This study used data from the Swedish Medical Birth Registry between 1982 and 1995 to address the question of whether there is higher mortality in twins in relation to singletons of the same gestational age and to examine the optimal gestational age range for twins. A “varying-coefficient approach” was adopted to estimate the gestational age-specific relative and absolute risks of mortality in twins and singletons, adjusting for size at birth and risk factors of short gestational duration. The models showed that twins born between 29 and 37 weeks of gestation had lower mortality than did singletons of the same gestational age. Twins born at older gestational age had higher mortality than did their singleton counterparts, because longer gestational duration was more advantageous to singletons than to twins. Without adjustment for size at birth, there was an upturn of mortality in twins born after 38 weeks. It is postulated that twins have better health than singletons initially, but they could not enjoy the benefit of a longer gestational duration as much as singletons could. The optimal gestational age for twins appeared to be 37–39 weeks according to neonatal and infant mortality. Am J Epidemiol 2000;152:1107–16.
analyzes included only those with gestational age of this paper, the descriptive analysis of gestational age twins (15); the corresponding figure in this study is 5,328. In of twins in the Registry during 1983–1985 identified 5,323 consisted of 32,942 twins and 1,466,068 singletons, excluding a same pregnancies. The study population of this paper con-
saronal identity number. We used four variables, that is, moth-
exact identification of mothers in terms of the Swedish per-
of confidentiality reasons, the data set does not allow for
newborn and his/her parents. A total of 34,907 babies were
recorded for 316 (0.02 percent), and they were consequently
1,505,736 babies registered; of these, the sex was not
maternal smoking. This study used data from 1982 to 1995.
Registry has widened to collect more information, such as
from 1973 to 1995. Since 1982 the Swedish Medical Birth
registry (14). Data are now available for the birth cohorts
1973. It covers virtually all (99 percent) births in the country
MATERIALS AND METHODS
Materials
The Swedish Medical Birth Registry came into effect in 1973. It covers virtually all (99 percent) births in the country (13). A wide range of demographic and medical data is collected from interviews with mothers and from the hospital medical records. The Swedish Medical Birth Registry is linked to the Registry of Population, which includes the death registry (14). Data are now available for the birth cohorts from 1973 to 1995. Since 1982 the Swedish Medical Birth Registry has widened to collect more information, such as maternal smoking. This study used data from 1982 to 1995.

During the study period from 1982 to 1995, there were 1,505,736 babies registered; of these, the sex was not recorded for 316 (0.02 percent), and they were consequently excluded from this study. Each record gives data about a newborn and his/her parents. A total of 34,907 babies were recorded as multiple births, that is, not singletons. Because of confidentiality reasons, the data set does not allow for exact identification of mothers in terms of the Swedish personal identity number. We used four variables, that is, mothers’ date of birth, hospital of delivery, the babies’ date of birth, and time of birth, for matching the newborns from the same pregnancies. The study population of this paper consisted of 32,942 twins and 1,466,068 singletons, excluding a small number of newborns (182 twins, 4,445 singletons) who were not recorded for gestational age. A previous study of twins in the Registry during 1983–1985 identified 5,323 twins (15); the corresponding figure in this study is 5,328. In this paper, the descriptive analysis of gestational age involved all the twins and singletons. Other descriptive analyses included only those with gestational age of ≥28 weeks (32,312 twins and 1,463,261 singletons). Because of consideration of computing resources, the multivariate analyses involved all the 32,312 twins, and for each twin two singletons of the same sex and gestational age were randomly taken. In the multivariate analyses, the sample size and ratio of twins to singletons could vary slightly in different models because of missing values.

Variables
Birth weight was measured to the nearest 10 g, and the length was measured to the nearest centimeter by nurses. Weight and length were converted into sex- and gestational age-specific standard deviation scores according to a set of Swedish reference standards (16). The standards were developed for the range of 28–42 completed weeks of gestation. The body size of babies born prior to 28 weeks is a topic for a particular investigation. In this study we focus on those born at or above 28 weeks. Extreme standard deviation score values (≤–6 or ≥6) were recoded as missing. Gestational age was estimated by second-trimester ultrasonography when available or by the method of last menstrual period otherwise. In 1982, 20 percent of the obstetric departments performed routine ultrasonography. This percentage increased steadily, and from 1991 onward all pregnant women were offered routine early ultrasound examinations (14). Deaths before the age of 365 days were recorded in the database.

