A Risk-Adjustment Technique for Comparing Prematurity Rates Among Clinic Populations

RICHARD HEBEL, Ph.D., GEORGE ENTWISLE, M.D., and MATTHEW TAYBACK, Sc.D.

RISK characteristics often vary appreciably from one clinic population to another so that comparison of the effectiveness of a health program among these populations is difficult. Failure to take such differences into consideration can lead to serious errors in the interpretation of patients' performance.

Although a health program would be expected to affect the measures of outcome which reflect patients' performance, there are usually other factors associated with populations of patients which also alter outcome. We shall refer to these concomitant sources of outcome variability as risk factors. A statistical treatment of outcome data which will account for the risk factors is highly desirable.

The authors are in the department of preventive medicine and rehabilitation, University of Maryland School of Medicine. Dr. Hebel is an associate professor of biostatistics, Dr. Tayback is a professor of biostatistics, and Dr. Entwisle is chairman of the department. The work described was supported by Public Health Service Grant No. Ph700. Tearsheet requests to Dr. Richard Hebel, Department of Preventive Medicine and Rehabilitation, University of Maryland School of Medicine, Baltimore, Md. 21201.

Several techniques may be used to adjust for the risk factors. A sophisticated statistical procedure, such as analysis of covariance, is sometimes used for this purpose. By analysis of covariance, one can isolate and measure the effect of each possible source of outcome variability which is identified. Although this method is a powerful means of controlling for the concomitant variables, it has the disadvantage of being computationally complex (usually requiring a computer) and is dependent on the specification of an appropriate mathematical model. A rigorous discussion of analysis of covariance is given by Cochran (*1*).

Alternatively, an intuitive method is often applied, in which the groups to be compared are stratified according to the concomitant variables and comparisons are made only within similar strata. Although this approach is straightforward computationally, interpretation of the results is complicated because a separate set of comparisons is required for each stratum. Additional difficulties arise when small numbers of cases occur in certain of the strata.

If only one concomitant variable is involved, a compromise tactic can be used in which the index of performance is adjusted according to the distribution of the concomitant variable in some standard population. This procedure is called standardization and has long been used by demographers for the comparison of vital rates among different populations. The technique is relatively easy to apply and does not rely upon the specification of a mathematical model. Care must be taken in the choice of the standard population so that the distribution of the concomitant variable is not drastically different from the distribution in the groups to be standardized. Extensive accounts of the standardization method are given in recent texts by Bogue (*2*) and Cox (*3*). The effectiveness of standardization relative to analysis of covariance has been examined by Cochran (*4*), and this method appears to be satisfactory for most applications.

It is frequently desirable, however, to adjust for the joint influence of several concomitant variables. Such is the case in a study to determine the effect of prenatal services on the outcome of pregnancy. We shall consider the methodology of a standardization approach which can be applied to the concomitant variables involved in the evaluation of maternity care. Since the concomitant variables to be considered are risk factors, we shall refer to our technique as risk adjustment.

Consideration of Risk

The rate of infants born weighing 2,500 gm. or less, commonly called the prematurity rate, is an important performance index in evaluating maternity care programs and will be used in constructing and discussing a model of risk adjustment. The need for risk adjustment became apparent from studies in the department of preventive medicine and rehabilitation, University of Maryland School of Medicine, that focused on evaluation of the U.S. maternity and infant care program. This program sponsors many clinics providing prenatal care for low-income women, and the differences in the characteristics of patients from clinic to clinic are considerable.

To compare fairly the prematurity rates among these clinics, adjustment must be made for those variables which influence the outcome of pregnancy. Such variables, identified at the time of the patient's registration, would include age, number of previous pregnancies (parity), smoking habits, preconception weight according to height, race, any prior history of premature delivery or fetal death, period of gestation, and the time elapsed since termination of the last pregnancy. If sufficient data were available for each patient, an adjustment procedure could take into account all of these risk factors. For illustration, however, we shall confine our discussion to a consideration of the first four factors—age, parity, preconception weight according to height, and smoking habits.

