

***MATERNAL RISK***

by  
**Beverly Winikoff, MD, MPH**

**The Population Council**

**Paper presented at  
Berzelius Symposium  
Stockholm, Sweden**

**May, 1991**

Copyright 1991

The Robert H. Ebert Program on Critical Issues in Reproductive Health  
The Population Council, One Dag Hammarskjold Plaza, New York, NY 10017 USA Telephone: (212) 339-0500 Facsimile: 212-755-6052

## MATERNAL RISK

Risk is a part of life; it has to do with chance and taking chances. In a technical and also in a practical sense it is a statement of probability. Probabilities also can be compared to each other, allowing us to make judgement, if we like, of risks that are more or less than other risks. Some things that we find particularly frightening may be labelled high or extremely high risk.

For example there are risks involved in seemingly innocuous activities, such as drinking diet soda or eating peanut butter. (SLIDE 1) Now, these risks both are rather remote, but if we compare them to each other we could say that eating peanut butter is a "high risk activity." Other activities can also be compared to each other, some of which can be seen as very high risk and other as low risk but only, of course, in comparison to some standard.

The risk of canoeing, for example, seems to be a low risk activity as compared to auto racing or cigarette smoking or even rock climbing. (SLIDE 2) On the other hand, there are many individuals who engage in these riskier activities, and for them, the balance of risks and benefits is in favor of continuing with the activity.

The concept of risk and "maternal risk" is one that has been used in maternal health care. I interpret maternal risk in the title here to be the probability or chance of dying or being seriously injured in pregnancy. *The concept of risk has been extremely useful as a tool for research and a didactic tool in medical and epidemiological education. Today, however, we are exploring its use in health services, and, in this respect, its use is certainly more problematic.*

I will begin by postulating that the concept of risk, as a tool for service delivery design, rests on three foundations. (SLIDE 3) First, that certain groups of women have a higher likelihood or chance of maternal death than other groups. Second, that we can locate these women; and, third, that if we can locate the women we can prevent the deaths. The appeal of this reasoning is, of course, that one can concentrate resources, especially in place where resources are scarce, on the people who are likely to have bad outcomes *and that in doing so we will prevent most of the bad outcomes.*

While the first statement may be true—we can define groups of women who have higher than average chances of poor outcomes—the second two statements are not as clearly true. There is little objective evidence to substantiate items 2 and 3.

The next slide presents a visual schematic picture of this problem. (SLIDE 4)  
This picture is a population of 100 women among whom will occur nine deaths. Let us assume we believe we have found a characteristic that concentrates our ability to locate those women who will die. These women we will call "high risk." We can identify nine such women in the population out of whom three will indeed be "cases" or deaths. So, instead of having a 9% risk of death, the "high risk" women have a 33% risk of death. The other 91 women will have only a 6.6% percent risk of death, and the ratio of risk between the "high risk" and "low risk" women will be about 5. (Incidentally, this is much better than the risk ratio that most good risk predictors will give in a population!) In other words, we can identify women based on some characteristics who will have, on average, about 5 times the risk of poor outcome as other women do.

Let us look at the meaning of that for our real ability to locate and prevent death. If we screen out 9 women, and, in fact, prevent 3 deaths, we will still not have attended to 91 women and will not have prevented 6 other deaths. In other words, we will only have prevented 1/3 of all the deaths, and we will have foregone the possibility of preventing the other 6 because we will not be serving those women at all. Let us presume that we want to expand our base of "high risk" to include more cases. Let us, instead of screening 9 women, screen 36 women. In this case, we are adding 27 more women to our care pool for a benefit of finding 2 more cases. This picture represents the very real problem of our inability to identify well where the cases will be and the problem of rationing care by ignoring a majority of women **and** of poor outcomes.

There are many prerequisites in order for "risk" systems to work to improve services to women, and these are **not** trivial problems. (SLIDE 5, 6)

First, there must be a reliable rating system. This, in itself, is a major problem. We are very poor at discriminating women who will have bad outcomes from those who will not. Second, the system must be able to be taught effectively, then it must be applied correctly. All of this should be done at a reasonable cost to constrained health care systems. Finally, in order to have any effect, the identification of risk must result in adequate referral and/or treatment.

I would like to discuss, for a few moments, some of the intrinsic problems of the development of the risk scoring systems that must be the basis of any attempt to use risk in a clinical or service delivery setting. In order to construct a risk system, one has to have reliable data on the relationships between individual characteristics and the outcomes being studied. This is very difficult to develop with regard to maternal health, because, for most communities, there are very poor data on vital events. In other words, we do not know who died of causes

related to pregnancy and cannot ascertain enough cases to find out whether certain characteristics of the individuals and their environment are related to the poor outcomes. It is even harder to assemble accurate data on morbidity.

