birth record have been completed in New York utilizing several information sources as check. Direct comparison of a random sample of nearly 13,000 live birth and 1,000 stillbirth certificates with hospital records revealed an underreporting on the former of 50 to 60 per cent for nonpuerperal and puerperal complications of pregnancy. Twenty-two per cent of the delivery operations and 33 per cent of the complications of labor were not recorded on the birth records. No attempt was made to evaluate the standard, in this case the hospital record. Beginning with a group of 3,000 cases receiving treatment for specific congenital malformations under the State Aid Program, examination of the corresponding birth records showed that reporting completeness varied between 10 and 85 per cent depending on the condition. As anticipated, internal defects were generally not noted at the time of birth, while external defects such as spina bifida and phocomelia tended to be picked up by the system. Matters are considerably improved by recourse to death certificate data. This is particularly true for congenital hydrocephalus and for those internal anomalies not compatible with extended postnatal survival.

The accuracy of such items as sex, parental ages, and birth weight was

deemed to be purely a function of the care with which the information was recorded and was not investigated. Gestation length is demonstrably of low accuracy with excessive digit preference at 36 and 40 weeks. The questions relating to parity on the form were badly stated and undoubtedly lead to inappropriate responses. Number of previous children cannot be equated with number of previous confinements which may include abortions and prior twin deliveries. On examination, members of twin sets were found to be variously recorded as either sharing the same birth order or occupying adjacent orders. Operationally, the item was regarded as a crude index of prior pregnancies with a systematic bias. It does not appear that matters can be improved by a simple rewording of the question. A short pregnancy history would be a highly desirable addition to the record.

Data Processing

The third problem is one of developing information on twin maternities from a deck of tabulating cards wherein the unit of classification is the individual birth. A single code indicating plurality permitted a 98 per cent reduction of the birth deck, the remaining 2 per cent comprising the 2 N individual members of N twin sets. Unfortunately, data on order of birth, first or second, within sets had not been key punched. At this point, the problem bore a superficial resemblance to the situation of a person attempting to select pairs of shoes from a dark closet containing over 40,000 single shoes, each of which hopefully possesses one and only one mate. Under random selection the probabilities involved have been well described by Mr. Feller and offer little comfort. Actually, the closet was only partially dark because twins by definition share the same parents, if superfetation is ruled out, usually on the same day and at the same place, and these data were available in digital form. A few duplicate records were removed from the file by testing for equality of the eight-digit certificate number field. The shared information was treated as a 40 character control field by which the records were sequenced. Since the probability of a false match was estimated to be, at most, one in one billion, mates should have occupied adjacent positions in the sequence, at least under ideal circumstances. Based on this presumption, a merge procedure was developed whereby adjacent records were sequentially tested for an equality condition in the control field. If equal, a combined record was written. If unequal, the record was transferred to an unmatched file for clerical resolution.

About 8 per cent of the cases fell into the unmatched group, a remarkable result in that the 8 per cent represent the probability of the realization of at least one discrepancy in each of the 40 characters. To improve matters further, certain discrepancies are legitimate and may be anticipated. For example, nearly ½ per cent of the time, twins will be born on different days since the interval between births may range from minutes to hours. The remaining errors were those of omission and commission, dropped cards, spelling permutations, miscoding, and keypunching.

Discrepancies in place of occurrence were most often of the type where the first twin was born at home, in a taxi, or in a small general hospital, while the second was born in a large medical center. Evidently, the twin delivery was either not anticipated or complications arose before the birth of the second, necessitating specialized care. Difference in race within twin sets was rare, the single instance being that of one white and one Indian child in the same set. Further investigation revealed that the residence was a reservation and by arbitrary decision Indian concordance was

ruled. The Gilbert-Ilbert twins exemplify the infrequent name variation discrepancies which can clearly be ascribed to keypunch error. In general, the major source of measurable discrepancy was found to arise in the hospital. Table 1 gives a list of the variables available for analysis.