With reference to previous studies about determinants of gestational age and preterm birth (17–19), we estimated the impact of nine factors on “short gestational duration” (gestational age between 28 and 34 weeks inclusive) versus gestational duration over 34 weeks in singletons and twins separately. The reason for this classification of gestational duration will be discussed later. The mother’s health status was classified as with or without any of the four following conditions: diabetes mellitus, hypertension, chronic kidney condition, and systemic lupus erythematosus. Maternal smoking was dichotomized as either smoker or nonsmoker at the first antenatal visit. The mother’s age at delivery was measured in completed years and handled as a linear variable. Mother’s prior stillbirth and prior neonatal death were included as two binary variables. Parity was classified as first parity or second/above. Any significant congenital conditions were noted by an obstetrician and binary coded as either present or absent in this study. The babies’ sex and year of birth were also included in the multivariate analysis. The data were extracted from medical records or enumerated in questionnaires by midwives.

Statistical analysis
The analysis will provide information by twins versus singletons and by gestational age. Statistical tests were carried out to compare twins with singletons. Comparisons of means were based on analysis of variance. Comparisons of medians were based on the Wilcoxon signed rank test. The F test was used for testing equality of variances. Analysis of determinants of short gestational duration was based on multiple logistic regression. The Hosmer-Lemeshow test (20) was used to assess the goodness-of-fit of the models.

For the mortality analysis, logistic regression was used. In one set of regression models, adjustments were made for the factors that may determine gestational age, as well as the standard deviation scores for weight and length and their square terms. In another set, the standard deviation score measures were not included. The differences reveal the impact of body size as a mediator between twin pregnancy

Am J Epidemiol Vol. 152, No. 12, 2000
and health outcomes. Differences in health outcomes between twins and singletons may interact with gestational age. To capture this potential interaction, regression models that allow the impact of twin pregnancy to be a nonlinear function of gestational age (GA) are formulated. The models are in the form of equation 1:

$$\log (p/1 - p) = a_1 + b_1 \times GA + b_2 \times GA^2 + (a_2 + c_1 \times GA + c_2 \times GA^2) \times Twin + D'X$$  \tag{1}$$

where $p =$ probability of death, Twin is dummy coded (singleton = 0, twin = 1), $D$ represents a set of regression coefficients, and $X$ represents a set of covariates. The varying-coefficient models can be seen as models of separate curves for different groups (12). That is, the function $(a_1 + b_1 \times GA + b_2 \times GA^2)$ represents the gestational age-specific curve for singletons; the function $[(a_1 + a_2) + (b_1 + c_1) \times GA + (b_2 + c_2) \times GA^2]$ represents the curve for twins. The impact of twin pregnancy can then be plotted against gestational age. We used linear combination of variances and covariances for the calculation of confidence intervals in the presence of interaction (21). Again, the Hosmer-Lemeshow test was used to check the goodness-of-fit of the models. The modeling was carried out in STATA 6.0 statistical software (22).

RESULTS

Descriptive analysis

The proportions of males among singletons and twins were both 0.51. The twin pregnancy rate in this study was 11.1 per thousand. Gestational age was measured in number of completed weeks. The median for twins was 37 weeks, while that for singletons was 40 weeks ($p < 0.01$). In addition, twins also have a higher level of dispersion in gestational age distribution. The standard deviations were 2.9 and 1.9 for twins and singletons, respectively ($p < 0.01$). These differences are revealed in a plot of the cumulative distributions (figure 1), which shows substantially different intercepts with the reference lines of 2.3, 50, and 97.7 percentiles. The 2.3 and 97.7 percentiles approximately correspond to ~2 and 2 standard deviations from the mean; they can give a sense of the “normal” range. The normal range for twins was from 28 to 40 weeks, whereas the range for singletons was from 35 to 42 weeks. The proportions of twins (0.03 percent, $n = 10$) and singletons (0.07 percent, $n = 1,063$) born over 43 weeks were very small. They were grouped into 43 weeks of gestational age subsequently.

