Frequent Attendance at Religious Services and Mortality over 28 Years

William J. Strawbridge, PhD, Richard D. Cohen, MA, Sarah J. Shema, MS, and George A. Kaplan, PhD

INTRODUCTION

Associations between religious affiliation and mortality have been analyzed since Durkheim's 1897 comparison of suicide rates among Catholics, Protestants, and Jews. Most commonly, mortality rates of selected religious groups are compared with those of the general population, and lower mortality rates have been found for members of behaviorally strict denominations, such as Mormons and Seventh Day Adventists. These groups prescribe such health practices as not smoking cigarettes or drinking alcohol. More generally, the emphasis placed by many religions on respect for one's body and moderation in behavior implies that adherents will be more likely to adopt good health practices. Social and psychological factors may also be important: a recent study attributed the considerably lower mortality in religious kibbutzim compared with secular kibbutzim in Israel to a social environment that caused less stress, enhanced host resistance, and improved well-being.

Over the past 2 decades, increased interest has been shown in measuring religiosity by the frequency of attending services (usually dichotomized as once a week or more vs. less) rather than by affiliation. Frequent attendance was associated with lower mortality for females (but not males) in the Tecumseh Community Health Study, lower 2-year mortality rates in a sample of elderly poor, and lower cause-specific mortality rates for arteriosclerotic heart disease, suicide, cirrhosis of the liver, and emphysema among a variety of groups. More frequent religious attendance has been associated with less depressive symptomatology, lower blood pressure, better perceived health, and higher life satisfaction.

For Mormon women, higher church activity was associated with lower rates of lung cancer. More broadly, higher activity has been associated with lower mortality rates among the elderly. Disabled persons may be less likely to attend religious services, yet health status is rarely assessed. The observed lower mortality rate for frequent attenders in one study disappeared after 6 years of follow-up, suggesting that better health status at baseline might have been a confounder. Other potential confounders include social connections and such health practices as exercising and not smoking. If persons with good health practices and stronger social connections also attend religious services, observed differences in mortality between attenders and nonattendees might be explained by the healthier practices and stronger connections of the attenders.

However, it is also possible that health practices and social connections are intervening variables that lie on the causal pathway between attendance and mortality. Frequent attendance may facilitate the adoption of better health practices and stronger social connections over time. Examining the timing of

William J. Strawbridge, Richard D. Cohen, and Sarah J. Shema are with the Human Population Laboratory, California Public Health Foundation, Berkeley, Calif. George A. Kaplan is with the Human Population Laboratory, California Department of Health Services, Berkeley, Calif.

Requests for reprints should be sent to William J. Strawbridge, PhD, Human Population Laboratory, 2151 Berkeley Way, Annex 2, Suite 300, Berkeley, CA 94704-1011.

This paper was accepted August 14, 1996.
adopting better health practices and strengthening social connections in relation to attendance would help clarify the causal pathways involved.

The analyses reported here address previous methodologic difficulties by examining the association between frequent religious attendance and mortality over 28 years and by including likely confounders. Changes in health, health practices, and social connections are assessed to determine whether these factors explain any observed association.

Methods

Study Population

A longitudinal study of health and mortality, the Alameda County [California] Study has followed 6928 persons aged 16 to 94 at baseline since 1965.22,23 Survivors were resurveyed in 1974, 1983, and 1994. Response rates for the four surveys starting in 1965 were 86%, 85%, 87%, and 93%, respectively.

Mortality analyses (n = 5286) are based on respondents aged 21 through 65 at baseline in 1965 who had no missing values on any of the variables used in the statistical models. Mean age was 39.8; 52.8% were female, and 12.7% were Black. Analyses for improvements in health practices, body mass index, and social contacts are based on 2540 survivors who responded to the 1994 questionnaire; their mean age was 65.3 (range 50 through 94).

Measures

Measures were selected from those that were included in all four surveys and that had demonstrated associations with mortality.

Attendance. Frequent attendees were defined as those who went to religious services once a week or more and constituted 25.1% of the sample in 1965. For infrequent attendees in 1965, 31% said they never attended, 38% went once or twice a year, and 31% went once or twice a month. The dichotomous split between frequent and infrequent attendees was relatively stable over time: among survivors, 58% of 1965 frequent attendees were still frequent attendees in 1994, and 86% of 1965 infrequent attendees were still infrequent attendees in 1994.

Adjustment variables. Because frequency of attendance was higher for some religious groups than others, religious affiliation was included as a control and coded as Protestant, Catholic, Fundamentalist, Seventh Day Adventist/Mormon, and others/none. Small numbers for some groups precluded more detailed coding. Sociodemographic variables included age, gender, ethnicity, and education. Health variables included mobility impairment (trouble climbing stairs or going outdoors), perceived health (good or excellent vs fair or poor), depression (score of 5 or more on the scale developed by Roberts and O'Keefe),24 and a count of the presence in the past 12 months of diabetes, cancer, stroke, heart disease, bronchitis, or high blood pressure.