Estimates of Zygosity

Given a sample of twin maternities classified by the usual birth record variables, a central problem is the development of estimates of zygosity, that is, how many and which sets have arisen from a single ovum and how many from two ova. Certainly, the birth certificate contains no information by which a direct diagnosis can be obtained. At the most, sex composition furnishes the sole criterion whereby monozygosity can be ruled out. Thus, sets consisting of a boy and a girl are taken to be dizygotic while like sexed sets, boy-boy and girl-girl, may be of either type.

Table 1—Twinship Data Format: A. First Twin; B. Second Twin

Variable No.	Variable Name	Code Range		Character	
		Min	Max	BCD	BIN
1	A. place of occurrence	0100	7000	4	13
2	A. place of residence	0100	7000	4	13
3	A. birth certificate number	—	—	6	36
4	B. birth certificate number	—	—	6	36
5	A. institution of birth	0	50	2	6
6	A. sex	0	1	1	1
7	B. sex	0	1	1	1
8	A. plurality	0	5	1	3
9	A. date of birth	010150	123160	6	18
10	A. age of father	15	99	2	7
11	A. age of mother	12	99	2	7
12	A. color and nativity	0	9	1	4
13	A. weeks of gestation	20	50	2	6
14	A. birthweight	0	12	2	4
15	B. birthweight	0	12	2	4
16	A. previous children, total	0	30	2	5
17	A. previous liveborn, now dead	0	10	1	4
18	A. previous stillborn	0	10	1	4
19	A. complications of pregnancy	0	12	1	4
20	A. nonpuerperal complications	0	12	1	4
21	A. complications of labor	0	12	1	4
22	B. complications of labor	0	12	1	4
23	A. delivery operations	0	12	1	4
24	B. delivery operations	0	12	1	4
25	A. congenital malformation	00	90	2	7
26	B. congenital malformation	00	90	2	7
27	A. survivorship	0	7	1	3
28	B. survivorship	0	7	1	3
29	A. age at death	0	59	2	6
30	B. age at death	0	59	2	6
31	A. cause of death	0	999	3	10
32	B. cause of death	0	999	3	10

Format specification: (2I4, 2I6, 1I2, 3I1, 1I6, 2I2, 1I1, 4I2, 8I1, 2L2, 2I1, 2I2, 2I3).

Table 2—Zygosity Estimation on the Basis of Sex Composition

Notation:

- Θ proportion of twin sets dizygotic
- $1-\Theta$ proportion of twin sets monozygotic
- p probability a zygote is of type XY (male) given fertilization has occurred

Assumptions:

a. p is constant between and within sets
b. "trials" are independent
c. no sex selective loss occurs between conception and that point of gestation when observation is possible

Twin Type	Observed Frequency	Theoretical Frequency DZ	Theoretical Frequency MZ	Total
MM	a	$p^2\Theta$	$p(1-\Theta)$	$p-pq\Theta$
MF	b	$2pq\Theta$	0	$2pq\Theta$
FF	c	$q^2\Theta$	$q(1-\Theta)$	$q-pq\Theta$
Total	1	Θ	$(1-\Theta)$	1

Maximum Likelihood Estimator:

Under the model, the likelihood of the sample values is

$$e^L=(p-pq\Theta)^a(2pq\Theta)^b(q-pq\Theta)^c$$

Maximizing L with respect to the two parameters leads to the equations

$$S_\Theta=\frac{b}{\Theta}-\frac{aq}{1-q\Theta}-\frac{cp}{1-p\Theta}=0$$

$$S_p=\frac{a+b}{p}-\frac{b+c}{q}+\frac{a\Theta}{1-q\Theta}-\frac{c\Theta}{1-p\Theta}=0$$

which admit as solution the maximum likelihood estimators

$$\hat{\Theta}=\frac{2b}{1-(a-c)^2} \qquad \hat{p}=a+\tfrac{1}{2}b$$

Weinberg Estimate:

The corresponding Weinberg estimate is $\tilde{\Theta}=2b$ which is the same as the MLE when $a=c$