Twins had a higher proportion of missing values and outliers (standard deviation scores of $<6$ or $>6$). The amounts of missing values and outliers among singletons were 0.4 percent and 1.0 percent for weight and length, respectively. The corresponding figures for twins were 0.9 percent and 6.0 percent, respectively. Figure 2 depicts the mean standard deviation scores of the anthropometric measures. There is an obvious pattern of differences in the weight and length measures between twins and singletons over the gestational age range from 28 to 43 weeks. Twins born from 28 to 32 weeks differed little from singletons. Nevertheless, twins born over 32 weeks had a reduced body size, and its magnitude increased continuously and steadily as gestational age increased.

The overall risk ratios of perinatal, neonatal, and infant mortality for twins in relation to singletons were 3.9 (95 percent confidence interval (CI): 3.6, 4.1), 5.4 (95 percent CI: 4.9, 5.8), and 4.2 (95 percent CI: 3.9, 4.5), respectively ($p$ values $<0.01$). Figure 3 shows that the mortality rates decreased as gestational age increased. The mortality curves for twins and singletons cross over at around 37 weeks of gestational duration. Singletons born between 28 and 36 weeks had higher perinatal, neonatal, and infant mortality than did their twin counterparts. Singletons born at older gestational ages, however, had lower mortality levels.

Multivariate analysis of risk factors of short gestational duration

Table 1 shows the outcomes of multiple logistic regression analyses of short gestational duration for singletons and twins separately. Mother’s health status, maternal smoking, maternal age, parity, prior neonatal death, and congenital problems were significant factors of short gestational duration common to singletons and twins ($p$ values $<0.05$). Maternal smoking had a stronger impact among singletons (odds ratio $= 1.39$) than among twins (odds ratio $= 1.16$). The 95 percent confidence intervals did not overlap. The impact of maternal age was qualitatively different in singletons and twins, and the confidence intervals did not overlap.
Older mothers of singletons and younger mothers of twins were more likely to give birth early. Prior stillbirth of the mothers was a significant risk factor of short gestational duration for singletons \((p < 0.01)\) but not twins \((p > 0.05)\). Female singletons were less likely to be born in the gestational age range of 28–34 weeks \((p < 0.01)\); in this same age range, female twins did not differ from male twins \((p > 0.05)\). The year of birth was significant among twins \((p < 0.05)\) but not among singletons \((p > 0.05)\). Hosmer-
Lemeshow tests indicated that both models achieved a sufficient goodness-of-fit \((p > 0.05)\) in predicting short gestational duration.

**Multivariate analysis on mortality**

Varying-coefficient models of the form of equation 1 were fitted for perinatal death, neonatal death, and infant death. Covariates adjusted for included mother’s health, maternal smoking, parity, earlier stillbirth, earlier neonatal death, significant congenital problems, sex, maternal age, year of birth, and standard deviation scores for weight and length. Square terms of the standard deviation scores for weight and length were also included to allow for a nonlinear relation. The results were presented in two ways: log odds ratios and gestational age-specific odds.

Figure 4, left panel, shows the log odds ratios and 95 percent confidence intervals of having the mortality outcomes among twins. Hosmer-Lemeshow tests indicate sufficient goodness-of-fit for the models on neonatal mortality and infant mortality \((p > 0.05)\). However, the fit for perinatal mortality was not sufficient \((p < 0.05)\). This suggests that there were some determinants of perinatal mortality not yet included in the regression models. Twins born between 29 and 38 weeks of gestation had lower perinatal, neonatal, and infant mortality than did singletons \((p < 0.05)\). The log odds ratio was around \(-1\) at 33 weeks, which was the lowest point. The log odds ratio increased with higher gestational age. The 95 percent confidence intervals for perinatal and infant mortality stretched across zero in the range between 39 and 41 weeks. Above that gestational age twins had significantly higher perinatal and infant mortality \((p < 0.05)\).