A common result is that several rare adverse outcomes are grouped as "poor pregnancy outcomes." This **group** of adverse events is then considered a single outcome. When this is entered as the dependent variable in a regression analysis, the resulting correlations of patient characteristics with poor outcomes do not describe the particular relationship between any particular characteristic and specific problems of pregnancy.

A major problem with the issue of maternal risk has been the complete confusion in most risk scoring systems of **whose** risk is being "scored" or looked at. Since poor maternal outcomes are so rare in a statistical sense, many risk systems have fallen back at looking at **infant** outcomes. In other words, still birth, low birth weight, and neonatal death are much easier to record and are much more common than adverse maternal outcomes. Thus, many risk systems really predict poor infant, not maternal outcomes. While maternal death is very often accompanied by poor infant outcome, it is not always true that poor infant outcome is associated with poor maternal outcome. Thus, predictors of maternal risk that are based on poor infant outcome are often wholly inappropriate as measures of maternal health.

A final problem I have listed as "inclusion of outcome within predictors." In fact, this operates on various levels. One is the use of actual proxies for the outcome in the predictor. In other words, a risk system that uses toxemia as a predictor of hospitalization, for example, is in effect a tautology if the practice is to hospitalize all women with toxemia. This kind of thinking is replete within many risk systems in which variables such as severe bradycardia are used in scoring for fetal demise, for example.

This issue also operates at a more subtle level. That is, predictors that are closer to the final event actually are more strongly correlated with outcome than predictors that one can assess further from the event. In other words, predictors of death in childbirth are much stronger during labor than they are before labor. They are also much stronger at 38 weeks of pregnancy than they are at 12 weeks of pregnancy. *The weakest and worst predictors of all are the predictors based on demographic and socio-economic characteristics, yet these are often important elements of risk systems.* The paradox, of course, is that when one is very close to the event, while prediction is better, there is precious little time or flexibility to intervene.

In addition to these generic problems of constructing risk scoring systems, there

are technical problems of scoring risks. (SLIDE 7)

One has to do with the inevitability of false positives and false negatives. These are linked inversely: if there are fewer false negatives, the system will produce more false positives. Another problem has to do with threshold. In other words, since risk systems are based on only two options, high risk or not high risk, and scales or scoring systems are often continuous, one has to choose one point at which high risk begins and low risk ends. This is obviously artifice and results in a large degree of misclassification.

Many risk scoring systems are based on a compilation of several variables. Frequently, these are simply added together. Sometimes they are multiplied by various weighting factors. But, in fact, there is no real information to assess the weight of different variables. For example, how does the weight of grand multiparity compare to the weight of previous gestational diabetes as a risk factor? Scores, or relative weights of scores, are often assigned on the basis of clinical intuition rather than on any studies. The same is true for interaction of variables, rarely accounted for in any risk scoring system. In other words, it is assumed that any two factors can be added together and that it is not worse to have problems A & B than problems A & E, if the initial "scores" assigned to B and E are equal.

Finally, risk scoring systems are predicting an overall problem, but not any specific difficulty. "Risk" then is actually an agglomeration of very different processes each of which have very different causes and interventions. Thus, a woman scored as "high risk" may be "high risk" because she has high blood pressure, or "high risk" because she has a twin pregnancy or because she is over 45, for example. Yet no specific intervention can be suggested without reference to specific clinical conditions, and these are not apparent from a risk score. Risk, itself, does not suggest any intervention and, therefore, there is no automatic plan of action for women designated as "high risk."

The sum of all these problems means that the application of risk systems is very imperfect. A graphic (SLIDE 8) illustrates this from a study in a U.S. hospital. Most of the "high risk" women do not suffer complications and most of those who suffer complications were rated as "low risk."

The second criterion for effective use of risk as a basis for services is that the system must be able to be taught properly to those who will use it. (SLIDE 9) This training is not simple. Training of clinical personnel to fill out risk scoring forms or in other ways to apply a mathematical scoring system is an intrinsically difficult exercise. Clinicians are not oriented to this kind of activity and often have

little or no interest in it. They think more in terms of individual patients, clinical findings and their sense of how a pregnancy is proceeding. In some way, the routinization of scoring systems is directly counter to the clinical orientations of the professional.