Health practices and conditions. The ones assessed included cigarette smoking (current, former, or never), physical exercise (often, sometimes, or never walk, swim, do physical exercise, or do sports for exercise), and alcohol consumption (abstain, 1 to 45 drinks per month, over 45 drinks per month). The body mass index (weight in kilograms divided by height in meters squared) was used to divide subjects into weight quintiles by gender: those in the upper quintile were considered overweight; those in the lower quintile were considered underweight. All of these practices and conditions have been shown to predict both mortality and morbidity.22,25-28

Social connections. These included three from the Social Network Index: marital status (married vs not married), close social contacts (seeing three or more close friends or relatives at least once a month vs fewer than three), and group memberships (belonging to three or more nonreligiously associated groups vs fewer than three). This index and its components have been shown to predict mortality and physical functioning.29-33

Improved health practices, increased social contacts, and stable marriages from 1965 to 1994. Improved health practices included quitting smoking by 1994 for those smoking in 1965; increased exercise by 1994 for those exercising never or only sometimes in 1965; being no longer in the top body mass index quintile in 1994 for those in the top body mass index quintile in 1965; and reduced alcohol consumption for those having more than 45 drinks per month in 1965. Increased social contacts included increased numbers of close friends and relatives seen each month for those seeing fewer than three in 1965 and increased nonreligious group memberships for those belonging to fewer than three in 1965. Marital stability was measured in terms of whether 1994 respondents were married to the same person they had been married to in 1965.

Statistical Analyses

Baseline associations between frequent attendance and adjustment variables were analyzed with percentages and unadjusted odds ratios.

Cox proportional hazards models with time-dependent covariates were used to analyze the relationship between attendance and mortality.34,35 This method takes into account changes in attendance and adjustment variables reported at each new survey during follow-up. Survival times were censored at loss to follow-up or at end of study. Deaths were included through 1993 and numbered 770. Four sequential models were used to assess the relative impacts of the adjustment variables on the relationship between attendance and mortality. The first model included age, gender, ethnicity, education, and affiliation as adjustments. The second model added health conditions; the third added social connections; and the fourth added health practices. Gender differences in outcomes and gender-specific associations between attendance and mortality were assessed by adding gender-by-attendance interaction terms to the models.

Multiple logistic regression was used to assess associations between attendance and 1965-through-1994 changes in health practices, body mass index, and social connections. Adjustments included age, gender, ethnicity, religious affiliation, education, and health conditions.

Results

Baseline Associations and Mortality

Table 1 presents associations between baseline characteristics and frequent attendance. Females and Blacks were more likely to attend frequently, as were those who were mobility impaired and not depressed. Marriage was not associated with frequent attendance, but both number of close social contacts and group memberships were. While there was no association for exercise, smokers and heavy drinkers were much less likely to be frequent attendees. Overweight persons were more likely to be frequent attendees.

Table 2 presents the results of the Cox proportional hazards analyses using time-dependent covariates. Frequent attendees had lower mortality rates than nonattendees when only age, gender, ethnicity, education, and affiliation were
included as adjustments (relative hazard [RH] = 0.64; 95% confidence interval [CI] = 0.53, 0.77). There were modest reductions in this relationship when health conditions (Model II) and social connections (Model III) were added. When adjustments for health practices and body mass index were included (Model IV), the relationship between frequent attendance and mortality became weaker (RH = 0.77; 95% CI = 0.64, 0.93) though still statistically significant.

The data from Table 2 also indicate that the relationship between frequent attendance and mortality was stronger for females than for males. For males, the inclusion of all covariates made the relationship between frequent attendance and mortality no longer statistically significant (RH = 0.90; 95% CI = 0.70, 1.15). The relationship between frequent attendance and mortality remained statistically significant for females even when all covariates were included (RH = 0.66; 95% CI = 0.51, 0.86).

Changes in Health Practices and Social Contacts for Survivors

Table 3 presents the results comparing frequent and infrequent attenders on improvements in health practices, body mass index, and social connections between 1965 and 1994. Frequent attenders who smoked in 1965 were nearly twice as likely as infrequent attenders to stop (odds ratio [OR] = 1.90; 95% CI = 1.27, 2.85). Frequent attenders who exercised never or only sometimes in 1965 were over a third more likely to increase their frequency of exercise (OR = 1.38; 95% CI = 1.08, 1.77). The odds ratios for reducing drinking by 1994 and for no longer being overweight favored frequent attenders, but small numbers made the resulting confidence intervals wide.

Frequent attenders showed greater stability or improvement on the three social measures than infrequent attenders. They were more likely to stay married to the same person (OR = 1.79; 95% CI = 1.36, 2.35); those with few group memberships were more likely to increase memberships (OR = 1.58; 95% CI = 1.21, 2.06); and those with few close contacts in 1965 were 50% more likely to increase their contacts (OR = 1.50; 95% CI = 1.02, 2.21).

Discussion

Using time-dependent covariate survival models, this study demonstrated lower mortality rates over nearly 3 decades for frequent religious attenders compared with infrequent attenders, even with adjustments for mental and physical health during follow-up. Adjusting for social connections had only a modest impact; the association between attendance and mortality was reduced when health practices were added as adjustments, but remained statistically significant.