An early solution to the problem of estimating zygosity by indirect means was proposed by Weinberg (1902) who assumed a sex ratio of unity and independent sex determination among the dizygotic sets. Thus he reasoned that like and unlike sets should be about equally represented among the two egg twins. On this basis, the number of monozygotic sets is simply taken as the residual after the unlike have been subtracted from the like sexed sets. Translated into probabilistic terms, the assertion is that sex is a binomial type variable or, specifically, that the probability of a child being male is constant between and within families. (See Table 2 for estimating procedure.) The validity of this assumption is critical and should be the subject of empirical confirmation.

Review of the extensive sex ratio literature on human sibships and mammalian litters leads to conflicting re-

sults. A bewildering variety of observations have been collected which suggest that the male probability may vary between families or be a function of such factors as maternal age and parity, nutritional status, time of fertilization during the menstrual cycle, background radiation, ABO blood type, and so forth. Primary variations in the ratio of X-bearing and Y-bearing spermatozoa produced, variation in maternal conditions favoring fertilization by one type of sperm, and sex selective early loss of embryos would all operate to upset the binomial pattern in the direction of overdispersal. An unequal load of recessive X-linked lethal mutations would lead to heterogeneity in the binomial probability. Under such conditions, the Weinberg assumption would result in an underestimate of the proportion dizygotic. On the other hand, data have been published which show an underdispersed sex distribution when compared to the binomial expectation. This result, if applicable, would mean that the Weinberg estimate is an overstatement. Elaboration of the mechanism underlying such a finding, as well as those showing overdispersal, must await definitive investigation obtainable only through direct zygosity classification. The present intention is merely to gauge the potential range of effect on estimation of the proportion dizygotic by departures from the binomial on the basis of pertinent observational data.

Utilizing a generalized two parameter form of the binomial, the conditions of positive and negative sex correlation within sibships, as well as those of heterogeneity in the male probability between and within sibships may be simulated. For a sequence of trials of length n, each trial having two possible outcomes, Skellan[4] has proposed the following probability distribution to account for heterogeneity in the binomial probability:

$$\Pr(M=k)=\binom{n}{k}\frac{p(p+z)\dots(p+kz-z)q(q+z)\dots(q+nz=kz-z)}{(1+z)(1+2z)\dots(1+nz-z)}.$$

The same distribution may be derived by assuming a constant correlation between successive outcomes. The first parameter corresponds exactly to the single binomial parameter p, the probability a birth is male, while the second parameter, z, may be termed a "dispersion coefficient." For z-positive, the distribution is overdispersed relative to the binomial; for z-negative, underdispersed; while for $z=0$, the two distributions are identical in form. Utilizing several series of observations, the dispersion coefficient z was found to fall in the range -0.03 to $+0.02$. In terms of the 21,128 twin sets born in Upstate New York during the study period, such a range in z corresponds to a range of 0.626 to 0.657 in the proportion dizygotic 0 for the entire sample. To our view this constitutes a small variability in terms of the limited purposes intended and gives us a measure of confidence in the zygosity estimation procedure (Table 3).

Measure of Concordance

A second type of analytic problem is the development of approaches for the study of concordance of traits within the two zygosity classes. To this end, consider the simple model wherein it is assumed the probability a child is affected with a given trait is a constant p and the correlation between sibs is a constant r. Under this setup, the random variable defined as the number of sibs affected within a twinship may be shown to have the distribution:

Number Affected (M)	Observed Frequency	Probability
2	a	p^2+pqr
1	b	$2pq-2pqr$
0	c	q^2+pqr
Total	1	1

Table 3—Observed Number of Twin Sets by Sex Composition and Estimated Number of Sets by Zygosity Class (Upstate New York, 1950-1960)

Twin Type	No. of Males	Relative Frequency	Observed No. Sets	Estimated No.	
				DZ	MZ
MM	2	a	7,300	3,507	3,793
MF, FM	1	b	6,818	6,818	0
FF	0	c	7,010	3,319	3,691
Total		1	21,128	13,644	7,484

Proportion male of all twin individuals

$$p = a + \tfrac{1}{2}b$$
$$= 0.5069$$

Proportion dizygotic of all twin sets

$$\Theta = 2b/[1-(a-c)^2]$$
$$= 0.6458$$

Maximum likelihood estimates of the parameters are:

$$p = a + \tfrac{1}{2}b \qquad r = 1 - \frac{2b}{1-(a-c)^2}.$$

The parameter r will be taken as a measure of concordance.