The curve for neonatal mortality showed a similar pattern but slightly more curvature. Twins born between 30 and 36 weeks had a log odds ratio below the null value; those born at 39 weeks or over had log odds ratios higher than the null value \((p < 0.05)\).

Figure 4, right panel, also shows the log odds of mortality outcomes among twins and singletons. To avoid visual confusion, we did not put in confidence intervals. Since dif-
FIGURE 4. Varying-coefficient models comparing the mortality of twins and singletons, Sweden, 1982–1995. Adjusted for mother’s health, maternal smoking, maternal age, parity, prior stillbirth, prior neonatal death, significant congenital problems, sex, year of birth, and second order polynomials of weight and length standard deviation scores. a, perinatal mortality ($n = 72,116$); b, neonatal mortality ($n = 71,174$); c, infant mortality ($n = 71,174$). Left panel, log odds ratios and 95% confidence intervals for twins in relation to singletons; right panel, log odds of mortality. Solid lines represent twins; dashed lines represent singletons.
ferences in log odds correspond to log odds ratios, significant differences between the curves can be read from where the 95 percent confidence intervals of the log odds ratio do not cross zero. Other factors being held constant, the longer the gestational age, the lower the mortality risk. This description applies to both twins and singletons. The negative slopes of the mortality curves were either roughly constant or accelerating for singletons. In contrast, the negative slopes for twins were decelerating. The curves crossed over around 38–40 weeks. Twins who were born after 38 weeks had high log odds ratios of mortality but not because they had worse health as compared with twins of younger gestational age. Instead it was because the decline of log odds in relation to gestational age among singletons was sharper than it was among twins.

Figure 5, left panel, displays the gestational age curves of log odds ratios of mortality outcomes estimated by models without adjustment for weight and length standard deviation score measures. The shapes of the curves are similar to those in figure 4. However, they climbed up faster after 33 weeks of gestation. Twins born between 29 and 36 weeks had lower perinatal, neonatal, and infant mortality than singletons had. The log odds ratios for a gestational age range over 38 weeks were higher than those estimated with adjustment for body size standard deviation score measures.

The log odds of mortality are also shown in figure 5, right panel. The curves for singletons were very similar to those with adjustment for body size standard deviation score measures. This is expected since the expected standard deviation score values for weight and length for singletons are always zero. The curves for twins had larger curvature than before. For all three mortality outcomes, the twin curves crossed over the curves of the singletons at 37 weeks, in contrast to the crossover points around 38–40 weeks in figure 4. The neonatal mortality and infant mortality U-shaped curves turned upward among twins born at 40 weeks or above.

DISCUSSION

Strengths and weaknesses

This is a population-based study using a large data set from the Swedish Medical Birth Registry. Not many studies have this kind of statistical power to make gestational age-specific comparison related to twins. The data quality was examined in a recent paper and was considered satisfactory (23).

Some twin pairs have one and some have both newborns suffering an adverse outcome. There are two interesting analytical issues. One is to use pregnancy instead of the individual as the unit of analysis. Apart from a slight shift of the theoretical perspective, this will have a problem in statistical adjustment for individual specific covariates, such as sex and body size. Another issue is the dependence of outcomes within twin pairs. Since adverse outcomes may cluster in some twin pairs, this will violate the independence assumption of most statistical methods. Nevertheless, in this study the number of clusters, or twin pairs, is very large (n = 16,471), and the number of subjects within a cluster is small (n = 2). So the practical impact would be small.

The present study could not exhaust all potential confounders. “People are not passive acceptors of threats to health” (24, p. 147). It will be interesting to check whether mothers of twins, as well as the medical personnel, take extra steps to prevent an adverse outcome. In addition, dizygotic twinning appeared to be related to socioeconomic differences, though the determinants of dizygotic twinning are not very well established (8). In contrast, monozygotic twinning seems to be unrelated to socioeconomic factors. Unfortunately, zygosity is not recorded in the Swedish Medical Birth Registry data set. Future studies taking zygosity into account will be helpful. This will require a very large database and well-designed data collection plan. Last but not least, because of limitation of space, this paper analyzed only all-cause mortality. Future investigation on the cause-specific mortality of twins and singletons may shed further light on the impact of twin pregnancy.