Before looking at the joint influence on the prematurity rate of the selected risk factors, let us first consider the effect of each factor by itself, making use of data collected at the outpatient department of the University of Maryland Hospital, Baltimore, on all black patients who registered for prenatal care during the period 1962–67. For this group of 7,945 women, the prematurity rate was 14.4 percent. From prenatal registration to delivery, the patients were carefully followed by Dr. Maureen Henderson as part of a study she was doing on bacteriuria.

Table 1 shows the prematurity rates for this group by each of the four selected risk factors. The height-weight classifications "small," "medium," and "large" correspond to categories 1–2, 3–5, and 6–7, respectively, on the 7-point scale described by Moore and associates (*5*). (These authors provide a chart from which one can determine the height-weight category for any combination of height and weight.) For our analysis, patients who smoked cigarettes in any amount

Table 1. Prematurity among black infants born alive in single births, according to mother's age group, parity, height-weight category, and smoking habits, University of Maryland outpatient department, 1962–67

Risk factors	Cases	Prematurity rate
Age group (years):		
Under 20	2,385	16.7
20–34	4,835	13.5
35 or over	724	13.1
Previous pregnancies:		
0	2,078	16.4
1–3	3,298	14.2
4 or more	2,569	13.1
Height-weight category:		
Small	1,459	19.8
Medium	5,603	14.1
Large	826	7.6
Smoking habits:		
Smoker	4,276	17.0
Nonsmoker	3,530	11.4

were regarded as smokers. Table 1 gives some clues for identifying those prenatal patients at highest risk of having a premature infant. We see that (*a*) teenagers are at higher risk than older women, (*b*) primiparas are at higher risk than multiparas, (*c*) slightly built women are at higher risk than heavier women, and (*d*) smokers are at higher risk than nonsmokers.

A more complete picture of risk is obtained by the four-way cross-classification given in table 2. Here all combinations of the four risk factors are delineated. As would be expected, some of the cells in table 2 contain few or no patients, and the highly variable prematurity rates in these fringe areas complicate interpretation of such cells. Nevertheless the impressions gained from this table are similar to those gained from table 1, except that some effects of interaction are now apparent. For example, we note that age and parity play a more important role for nonsmokers than for smokers.

If we could obtain a cross-classification, as in table 2, for all the clinic populations to be considered, then by restricting comparisons of prematurity rates between clinics to identical cells, we could simultaneously control for the four risk factors. We could compare prematurity rates among small teenage primiparas who smoke cigarettes, for example. To use all the data, however, we would require 54 such comparisons, and many of these would be uninformative because of the small numbers of patients. As an alternative, we might consider an adjustment procedure which would control for the risk factors while simplifying our task of interpretation. A weighted prematurity rate would be computed for each clinic, the weighting based upon the distribution of all combinations of the risk factors in a standard population. Prematurity rates thus adjusted would be subject, however, to a great deal of variability since few patients in the clinic populations would be expected to fall in certain of the categories. It would therefore be highly desirable to obtain broader classifications of risk which would contain appreciable proportions of the clinic populations.

Classification of Risk

One way to arrive at broader classifications of risk would be to combine those cells of table 2 which appear to have a similar outcome. Table 3 illustrates a grouping which leads to three classes of risk, identified as high, intermediate, and low. As is seen by comparing table 1 with table 2, cells with expected prematurity rates near or above 20 percent are defined as high risk; cells with expected rates near or below 10 percent are defined as low risk; all others are put in the intermediate risk class. The prematurity rates and the numbers

Table 2. Prematurity rates for black infants born alive in single births, according to a cross-classification of selected risk factors in their mothers, University of Maryland outpatient department, 1962–67