Evidently, the previously mentioned zygosity estimation procedure is closely related. The proportion monozygotic $(1-\Theta)$ has the identical estimator as the concordance coefficient r. This is another way of saying that monozygosity is estimated as the basis of sex concordance within sets.

To fix ideas, consider the circumstance where all the pairs are concentrated in the extreme two classes $M=2$ and $M=0$; i.e., whether both members have or both members do not have the trait. In such a case, common parlance would dictate that the situation is one of perfect concordance and the coefficient r would have the value of unity. At the other extreme, where all pairs are concentrated in the center class $(M=1)$, the situation would be one of complete discordance and r would have the value of minus one. Two examples will suffice: Twins are perfectly concordant for the variable race, at least in that the issue has been forced by definition. Similarly, twins are completely discordant for the variable birth order within set in that if one is first born, the other must be second born and vice versa. The intermediate situation of $r=0$ (no correlation) leads to independence and the simple binomial distribution.

A practical example of complete concordance arose in our attempt to estimate the coefficient r for each category of congenital anomaly appearing in the International List. Initially the highest value of r was noted for the code number 750 which is titled "Monstrosity" and includes such diverse entities as anencephalus, hemicephalus, ischiopagus, podencephalus, macrocephalus, and megalocephalus, monster, thoracopagus, and syncephalus. The latter two are twins conjoined at the thorax and head, repectively. Certainly Siamese twins must be concordant for the characteristic of being conjoined, since one cannot be joined with the other without the other being joined with the one.

A general notion of how the frequen-

Table 4—Twin Sets by Sex Composition and Number Affected for Selected Characteristics (Upstate New York, 1950-1960)

Characteristic	Like Sexed (MM, FF)				Unliked Sexed (MF, FM)			
	0	1	2	r	0	1	2	r
Birth order (1st, 2nd)	0	14,310	0	−1.00	0	6,818	0	−1.00
Race (Wh, N Wh)	831	0	13,479	1.00	491	0	6,327	1.00
Sex (maleness)	7,010	0	7,300	1.00	0	6,818	0	−1.00
Thoracopagus, etc.	14,298	0	12	1.00	6,818	0	0	—
All malformations:	13,901	365	44	0.18	6,630	186	2	0.00
CNS	14,196	102	12	0.19	6,775	42	1	0.04
Heart	14,234	70	6	0.14	6,791	27	0	0.00
GU	14,278	27	5	0.27	6,806	12	0	0.00
GI	14,255	49	6	0.20	6,789	29	0	0.00
Other	14,194	102	14	0.20	6,746	71	1	0.02
Mongolism	14,304	5	1	0.29	6,813	5	0	0.00
Postnatal asphyxia	14,260	40	10	0.33	6,799	18	1	0.10
Erythroblastosis	14,290	8	12	0.75	6,812	2	4	0.80
Intracranial injury	14,284	16	10	0.56	6,813	5	0	0.00
Fetal death	13,246	695	369	0.49	6,553	216	49	0.30
Perinatal death	11,656	1,192	1,462	0.66	5,922	495	401	0.58

cies are distributed for selected characteristics may be obtained from Table 4. Concordance among the liked sexed sets is in general higher than among the unlike sets. The magnitudes of the differences in concordance coefficients undoubtedly are understatements since about half of the like sexed sets are dizygotic and should behave in a similar fashion to the MF dizygotic sets. Erythroblastosis is an interesting exception as it is the single trait with a well described genetic basis. Theoretically, monozygotic twins should always be concordant while dizygotic twins may, or may not, be discordant if the father is an Rh heterozygote. Unfortunately, the birth certificate is not an ideal source of information on erythroblastosis, a condition which may not be evident immediately at the time of birth.