Descriptive findings

Twins and singletons had different gestational age distributions. Using 2.3 and 97.7 percentiles to suggest an empirical “normal” range, we found that the range for twins was 28–40 weeks and that for singletons was 35–42 weeks. Consistent with previous studies (10, 25–27), in the present analysis twins and singletons born between 28 and 32 weeks showed little difference in body size. However, those born after 32 weeks of gestational age diverged. The gap of mean size at birth widened steadily with higher gestational age. The present empirical findings generate an interesting question of how to explain this phenomenon. One possibility is that growth is limited by physical constraints in utero, whose impact is particularly pronounced in late gestation (28). Placental function may also impose a similar limitation (29). Another explanation may be that small twins stay longer in utero in order to improve their survival chance. However, this can also be applied to singletons.

The relative risks of mortality were dependent upon gestational age. The perinatal, neonatal, and infant mortality curves crossed over at around 37 weeks of gestation (figure 3). This crossover point is similar to the findings of large-scale studies in Japan and the United States (5, 9). Singleton births have different causes of preterm delivery. In singletons, preterm birth may be more pathological in origin (1), so the crossover might be partly confounded by these pathological factors.

Risk factors of short gestational duration

According to the statistical distribution of gestational age, we suggested separate definitions of preterm birth for twins and singletons, that is, birth prior to 28 weeks and 35 weeks for twins and singletons, respectively. Using these definitions, we investigated the risk factors of short gestation (28–34 weeks), which was quite common in twins but not in singletons. Among the nine factors investigated, four factors had nonoverlapping confidence intervals for twins and singletons. Maternal smoking had an odds ratio of 1.16 (95 percent CI: 1.08, 1.25) in twins and 1.39 (95 percent CI: 1.32,
FIGURE 5. Varying-coefficient models comparing the mortality of twins and singletons, Sweden, 1982–1995. Adjusted for mother's health, maternal smoking, maternal age, parity, prior stillbirth, prior neonatal death, significant congenital problems, sex, and year of birth. Adjustments not made for weight and length standard deviation scores. a, perinatal mortality (n = 75,962); b, neonatal mortality (n = 74,927); c, infant mortality (n = 74,927). Left panel, log odds ratios and 95% confidence intervals for twins in relation to singletons; right panel, log odds of mortality. Solid lines represent twins; dashed lines represent singletons.
1.46) in singletons; prior stillbirth had an odds ratio of 0.87 (95 percent CI: 0.60, 1.25) in twins and 1.57 (95 percent CI: 1.31, 1.88) in singletons. It suggests that delivery during this range of gestational age was more pathological in origin among singletons than among twins. Maternal age was negatively related to early delivery in twins but positively related to early delivery in singletons. The odds ratios were 0.96 (95 percent CI: 0.96, 0.97) and 1.01 (95 percent CI: 1.01, 1.02), respectively. This difference remains unexplained in the present paper. Twins born in more recent years were more likely to be delivered during 28–34 weeks, the odds ratio being 1.03 (95 percent CI: 1.02, 1.04). No significant secular trend was found among singletons, the odds ratio being 1.01 (95 percent CI: 1.00, 1.01). The use of fertility therapy has been increasing in Europe (30, 31). It may have caused an impact on the trend of gestational age in twins.

Multivariate analysis of mortality

Studies on perinatal mortality and morbidity sometimes matched singletons for twins according to gestational age. Given the knowledge that twins and singletons have different gestational age distributions as well as different causes of preterm delivery, we should further explore whether such matching is valid. There is a similar situation in comparing birth weight-specific mortality rates between twins and singletons. It is often found that small twins had lower mortality than small singletons (10, 32). In view of the different distributions of birth weight of twins and singletons, Buekens and Wilcox (32) suggested an approach of comparing mortality at the same “relative birth weight,” which was defined separately for twins and singletons. This approach may shed light on studies of gestational age-specific health outcomes in twins and singletons. However, its obvious limitation is that it cannot handle multivariate analysis. Our approach is to adjust for factors that can adequately predict short gestational duration. This helped to remove the underlying confounding process aforementioned.