Several potentially important variables to further assess the pathways by which religiosity might have an impact on health were not available to us. Religiosity may affect health through psychological means, such as an improved sense of coherence, a belief in the therapeutic value of faith, or by a stronger host resistance to the impact of stressors on mental and physical health.36 Quality, rather than quantity, of social relationships may also be important. Such concepts stress the intrinsic aspects of religiosity as opposed to the extrinsic or organizational aspects measured in our analysis.37 Using such measures with our methodology might prove fruitful.

Gender Differences

The observed associations between frequent attendance and mortality were stronger for females than for males, a finding consistent with results in the Tecumseh Community Health Study.3 Past studies of American religiosity have reported that women attend services more frequently than men and evidence a
TABLE 2—Multivariate Sequential Models Using Time-Dependent Covariates to Compare Mortality Rates for Frequent vs Infrequent Attenders of Religious Services during 28 Years of Follow-Up

<table>
<thead>
<tr>
<th>Model and Covariates</th>
<th>Total Sample (n = 5290)</th>
<th>Females* (n = 2789)</th>
<th>Males* (n = 2479)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RH 95% CI</td>
<td>RH 95% CI</td>
<td>RH 95% CI</td>
</tr>
<tr>
<td>I: Age, gender, ethnicity, education, religious group</td>
<td>0.64 0.53, 0.77</td>
<td>0.54 0.42, 0.70</td>
<td>0.76 0.60, 0.97</td>
</tr>
<tr>
<td>II: Model I plus health conditions</td>
<td>0.67 0.56, 0.80</td>
<td>0.56 0.43, 0.73</td>
<td>0.80 0.62, 1.02</td>
</tr>
<tr>
<td>III: Model II plus social connections</td>
<td>0.69 0.57, 0.83</td>
<td>0.57 0.44, 0.74</td>
<td>0.82 0.64, 1.05</td>
</tr>
<tr>
<td>IV: Model III plus health practices, body mass index</td>
<td>0.77 0.64, 0.93</td>
<td>0.66 0.51, 0.86</td>
<td>0.90 0.70, 1.15</td>
</tr>
</tbody>
</table>

Note: RH = relative hazard; CI = confidence interval.
*RH and CI for males and females estimated from a single model containing gender by attendance interaction term; P values for this interaction term were .05 (Models I and II), .04 (Model III), and .08 (Model IV).

TABLE 3—Improved Health Practices, Social Connections, and Stable Marriages over 29 Years for Frequent Attenders of Religious Services in 1965 Compared with Infrequent Attenders

<table>
<thead>
<tr>
<th></th>
<th>Frequent Attenders</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No.</td>
</tr>
<tr>
<td>Improved health practices/conditions</td>
<td></td>
</tr>
<tr>
<td>Stopped smoking (for those smoking in 1965)</td>
<td>1061</td>
</tr>
<tr>
<td>Reduced drinking ≤45 drinks/mo (for those &gt;45 drinks/mo in 1965)</td>
<td>350</td>
</tr>
<tr>
<td>Increased exercise (for those exercising never or only sometimes in 1965)</td>
<td>1486</td>
</tr>
<tr>
<td>No longer in top body mass index quintile (for those in top quintile in 1965)</td>
<td>355</td>
</tr>
<tr>
<td>Stable marriages, increased social connections</td>
<td></td>
</tr>
<tr>
<td>Stayed married to same person (for those married in 1965)</td>
<td>2007</td>
</tr>
<tr>
<td>Increased nonchurch community group memberships to 3+ or more (for those belonging to &lt;3 in 1965)</td>
<td>1701</td>
</tr>
<tr>
<td>Increased number of close friends/relatives seen each month to 3+ (for those seeing &lt;3 in 1965)</td>
<td>779</td>
</tr>
</tbody>
</table>

Note. Odds ratios (ORs) and confidence intervals (CIs) are based on logistic regression models comparing those who attended religious services once a week or more in 1965 with those who attended less often or not at all. Models adjust for 1965 age, gender, ethnicity, education, religious affiliation, chronic conditions, mobility impairment, perceived health, and depression.

stronger religious commitment. More salient to the analyses here is that more women report using religion as a coping mechanism for dealing with life stress. There is also evidence that religious involvement has stronger protective associations for disability and depression among older women than among older men. Given the much higher proportion of widowhood among older women, religious organizations may act to fill an otherwise unmet social-support need. Koenig suggests that religiosity is more important for women because of their lower social status; Idler suggests that such differences stem from the more general tendency of women to seek and use social interaction to cope with illness. Further analyses of this gender difference could provide etiologic insights to better explain the relationship between religiosity and mortality.

Are Health Practices and Social Connections Confounding or Intervening Variables?