Age group of mothers (years)	Small-sized [1] mothers with parity of—			Medium-sized [1] mothers with parity of—			Large-sized [1] mothers with parity of—		
	0	1–3	4 or more	0	1–3	4 or more	0	1–3	4 or more
	Smokers								
Mothers under 20:									
All infants	226	110	4	520	336	7	17	14	0
Percent premature	24.3	18.2	75.0	16.9	16.1	28.6	11.8	7.1	....
Mothers 20–34:									
All infants	77	262	129	209	918	803	28	93	141
Percent premature	11.7	19.1	23.2	13.4	18.0	16.2	7.1	9.2	8.5
Mothers 35 or over:									
All infants	1	3	32	6	60	164	3	18	67
Percent premature	0.0	66.7	31.2	16.7	16.7	20.1	33.3	11.1	9.0
	Nonsmokers								
Mothers under 20:									
All infants	199	79	2	461	307	7	10	14	2
Percent premature	20.1	21.5	0.0	15.4	11.1	28.6	0.0	0.0	0.0
Mothers 20–34:									
All infants	56	149	82	179	687	629	17	116	155
Percent premature	23.2	15.4	12.2	13.4	8.3	9.5	11.3	6.9	7.7
Mothers 35 or over:									
All infants	1	6	12	5	36	177	1	11	97
Percent premature	0.0	50.0	8.3	20.0	5.6	7.9	0.0	0.0	6.2

[1] Height-weight category.

of cases for the population of the University of Maryland outpatient department according to this risk classification are as follows:

Risk class	*Prematurity rate (percent)*	*Number of cases*
High	21.1	1,181
Intermediate	16.0	4,230
Low	8.4	2,338

A more objective statistical procedure, such as discriminant analysis, might also be used to establish the risk classification. Norris and associates (*6*) discuss a method of discriminant analysis as it is applied in a study of acute myocardial infarction. Greenberg and Wells (*7*) have also used discriminant analysis in a study of perinatal mortality. The effectiveness of the discriminant function as a classification tool, however, depends upon the accuracy of its model. If the effects of certain factors are nonlinear or if the interactions among the factors produce significant effects, discrimination techniques such as those described by Hills (*8*) and Belson (*9*) would perhaps be more appropriate.

Having defined the classes of risk, we can identify women in any clinic population with similar prognoses of outcome. This concept alone could be very valuable in clinical practice, for example, as an indication of the patients in need of the more intense care. From the analytical point of view, however, we look upon risk classification as

Table 3. A risk classification system for prenatal patients based upon smoking habits, age group, size (height-weight category), and parity

Age group (years)	Small-sized with parity—			Medium-sized with parity—			Large-sized with parity—		
	0	1–3	4 or more	0	1–3	4 or more	0	1–3	4 or more
Smokers:									
Under 20	H	H	H	I	I	I	*L*	*L*	*L*
20–34	H	H	H	I	I	I	*L*	*L*	*L*
35 or over	H	H	H	I	I	I	*L*	*L*	*L*
Nonsmokers:									
Under 20	H	H	H	I	I	I	*L*	*L*	*L*
20–34	H	I	I	I	*L*	*L*	*L*	*L*	*L*
35 or over	H	I	I	I	*L*	*L*	*L*	*L*	*L*

H—High-risk I—Intermediate risk *L*—Low risk

Table 4. Comparison among 8 maternity and infant care clinics of prematurity rates for black infants born alive in single births in 1967