The possibility of a sex selection incidence of certain characteristics must be borne in mind. Hypospadias, the most common GU malformation reported, is limited to males and MF pairs must be discordant for the trait. On the other hand, anencephalus and isolated cleft palate show a female predominance of nearly two to one. One approach to this problem is presented in Table 5 wherein the modified binomial was applied twice, first for estimating sex concordance and second for estimating fetal death concordance. Further, it was assumed that male and female dizygotic twins had comparable stillbirth rates and the same value of r whether in like or unlike sexed sets. On such a basis, the numbers of monozygotic MM and FF sets with 0, 1, and 2 stillbirths were obtained by subtraction. Clearly, the assumption regarding the behavior of dizygotic twins cannot be tested with available observations and must be taken a priori. Monozygotic stillbirth rates were estimated to be about three times as high as dizygotic rates. Similarly, concordance was higher among the former. The same approach may be applied to the frequency distribution of any qualitative trait within twin sets classified by sex composition. Since the present purposes are merely

Table 5—Observed and Estimated Number of Stillbirths Within Twin Sets by Sex and Zygosity Class (Upstate New York, 1950-1960)

	Observed			Dizygotic		Monozygotic	
	MF	MM	FF	MM	FF	MM	FF
Total Sets	6,818	7,300	7,010	3,507	3,319	3,793	3,691
0 SB	6,553	6,741	6,505	3,364	3,197	3,377	3,308
1 SB male	116	361	—	117	—	244	—
1 SB female	100	—	334	—	100	—	234
2 SB	49	198	171	26	22	172	149
p_m	0.024	0.052	—	0.024	—	0.078	—
p_f	0.022	—	0.048	—	0.022	—	0.072
r	0.296	0.497	0.481	0.296	0.296	0.550	0.526

p_m male stillbirth rate
p_f female stillbirth rate
r stillbirth concordance coefficient

illustrative, a more detailed analysis and interpretation will be left to a later communication.

Conclusion

While vital record information is of limited depth and quality, the registration system at least provides a central repository for combining the experience of all institutions throughout the state. To this end, a sample of over 20,000 twin sets was collected with a minimal effort that was absorbed by the staff and required no extradepartmental support. Development of twinship data from a unit record processing system presented no major difficulties other than the somewhat laborious clerical effort required to reconcile discrepancies by reference to the individual birth and death certificates.

Ascertainment of the critical variables of sex, survivorship, and certain gross anomalies evident at birth are regarded with sufficient confidence to justify the effort and expense involved in processing and analyzing the data. Of particular significance are those rare events which can only be studied by pooling results.

REFERENCES

1. Gittelsohn, A. M., and Milham, S. Declining Incidence of Central Nervous System Anomalies in New York State. Brit. J. Prev. & Social Med. 16:153–158 (July), 1962.
2. Weinberg, W. Beitrage zur Physiologie und Pathologie der Mehrlingsgeburten beim Menschen. Pflüger's Arch. f.d. ges. Physiol. 88:346–379, 1902.
3. Milham, S. Increased Incidence of Anencephalus and Spina Bifida in Siblings of Affected Cases. Science 138:593–594 (Nov. 2), 1962.
4. Skellam, J. G. A Probability Distribution Derived from the Binomial Distribution by Regarding the Probability of Success as a Variable Between Sets of Trials. J. Roy. Stat. Soc. B10:257–265, 1948.

Dr. Gittelsohn is director of health statistics and Dr. Milham is development consultant, New York State Department of Health, Albany, N. Y.

This paper was presented before the Statistics Section of the American Public Health Association at the Ninetieth Annual Meeting in Miami Beach, Fla., October 16, 1962.