Health practices and social connections could either confound the relationship between attendance and mortality (persons with good health practices and stronger social connections are frequent attenders of religious services) or act as intervening variables on a causal pathway between attendance and mortality. We found support for both possibilities but somewhat stronger evidence for the intervening model. At baseline, frequent attenders were much less likely than infrequent attenders to smoke cigarettes or drink heavily, and among those frequent attenders who smoked or drank heavily at baseline, more were likely to stop these activities during follow-up. Both baseline differences and differential improvements during follow-up were also found for the two social-connections measures. For exercise and marital status, there were no baseline differences, yet frequent attenders who exercised little at baseline were more likely to increase, and those married more likely to stay married to the same person. Frequent attenders may have lower mortality rates because they adopt better health practices, increase their social connections, and have more stable marriages in conjunction with their religious attendance. Attendance at religious services could also have influenced health practices and increased social contacts before the study began.

The adjustment for health practices reduced the relationship between frequent attendance and mortality more than did the adjustment for number of social connections, although it is possible that the latter relationship would have been stronger if information had been available on the quality of social connections. The stronger impact of health practices in our analyses is consistent with cross-sectional results from the Yale Health and Aging Project, which used health rather than mortality as an outcome. On the other hand, Kark et al. reported that health practices were not responsible for the sharply lower mortality rate in religious vs secular Israeli kibbutzim. Differences in relative importance between health practices and social ties could reflect differen-

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
tial measurement, the comparison groups used, or a real difference in impact. This issue should be researched further.

Public Health Implications

If, as our analyses indicate, frequent religious attendance does facilitate the adoption of better health practices, the mechanisms involved have broad public health implications. Religious organizations are frequently involved in public health campaigns and supportive programs to assist marginalized members of their communities. The American Public Health Association has set up an initiative to form new partnerships with faith communities to better coordinate such activities. The initiative stresses the roles of religious organizations in communities as well as impacts on their members. Understanding how these organizations affect behavior and attitudes of their own members could help us understand why some individuals adopt good health practices while others do not. Possible mechanisms by which such organizations influence good health practices include peer influence, increased self-esteem, increased sense of perceived control, prescribed practices, and a general philosophical outlook that values social ties and treating one's body with respect. Further research is needed to identify the specific mechanisms involved and determine whether they are limited primarily to active members of these organizations or extend more broadly into the communities around them.

Acknowledgment

This study was supported by grant R37AG11375 from the National Institute on Aging.

References


American Journal of Public Health 961

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Routine Prenatal Screening for Congenital Heart Disease: What Can Be Expected? A Decision-Analytic Approach

Abstract

Objectives. This study assessed the potential impact of fetal ultrasound screening on the number of newborns affected by cardiac anomalies.

Methods. A decision model was developed that included the prevalence and history of congenital heart disease, characteristics of ultrasound, risk of abortion, and attitude toward pregnancy termination. Probabilities were obtained with a literature survey; sensitivity analysis showed their influence on expected outcomes.

Results. Presently, screening programs may prevent the birth of approximately 1300 severely affected newborns per million second-trimester pregnancies. However, over 2000 terminations of pregnancy would be required, 750 of which would have ended in intrapartum death or spontaneous abortion. Further, 9000 false-positive screening results would occur, requiring referral. Only the sensitivity of routine screening and attitude toward termination of pregnancy appeared to influence the yield substantially.

Conclusions. The impact of routine screening for congenital heart disease appeared relatively small. Further data may be required to fully assess the utility of prenatal screening. (Am J Public Health. 1997;87: 962–967)

Erik Buskens, MD, PhD, Ewout W. Steyerberg, PhD, John Hess, MD, PhD, Jury W. Wladimiroff, MD, PhD, and Diederick E. Grobbee, MD, PhD

Introduction

Over the last decade many reports have appeared on the possibilities of ultrasound for prenatal detection of congenital heart disease, but few authors have tried to assess the actual yield of a screening program for cardiac anomalies. It is important to distinguish between a limited screening procedure offered routinely to all pregnant women (population approach) and extensive fetal echocardiography offered to those at (high) risk for fetal congenital heart disease.

In the population approach, all pregnant women undergo routine ultrasound examination at a certain optimal gestational age. When a congenital anomaly is suspected, referral to establish a diagnosis and appropriate obstetric policy follows. The screening procedure currently advocated for routine evaluation of the fetal heart is the four-chamber view at 16 to 24 weeks of pregnancy.

The high-risk approach is generally accepted and offers extensive fetal echocardiography to selected women, in particular to those women with a history of congenital heart disease in their offspring or those who appear during routine screening to be carrying an affected fetus. Anomalies encountered in the high-risk group tend to be more serious and complex. In addition to the four-chamber view, the cardiac connections and functional status are evaluated. Obviously, this can only be accomplished by skilled experts during a lengthy and detailed examination.

To justify routine prenatal screening in low-risk pregnancies with subsequent extensive ultrasound examination in case of suspected fetal pathology, an assessment of the efficiency of such a program is needed. Presently, however, a favorable effect of routine fetal ultrasound including a four-chamber view evaluation is assumed. Routine fetal ultrasound is now offered to the majority of pregnant women in several countries, including the Netherlands. To our knowledge, this policy has not been preceded by an appropriate evaluation. Medical decision analysis offers a possibility for integrating and analyzing the influence of the efficacy of screening, the risks for the affected fetus, and societal or parental attitudes on the expected distribution of outcomes of pregnancy in a low-risk population.