Clinic	Total number of infants and percent premature	Risk-classified rates: High—r_1	Risk-classified rates: Intermediate—r_2	Risk-classified rates: Low—r_3	Risk—adjusted rates—r_a	Standard errors: $SE(r)$	Standard errors: $SE(r_a)$
Clinic A:							
Number	156	39	91	26			
Percent	19.2	25.6	18.7	11.5	17.6	2.81	2.84
Clinic B:							
Number	166	24	76	66			
Percent	12.6	20.8	14.4	7.6	13.3	2.72	2.82
Clinic C:							
Number	198	45	108	45			
Percent	16.6	31.1	12.0	13.3	15.3	2.50	2.47
Clinic D:							
Number	857	174	446	237			
Percent	14.8	20.1	16.6	7.6	14.4	1.20	1.19
Clinic E:							
Number	128	21	56	51			
Percent	7.8	23.8	7.1	2.0	8.1	3.10	3.22
Clinic F:							
Number	228	62	133	33			
Percent	15.8	17.7	16.5	9.1	14.4	2.32	2.34
Clinic G:							
Number	401	84	199	118			
Percent	13.2	17.9	14.6	7.6	13.0	1.75	1.75
Clinic H:							
Number	214	47	115	52			
Percent	15.0	23.4	14.8	7.7	14.0	2.40	2.38

a means of controlling for the effect of the concomitant variables. If we use the risk classification described in table 3 and restrict the comparisons of rates between clinics to the same risk classes, then we have reduced the required number of comparisons to three. These separate comparisons may be highly informative in clinics focusing their programs on selected segments of the patient population.

Risk Adjustment

Nonetheless, the convenience offered by the computation of a single prematurity rate for each clinic, adjusted for risk, remains appealing, and a technique to determine such risk-adjusted rates will be demonstrated.

The risk-adjusted rate r_a is given by

$$r_a = \sum w_i r_i \qquad (1)$$

where

w_i = the fraction of the standard population in risk class i.

r_i = the rate for the clinic population in risk class i.

Using the University of Maryland outpatient department as the standard population and with the risk classification defined as in table 3, we have

$$r_a = 0.152r_1 + 0.546r_2 + 0.302r_3, \qquad (2)$$

where the subscripts 1, 2, and 3 refer to the high, intermediate, and low risk classes.

Prematurity rates for 1967 for eight different maternity and infant care clinics are listed in table 4. Also shown are the rates for these clinics in each of the three risk classes and the risk-adjusted rates as determined from formula 2. Furthermore, if we hypothesize that the clinic experiences are samplings from a population identical to the standard population, we can determine standard errors for the adjusted rates as well as for the unadjusted rates. The standard errors thus determined are given in the last two columns of table 4. It might be noted that the standard error of the adjusted rate was computed as

$$SE(r_a) = \sqrt{\sum \frac{W_i^2 R_i (100-R_i)}{n_i}}, \qquad (3)$$

where

R_i = the rate for the standard population in risk class i.

n_i = number of cases for the clinic population in risk class i.

The effect of the risk-adjustment procedure is shown in the figure, in which unadjusted and risk-adjusted prematurity rates for the eight clinics are compared. Baseline rates as determined from the standard population are included in this figure, as well as 95 percent confidence bounds. These bounds are based on samples taken from the standard population which correspond in size to the number of cases for each clinic.