A commonly used method of stratification reduces statistical power and has to assume homogeneity within strata. With the use of varying-coefficient logistic regression models with adjustment for potential confounders, substantial interaction between twins and gestational age was discovered. In the presence of interaction between the primary variables and covariates, there is no straightforward interpretation of the regression coefficients (33). By deriving the log odds ratios as functions of gestational age and deriving confidence intervals for the functions, we are able to make a detailed comparison of twins and singletons in relation to important clinical outcomes. Researchers using logistic regression sometimes focus on odds ratios disregarding other information revealed by their regression models. This practice can throw away important information. By seeing a varying-coefficient model as a model of separate curves, we demonstrated the gestational age-specific changes in mortality.

Controlling for the covariates, as well as weight and length standard deviation score measures, twins born between 29 and 37 weeks had lower perinatal, neonatal, and infant mortality levels. After 38 weeks of gestation, the log odds ratios for twins went up as gestational age increased. This increase was most visible in neonatal mortality. Without controlling for body size standard deviation score measures, the log odds ratios rose more sharply after 38 weeks of gestation. As shown in the descriptive analysis, body size differences between twins and singletons widen as the gestational age increased. Without adjustment for size at birth, twins born at higher gestational age showed obvious disadvantages compared with their singleton counterparts. Fetal growth as indicated by body size measures appeared to be a mediator between twin pregnancy and health outcomes.

Multiple logistic regression with adjustment for size at birth has shown that a longer gestational duration was either good or at least not hazardous to both twins and singletons (figure 4). Twins born at a higher gestational age had higher mortality than did their singleton counterparts. This was not because a higher gestational age was hazardous to them, but because a higher gestational age was relatively more advantageous to singletons than to them. In the analysis without adjustment for size at birth, twins’ log odds of neonatal and infant mortality turned upward after 38 weeks of gestation. Their log odds of perinatal mortality decreased monotonically with gestational age.

Putting together the findings, we postulate that twins had better health than did singletons initially, but that they could not enjoy the benefit of longer gestational duration as much as singletons could. Some researchers observed more rapid placental maturation (34), earlier fetal pulmonary maturity, and less bronchopulmonary dysplasia in twins (4, 35). This may partly explain why twins can have better health. Bleker and Oosting (29) suggested that placental function might set a limit in term and postterm twin gestations. Twins may also experience more physical constraints in utero during late gestation (28); they seemed to stop growing after 39 weeks of gestation (36). These limitations may impose a ceiling on the development of twins and explain the interaction observed. When weight and length standard deviation scores were not controlled, gestational age above 38 weeks was associated with increasing log odds in neonatal and infant mortality in twins. On the basis of these two indicators, we would suggest defining the optimal gestational duration for twins as 37–39 weeks. This is slightly longer than the optimal range of 35–38 weeks suggested by Luke et al. (7), which was based on size at birth and hospital length-of-stay. If we used perinatal mortality as an indicator, however, it seems that the longer the gestational duration the better, regardless of multiplicity status. Nevertheless, our analysis on perinatal mortality did not reach a satisfactory goodness-of-fit. Further investigation is needed on this aspect. It also highlights the need for defining what are the health outcomes we want to optimize in twins. Perhaps a quantitative evaluation and combination of several health indicators are required. Finally, the conventional definition of preterm birth (<37 weeks) is adjacent to the lower bound of the optimal gestational age range we mentioned. Again, this seems inappropriate for twins in relation to neonatal and infant mortality.
ACKNOWLEDGMENTS

The project is supported by a CRCG grant (no. 10202179/21317/21100/323/01) from the University of Hong Kong and by the Faculty of Medicine, University of Hong Kong.

The authors thank Dr. Daniel Fong and James Thorburn, both at the Clinical Trials Centre, University of Hong Kong, for constructive comments on earlier versions of this paper.

REFERENCES