We set out to assess whether the advantages of prenatal detection of cardiac anomalies by means of ultrasound examination are sufficiently clear to merit screening of pregnant women at low risk for congenital heart disease in their offspring.

Erik Buskens and Diederick E. Grobbee were with the Department of Epidemiology and Biostatistics; Ewout W. Steyerberg is with the Department of Public Health; John Hess is with the Division of Pediatrie Cardiology, Sophia Children's Hospital; and Jury W. Wladimiroff is with the Department of Obstetrics and Gynaecology, all at the University Hospital Rotterdam Dijkzigt, Erasmus University Medical School, Rotterdam, the Netherlands. Erik Buskens and Diederick E. Grobbee are now with the Julius Center for Patient-Oriented Research, Utrecht University Medical School, Utrecht, the Netherlands.

Requests for reprints should be sent to Erik Buskens, MD, PhD, Julius Center for Patient Oriented Research, Utrecht University, PO Box 80035, 3508 TA Utrecht, the Netherlands.

This paper was accepted September 13, 1996.
Methods

Structure of the Model

The present report proceeds from an inventory of data available in the literature. These data have been introduced into a decision analysis model centered on the problem of whether to offer routine fetal echocardiography to pregnant women at low risk. Decision Maker software (New England Medical Center, Boston, Mass, 1988) was used to structure the model. The options and chances future parents, clinicians, and policymakers face at 16 to 24 weeks' gestational age are conveyed (Figure 1). The high-risk approach as a separate option was not included in the model, as its merits have been established. 6,8,17-20 In addition, as spontaneous abortion or intrauterine death prior to a prenatal diagnostic procedure is not amenable to intervention, this possibility was not evaluated. The probability of an affected fetus (with minor or major congenital heart disease) was considered in the model at the first chance node. The subsequent chance node in the model represents routine fetal echocardiography. Given a malformation in the fetus, the chance of a positive test result is the sensitivity. Similarly, the chance of a negative test given the absence of malformations is the specificity.

The subsequent branch of the model represents referral on suspicion of an anomaly at initial screening. Extensive fetal echocardiography is offered to those screened and found to have a suspected fetal anomaly. Also, those presenting pregnancy pathology—for example, growth discrepancy, lack of fetal movements, or abnormal fetal heart rate—are referred. For the model we assumed that extensive fetal echocardiography would in general reveal previous false-positive diagnostic errors (100% specificity). This appears to be a reasonable simplification of the model, since further diagnostic tests or termination of pregnancy are not offered unless a serious fetal anomaly is indeed suspected. Nonadoption of screening implies that extended ultrasound is available only in the event of clinically suspected fetal or pregnancy pathology. Similar to the test characteristics of routine screening, the test characteristics of extensive fetal echocardiography have been applied in the model.

In case an anomaly is confirmed at extensive fetal ultrasound examination, chorion villus sampling, amniocentesis, or cordocentesis is offered. These techniques have a low risk of induced abortion. This risk is represented by the corresponding chance node in the model.

The next step is the decision parents face when confronted with the diagnosis and prognosis of their fetus. They may choose to terminate the pregnancy or carry to term. Obviously, an unaffected fetus or a fetus with minor anomalies is likely to be carried to term. In case the gestation of a fetus with congenital heart disease is continued, two outcomes are possible: intrauterine death or a live infant with a cardiac anomaly. As a result of the anomaly the infant may die postnatally. The risk of a fetal outcome is again represented by a chance node. The situation is essentially similar if parents informed of the presence of a fetal anomaly decide not to terminate the pregnancy, if an anomaly is not detected, or if screening is not offered.

Assignment of Probabilities

All variables used in the model are summarized in Table 1. A number of problems in the assignment of probabilities need to be discussed. An estimate of the chance of a fetus with congenital heart disease is preferably based on the prevalence of cardiac anomalies at 16 to 24 weeks' gestation. However, while the prevalence of congenital heart disease at birth is well documented, reliable estimates at about 20 weeks' gestational age are sparse. We assumed that newborns with a birth prevalence of cardiac anomalies of approximately 0.008 originating from a larger cohort of fetuses of which a proportion aborted spontaneously or ended in premature death. In addition, it should be noted that only about half of the cardiac anomalies found in neonates are major anomalies. 32 Assuming that approximately 37% of the major anomalies end in premature death, 19,20 the number affected at 20 weeks' gestation was calculated at 10.3 per 1000 (0.004 + 0.004/1/1-0.37).