In addition, training takes time. It takes more time than one may first imagine. In a training scheme for risk approach in Mexico among family planning workers, physicians spent three days of time each in learning the system and its justifications. Other personnel spent even more time. Each person had to be trained in the system before it could be instituted. Training costs money. It costs money in dollar terms. Trainers have to be hired and need to be paid for training sessions, and clinical tasks must be absorbed by someone else.

These systems also need to be applied effectively, yet this is rarely the case. (SLIDE 10) In the first place, clinicians are often over burdened with clinical care, and use of the risk scoring system takes time away from clinical encounters with patients. Clinicians who are pressed for time rarely will fill out complicated forms accurately or completely. Finally, of course, the more complicated the risk scoring system or form, the more errors there will be in use and the less accurate it will be, even if clinicians spend time completing it.

The costs are not trivial. I have identified 4 types of costs. (SLIDE 11) In addition to the costs of training, there are opportunity costs involved not only in the training but also in the utilization of the system. In other words, when time is taken from clinical encounters, this time could have been used to see more patients or for counselling patients or for other worthwhile endeavors. We need to be sure that risk scoring is, in fact, an activity worthy of this expenditure of a very valuable resource, professional time.

Another cost aside, from the costs of training and the opportunity costs, is the incentive to use technology. Risk systems tend to over refer and imply the use of secondary, and in many cases tertiary, care. In some cases, use of risk systems have led even "low risk" women to demand higher technology services. A perinatal regionalization scheme in the United States with extensive referral of "high risk" cases resulted in a situation in which there soon were no rural practitioners available to give primary level care to pregnant women: the demand for primary care had dropped rapidly—as most women self-referred into secondary and tertiary care centers. This resulted in a decrease of care near women—at the same time as we are learning that, even in the United States, longer distance to delivery facilities is highly correlated with poorer outcome. One may question whether this system of referral was, in fact, a net service to the population of the region or not.

Finally, there are psychological and user costs to the institution of a risk system. One issue related to this is the issue of being labelled as a "high risk" individual. How is this viewed by women? What does it do to their peace of mind during pregnancy? How can any stresses caused by being labelled "high risk" affect pregnancy? Is there stigmatization associated with "high risk" status, and is this reflected in a feeling of guilt by women? Are we undermining both the support and empowerment that are components of good pregnancy care?

A second consumer issue, is that risk systems may make it difficult to communicate truthful and accurate information and ignore womens' own valuation of risk. The presentation of information in risk system may itself be distorting of the real risks women run. To illustrate this I would like to show the following three slides, based on real data but drawn in different ways. (SLIDES 12, 13, 14) How risky is "high risk"?

Risk systems furthermore, do not easily respond to the valuation that women themselves put on risks and benefits. For example, a woman who has not been pregnant and did not marry until quite late in life may be fully willing to bear any risks associated with being pregnant above age 40. For this woman, the only chance to have children is to have them when she is over 40. Is it fair, is it appropriate—or even true—to label her "high risk" ?

Finally, there are ethical issues involved in use of risk systems. In some cases, risk systems are seen as an end in themselves, in which it is adequate to identify women at risk but there is really nothing to offer them. As a final comment on the utility of maternal risk I would suggest that it is *only useful as a means to an end, and that is to do something to help women*. (SLIDE 15) Therefore, I would propose that it is not useful, not ethical, not proper, not sensible to identify problems unless there is a functional referral and treatment system for the problems identified. The lack of referral and treatment is a fundamental ethical and practical problem with risk systems, especially in resource-constrained environments—where using such systems may merely foster the appearance of doing something.

As it is much more likely that there is something that can be done about *specific problems* than about *risk*, I would like to propose here that we change our vocabulary. (SLIDE 16) I would like to distinguish among: looking for risk; population screening for asymptomatic disease; and solving problems. Looking for risk is probably the least helpful of all possible activities. Screening patients may be useful depending on what is being screened for, how much it costs, and what can be done about it. But certainly identifying problems, particularly when

they are early and manageable at lower costs with less morbidity, is a goal we should seek to achieve.

I would propose, then—and we could test this hypothesis—that rather than looking for *risks*, we should look for *problems*. We should find people who have problems that we can do something about. We should let clinicians do clinical work, and we should seek to find not women at high risk but problem pregnancies. With this perspective, we probably can do more to help women than by talking about and creating elaborate risk systems that tend to be inaccurate, poorly taught, haphazardly used, and distorting to the resource-constrained medical systems of developing countries.