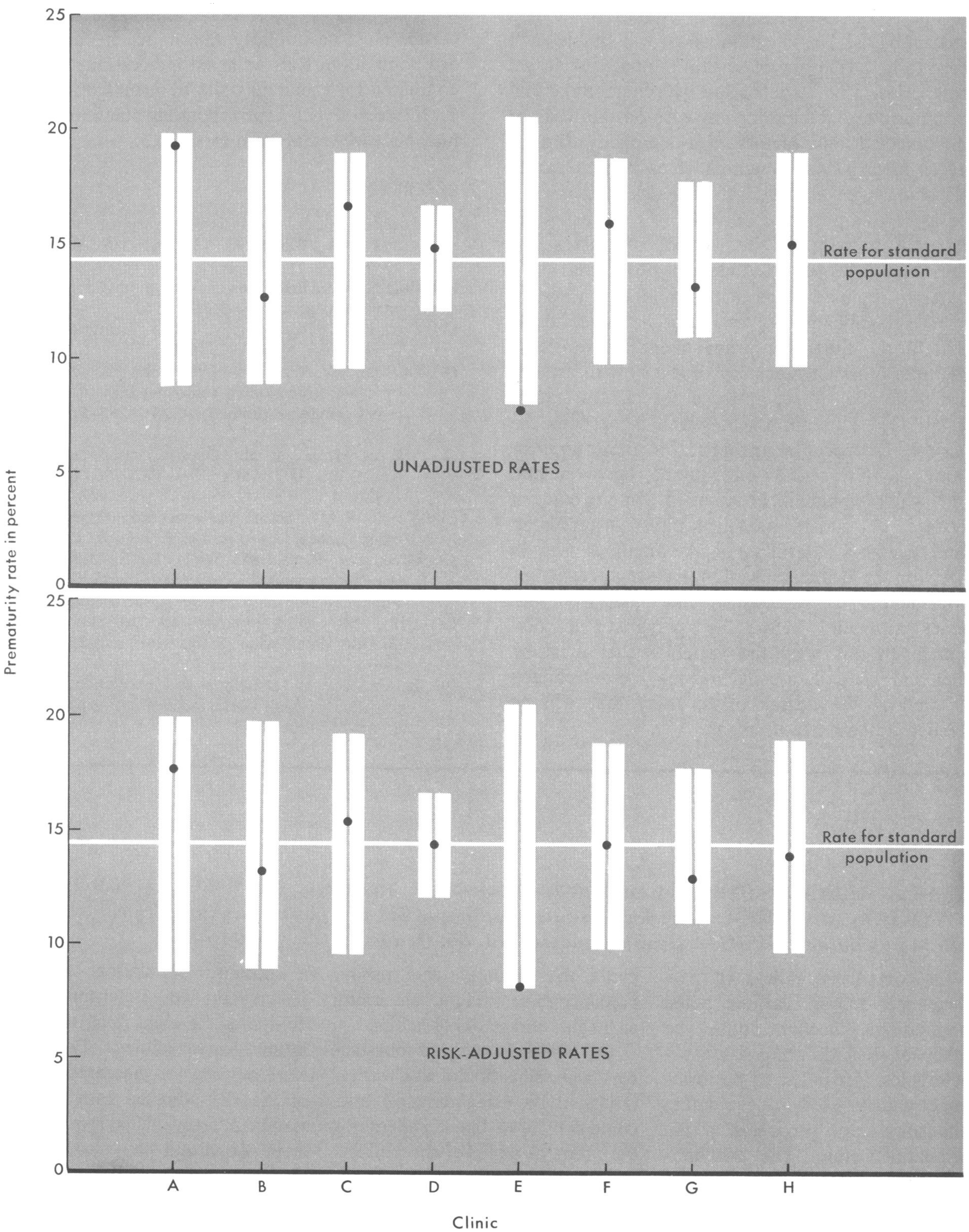
We notice that in most instances the clinic rates are brought closer to the standard rate after the risk adjustment is made. In fact, the prematurity rate for clinic E falls outside the 95 percent confidence bounds before the adjustment but is within the limits of sampling error after being adjusted for the selected risk factors. Then, at least for the clinics under consideration, risk adjustment strengthens the hypothesis that there is no real difference between the clinic rates and the standard rate. Of course, if biases which were presumably caused by an unbalanced distribution of risk factors actually masked a true difference between these rates, the risk adjustment would then operate to reject the "no difference" hypothesis.

Extension of Technique to Other Studies

The study of prematurity used in our paper to illustrate the risk-adjustment technique has much in common with other studies of medical and public health problems. To attain the numbers of cases needed for statistical validity and to provide the diversity of experience required for a realistic evaluation, multiple sources of data are required. Yet, unless random allocation is possible, some way of controlling for the effect of nuisance variables must be sought so that the results from multiple sources can be fairly compared. If such control cannot be attained through the study design, it must be attempted statistically. Since the risk-adjustment procedure represents such an attempt, we shall review in a general way the steps in this procedure.