The test characteristics of routine and extensive ultrasound examination used in the analysis are based on a literature survey. Some studies reported
TABLE 1—Estimates of the Various Probabilities Applied in the Decision Model on Routine Fetal Ultrasound Examination for Congenital Heart Disease (CHD)

<table>
<thead>
<tr>
<th></th>
<th>Point Estimate</th>
<th>Lower Value</th>
<th>Upper Value</th>
<th>Major Anomalies</th>
<th>Minor Anomalies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prevalence(^{21,22})</td>
<td>0.008</td>
<td>0.003</td>
<td>0.012</td>
<td>0.004</td>
<td>0.004</td>
</tr>
<tr>
<td>a. Prenatal prevalence(^a)</td>
<td>0.0103</td>
<td>...</td>
<td>0.0003</td>
<td>0.004</td>
<td>...</td>
</tr>
<tr>
<td>b. Sensitivity routine screening(^{1-15})</td>
<td>0.07</td>
<td>0.07</td>
<td>0.50</td>
<td>(0.20–0.80)</td>
<td>0.07 (0.01–0.20)</td>
</tr>
<tr>
<td>c. Specificity routine screening(^{1-15})</td>
<td>0.99</td>
<td>0.99</td>
<td>1.0</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>d. Sensitivity routine ultrasound(^{6,17-19})</td>
<td>0.95</td>
<td>0.927</td>
<td>0.974</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>e. Specificity routine ultrasound(^{6,17-19})</td>
<td>0.99</td>
<td>0.987</td>
<td>0.993</td>
<td>1.00</td>
<td>...</td>
</tr>
<tr>
<td>f. Probability abnormal pregnancy given CHD</td>
<td>0.01</td>
<td>...</td>
<td>0.01</td>
<td>(0–0.03)</td>
<td>0.01 (0–0.03)</td>
</tr>
<tr>
<td>g. Probability karyotyping induced abortion(^3)</td>
<td>0.678</td>
<td>0.571</td>
<td>0.773</td>
<td>0.88</td>
<td>0.6</td>
</tr>
<tr>
<td>h. Probability termination of pregnancy given CHD(^{19,20})</td>
<td>0.370</td>
<td>0.194</td>
<td>0.576</td>
<td>0.37 (0.20–0.60)</td>
<td>0.03</td>
</tr>
<tr>
<td>i. Probability intrauterine death given CHD(^{19,20})</td>
<td>0.588</td>
<td>0.329</td>
<td>0.816</td>
<td>0.59</td>
<td>0.1</td>
</tr>
</tbody>
</table>

**Note.** In the analysis a distinction is made between major and minor anomalies. The plausible range used in sensitivity analysis is given in parentheses. Major and minor are categories of congenital malformation according to which sensitivity and clinical course is varied in the model. Ellipsis points indicate estimates of variables or ranges of variables not examined in the current analysis. The letters a through f refer to corresponding probabilities in Figure 1.

*Expected prevalence at 20 weeks' gestation calculated; varies according to probability of intrauterine death given CHD.

In the model a specificity of extensive fetal echocardiography of 100% is applied.

on a high-risk population and a more extensive screening procedure, whereas others reported on a low-risk population and a simple screening procedure performed once during pregnancy. This resulted in a wide range of published results.\(^{1-15}\) To account for this variability we assumed the test characteristics, especially the sensitivity, to vary according to case severity. Accordingly, the upper range of sensitivity reported (around 50%) was taken to apply to serious cases, whereas the lower range of literature estimates (below 10%) was taken to apply to minor anomalies. The specificity of routine screening is reported to be very high (99%) and has a narrow range. The sensitivity of extensive fetal ultrasound evaluation appears to be much better and has been reported at around 95% with a narrow range.\(^{6,17-20}\) Also, a near 100% specificity of extensive fetal ultrasound examination has been reported.\(^{6,8,17-20}\) Detailed data on extensive fetal echocardiography enabled us to calculate the test characteristics with the specific indication of suspected fetal (cardiac) pathology.\(^{6,19,20}\) The likelihood of abortion as a complication of fetal karyotyping is low (less than 1%).\(^{23}\) Data on the probability of pregnancy pathology in relation to congenital heart disease could not be found. A panel of obstetricians at Rotterdam University Hospital estimated the probability at less than 1%. In case a malformation is detected parents have to decide whether to terminate the pregnancy. Specific literature on this subject is scarce. Pryde et al. evaluated several factors influencing parental decisions regarding pregnancy outcome of congenitally malformed offspring.\(^{24}\) The prognosis appeared to be of major importance: two out of three couples opted to terminate pregnancy if a major anomaly was detected. Termination was never opted for in case of minor anomalies. Similar results have been reported in relation to extensive fetal echocardiography.\(^{6,19,20,24}\)

With regard to the prognosis of affected fetuses, a distinction between serious and mild cases was also made. Cases detected prenatally are likely to be a subsample with severe anomalies. Accordingly, the outcomes of pregnancy reported in case-series on cases detected prenatally are taken to apply to serious cases. Intrauterine death occurred in 37% of such cases and infant death in 59%.\(^{19,20}\) On the other hand, children with congenital heart disease born alive may represent a subsample with relatively mild anomalies. Live-born children with congenital heart disease have a mortality of less than 10%. This is the estimate of survival used in the model in case of minor cardiac malformations. In addition, we assumed that the majority of minor anomalies do not cause hemodynamic problems prenatally. Accordingly, the fetuses survive to term. Finally, a sensitivity analysis was conducted over plausible ranges of the probabilities (as presented in the right-hand side of Table 1) to assess the influence of variability in the estimates on the outcome of the model.

**Results**

The impact of routine screening, in numbers per million second-trimester pregnancies, is given in Table 2. With an assumed low sensitivity of 50% for major anomalies, it is estimated that the number of children born with severe congenital heart disease decreases by a third. A similar effect on the number of cases of intrauterine death and neonatal death is observed. Some of the intrauterine deaths and neonatal deaths that would otherwise have occurred are avoided if pregnancy is terminated in cases detected prenatally.

The impact of routine screening on the number of children born with minor congenital heart disease is negligible. The number of terminations of pregnancy would, however, increase 50-fold. Also, approximately 9900 false-positive screening tests would result. Moreover, screening would lead to a loss of 32 fetuses. 28 with major anomalies and 4 with only minor anomalies, owing to karyotyping. As we presumed a specificity of extensive fetal echocardiography of 100%, karyotyping is not offered to unaffected fetuses. An increase in sensitivity (to 80% and 20% for major and minor anomalies, respectively) will reduce the number of cases with an unfavorable outcome (from 525 fewer to 841 fewer for births of infants with serious anomalies and from 757 fewer to 1211 fewer for neonatal deaths).

For minor anomalies the change is negligible. Obviously, more pregnancies will have to be terminated to achieve this
reduction (from 2005 to 3208). In addition, as a result of a higher sensitivity, fetal karyotyping is performed more often, causing an additional loss of 23 (affected) fetuses (increasing the number lost from 32 to 55). Also, a sizable number of the cases that would result in intrauterine death or spontaneous abortion are now terminated (from 752 fewer to 1205 fewer). If sensitivity decreases, only a marginal effect remains. However, an identical number of women would have to be referred and go through an emotionally difficult period owing to a false-positive screening test.

Parental inclination toward termination of pregnancy in case of severe malfunction appears to have an effect that is numerically comparable to increased sensitivity of routine ultrasound. Obviously, with an increased proportion of the parents opting for termination of pregnancy in case of a major anomaly, fewer affected neonates are born. If all pregnancies are carried to term very little effect remains.

Figure 2 shows the results of a two-way sensitivity analysis of the two major determinants of the impact of screening. The effect of simultaneously varying estimates of the sensitivity of routine fetal ultrasound and estimates of the proportion of parents opting for termination of pregnancy, both in case of a severely affected fetus (i.e., cases resulting in neonatal survival with severe congenital heart disease or cases ending in neonatal death), is demonstrated. The number of newborns with serious anomalies prevented increases with increasing sensitivity and with an increasing probability of termination of pregnancy. Sensitivity and probability of termination show a combined (multiplicative) effect.

Variation in the probability of abnormal development of pregnancy or pregnancy pathology in affected fetuses does not appear to have any significant impact. Neither does variation in the probability of intrauterine death. We assumed the birth prevalence of congenital heart disease to remain stable. Accordingly, an increased probability of premature death implies that a larger number of pregnancies are terminated that would otherwise result in intrauterine death. Also, a larger number of affected fetuses would be lost owing to karyotyping. (Additional details and results are obtainable from the authors.)

<table>
<thead>
<tr>
<th>Outcome</th>
<th>No Screening</th>
<th>Screening</th>
<th>Difference</th>
<th>% Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Major congenital heart disease</td>
<td>1 629</td>
<td>1 104</td>
<td>−525</td>
<td>−32</td>
</tr>
<tr>
<td>Minor congenital heart disease</td>
<td>3 800</td>
<td>3 587</td>
<td>−23</td>
<td>−0.1</td>
</tr>
<tr>
<td>Neonatal death</td>
<td>2 745</td>
<td>1 988</td>
<td>−757</td>
<td>−28</td>
</tr>
<tr>
<td>Termination of pregnancy</td>
<td>41</td>
<td>2 046</td>
<td>2 005</td>
<td>49b</td>
</tr>
<tr>
<td>Karyotyping-induced abortion</td>
<td>1</td>
<td>33</td>
<td>32</td>
<td>32b</td>
</tr>
<tr>
<td>Intrauterine death</td>
<td>2 334</td>
<td>1 582</td>
<td>−752</td>
<td>−32</td>
</tr>
<tr>
<td>No congenital heart disease</td>
<td>0</td>
<td>9 897</td>
<td>9 897</td>
<td>=</td>
</tr>
<tr>
<td>(false-positive)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No congenital heart disease</td>
<td>989 651</td>
<td>979 754</td>
<td>−997</td>
<td>−1</td>
</tr>
</tbody>
</table>

#Percentage change relative to no screening.
#As the percentage change would exceed 100, a multiplier is presented.

Figure 2—Results of the two-way sensitivity analysis for the sensitivity of routine screening for major anomalies and the probability of termination of pregnancy (impact of screening vs no screening on number of newborns with serious anomalies prevented).

Discussion

We have shown that routine fetal ultrasound screening does not meet the generally held expectations. The fundamental idea of screening is that parents, neonates, and society in general may benefit from effective screening. One of the initial assumptions in the model was that approximately half of all severe anomalies would be detected prenatally. On prenatal detection of an anomaly, the obstetric policy may be adapted. If, for instance, a case of Fallot's tetralogy with severe pulmonary stenosis were detected prenatally, soon after birth the diminished pulmonary flow could result in a life-threatening situation. Birth should preferably take place in a setting able to provide pediatric intensive care. With prior knowledge, timing, mode, and location of delivery can be optimized to improve chances of neonatal survival. In cases with a (near) fatal prognosis, termination of pregnancy presently is the main alternative.