Obviously, the first step would be selection of the relevant risk variables. Since the risk classification will be simplified if the number of these variables is kept to a minimum, only those factors which significantly influence outcome should be considered. A note of caution must be interjected here. It is important that the risk variables be independent of the effects of medical treatment or whatever kind of intervention is involved (in the prematurity study, these were the effects of the maternity and infant care services). Otherwise,

Comparison among 8 maternity and infant care clinics of unadjusted and risk-adjusted prematurity rates for black infants born alive in single births in 1967



NOTE: Intervals indicate 95 percent confidence bounds.

such effects could be unwittingly removed or falsely exaggerated by use of the risk-adjustment procedure.

The second step entails the establishment of a system of risk classification which will enable each case to be put in one of several mutually exclusive categories. This classification will characterize the prognosis of outcome for any case on the basis of the selected risk factors. For simplicity and to avoid having small numbers of cases in separate classes, we recommend that the risk classes used be as few as possible. Cochran (*4*) discusses the relationship between the number of classes and the effectiveness of an adjustment procedure.

As to the method used in formulating a system of risk classification, several approaches are available. In the illustration provided earlier, we used a multiple cross-classification of data on the standard population to set up three risk classes (tables 2 and 3). This approach, which is perhaps the easiest, is probably adequate for most applications. A more objective technique, however, and one which would likely result in a more efficient system of classification, requires discriminant analysis. But regardless of the method used in deriving the risk classification, we suggest that this classification be established with data from another population before the study data are collected. In this way, the variables that must be included in the data collection may be determined as well as the degree of accuracy that will be needed in their measurement.

Having obtained the appropriate data from each study site, we must determine indices of outcome for the various risk classes (as in table 4, for example). Indices of outcome for the risk classes must be similarly determined in the standard population. Risk-adjusted indices may then be computed for each study site by means of formula 1. If necessary, the corresponding standard errors may be worked out from formula 3.

REFERENCES

(*1*) Cochran, W. G.: Analysis of covariance: its nature and uses. Biometrics 13: 261–280, September 1957.

(*2*) Bogue, D. J.: Principles of demography. John Wiley and Sons, Inc., New York, 1969.

(*3*) Cox, P. R.: Demography. Ed. 4. Cambridge University Press, New York, 1970.

(*4*) Cochran, W. G.: The effectiveness of adjustment by subclassification in removing bias in observational studies. Biometrics 24: 295–313, June 1968.

(*5*) Moore, M. E., et al.: Obesity, social class, and mental illness. JAMA 181: 962–966, Sept. 15, 1962.

(*6*) Norris, R. M., et al.: A new coronary prognostic index. Lancet 7589: 274–278, Feb. 8, 1969.

(*7*) Greenberg, B. G., and Wells, H. B.: Linear discriminant analysis in perinatal mortality. Am J Public Health 53: 594–602, April 1963.

(*8*) Hills, M.: Discrimination and allocation with discrete data. Appl Statist 16: 237–250, June 1967.

(*9*) Belson, W. A.: Matching and prediction on the principle of biological classification. Appl Statist 8: 65–75, June 1959.

HEBEL, RICHARD (University of Maryland School of Medicine), ENTWISLE, GEORGE, and TAYBACK, MATTHEW: ***A risk-adjustment technique for comparing prematurity rates among clinic populations. HSMHA Health Reports, Vol. 86, October 1971, pp. 946–952.***

In comparing indices of performance among various clinic populations, account must be taken of their differing risk characteristics. An index of performance widely used in evaluating maternity care programs is the prematurity rate. The concomitant variables considered to be risk factors for patients of maternity and infant care clinics include the mother's age, parity, preconception weight according to height, and smoking habits.

On the basis of these four factors, a classification of risk was established by using data on outcome for 7,945 black patients of the prenatal clinic of the University of Maryland outpatient department. With these patients as a standard population, risk-adjusted rates were determined and compared for eight maternity and infant care clinics. Formulas were devised for the computation of the risk-adjusted rate and its standard error. This risk-adjustment technique can also be extended to other studies of medical and public health problems.