According to the model, the birth of approximately 30% of neonates with serious congenital heart disease may be prevented. However, as a consequence 49 times as many pregnancies have to be terminated, of which again 30% would have ended in spontaneous abortion. Also, some affected fetuses will be lost in any case owing to karyotyping. The hypothetical yield of routine fetal ultrasound
appears to be relatively small. Moreover, in view of recently reported realistic estimates of the sensitivity of routine fetal ultrasound, only a few cases with an unfavorable outcome of pregnancy may actually be prevented.28 Even if optimal estimates of sensitivity apply, the impact of prenatal screening for congenital heart disease on the health of neonates in general appears limited. This conclusion is in agreement with that of Nelson et al., who reported that a large proportion of fetal anomalies are not likely to be diagnosed, especially in a nonreferral setting.29 The nature of the anomalies encountered prenatally predicts, to a large extent, the possibilities for early detection. Moreover, routine fetal ultrasound has additional disadvantages. Even a false-positive rate of only 1% results in a large number of parents being told that their future baby may have congenital heart disease. If the specificity in reality is 1% lower, doubling the false-positive rate, twice as many such cases occur. Apart from the anxiety caused by such news we must consider the implications for the referral institutions. All of these women will subsequently be referred for extensive ultrasound evaluation. The investments in personnel, appliances, and clinics required are sizable and revenues are small.

A second variable that was shown to affect the attainable yield was the opinion future parents (or society in general) hold on termination of pregnancy. The proportion of pregnancies terminated on account of serious congenital heart disease will depend strongly on such an opinion or preference. In other words, the impact of screening will vary with local legal and ethical standards and attitudes.

We are aware of the fact that the basis for some of the other variables examined was uncertain. However, the impact of variability of these estimates on the expected distribution of outcomes of pregnancy in a low-risk population appeared to be limited. Moreover, the test characteristics of extensive structural ultrasound can already be considered more or less optimal, as are the complication rates of fetal karyotyping. Also, the actual prevalence, the natural history, and factors affecting the natural history of congenital malformations are hardly accessible to intervention. If they were accessible, this would imply possibilities for primary prevention.

Finally, an additional remark should be made regarding the specificity of extensive fetal echocardiography. The specificity applied in the model was presumed to be 100%. Therefore, far-reaching consequences of false-positive routine screening tests are absent in the model and in the results presented. This may not be quite correct. Recently, cases have been described of fetuses with apparently severe congenital heart disease that proved to have only mild to moderate anomalies postnatally.30 A false-positive rate of extensive fetal echocardiography of 1% (specificity 99%) would result in approximately 100 such cases (1% of 9897) in the hypothetical cohort. Future parents erroneously presented with a serious prognosis who opted to terminate pregnancy would lose a normal fetus.

Overall, the yield of prenatal screening for congenital heart disease by means of the fetal four-chamber view, expressed as the prevention of the birth of a critically ill neonate, appears to be numerically small. With substantial effort the efficacy of routine screening may be improved. However, the results would still be modest. Moreover, the final decision parents make once a serious fetal anomaly is detected is culturally, socially, and economically determined. Evidently, a generally applicable protocol for termination of pregnancy is unrealistic.

The significance and valuation of the cases detected, cases not detected, and false-positive test results and their subsequent outcomes have not been assessed in the present analysis. Such an assessment may be attempted by estimating the psychological relief or burden perceived by the parents. Additionally, the costs of screening and postnatal costs could be weighed against the effects. Heckering and Verp31 and Pauker and Pauker31,32 assigned specific values or utilities to the various outcomes of pregnancy in a prenatal Down's syndrome screening program. Subsequently the utility of screening was estimated. Ekwo et al. also reported on the outcome of pregnancy with regard to congenital anomalies and the perceived consequences or burden.33 The parents' opinion on the outcomes of pregnancy could be expressed on a preference scale. At present, reliable data on parental attitudes toward congenital heart disease are lacking, as are data on the costs of congenital heart disease. Yet parental assessment of the various outcomes may have a substantial impact on the appreciation and efficiency of a prenatal screening program for congenital anomalies. After all, the parents decide whether or not to have prenatal screening and how to respond to the outcome of the test.

In conclusion, routine fetal ultrasound screening for congenital heart disease does not seem warranted at present. Public health, in particular neonatal health, is not likely to improve if prenatal screening is offered in low-risk pregnancies. The public expenditures involved could more effectively be spent otherwise.

References

13. Sharland GK, Allan LD. Screening for congenital heart disease prenatally. Results of a 21/2-year study in the South East.
14. Tegnander E, Eik-Nes SH, Linker D. Four-chamber view of the fetal heart—can more heart defects be detected on the routine scan? Presented at the Sixth World Congress on Ultrasound, World Federation for Ultrasound in Medicine and Biology, Copenhagen, Denmark; 1991.


