CP—75—5

INTERNATIONAL INSTITUTE FOR IIASA APPLIED SYSTEMS ANALYSIS

CONFERENCE PROCEEDINGS

# PROCEEDINGS OF THE JOINT IIASA/WHO WORKSHOP ON SCREENING FOR CERVICAL CANCER

April 1—2, 1975

SCHLOSS LAXENBURG
A-2361 AUSTRIA

CP—75—5

# PROCEEDINGS OF THE JOINT IIASA/WHO WORKSHOP ON SCREENING FOR CERVICAL CANCER

## April 1—2, 1975

## Preface

On April 1-2, 1975, a Workshop on Screening for Cervical Cancer was held at the International Institute for Applied Systems Analysis (IIASA), Laxenburg, Austria. The Workshop was jointly sponsored by the Cancer Unit of the World Health Organization (WHO), and the Bio-Medical Project of IIASA.

There were sixteen participants at the Workshop (see List of Participants). Five papers were presented for discussion at the Workshop. The papers and the summaries of the discussions that followed their presentation are reproduced here.

The participants heard a welcoming address by Dr. A. Garin, Head of the Cancer Unit, WHO. Introductory remarks were also given by Dr. R. Levien (US), Project Leader of the Handbook Project, IIASA; by Dr. A. Kiselev (USSR), Deputy Project Leader of the Bio-Medical Project, IIASA; and by Dr. J. H. Bigelow, member of the Bio-Medical Project, IIASA, and Workshop Coordinator.

## List of Participants

Workshop Coordinator: Dr. J.H. Bigelow

Dr. A.A. Afifi
Bio-Medical Project, IIASA

Mr. H. Brown
Bio-Medical Project, IIASA

Prof. William Christopherson
School of Medicine,
University of Louisville,
Kentucky, USA

Dr. L. Dobrossy
Head, Cytopathology Department
Research Institute of Oncopathology
Budapest, Hungary

Ms. D. Ellis
Bio-Medical Project, IIASA

Dr. B. Fraser
Cytological Diagnostic Laboratory,
Vancouver General Hospital,
Vancouver, British Columbia,
Canada

Dr. A. Garin
Head, Cancer Unit,
WHO, Geneva,
Switzerland

Mr. N.J. Glass
Bio-Medical Project, IIASA

Dr. A. Kiselev
Deputy Project Leader,
Bio-Medical Project, IIASA

Dr. R. Levien
Project Leader
Handbook Project, IIASA

Mrs. R. Lunt,
Cancer Unit,
WHO, Geneva,
Switzerland

Prof. G. Riotton
Director of Pathology,
WHO, Geneva,
Switzerland

Prof. N. Stormby
Director,
Cytodiagnostic Laboratory,
Malmo General Hospital,
Sweden

Dr. A. Winkler
Cancer Unit,
WHO, Geneva,
Switzerland

Dr. H. Wrba
Director,
National Cancer Institute,
Vienna, Austria

Table of Contents

# INTRODUCTORY REMARKS

## What is IIASA?

R. Levien

On October 4, 1972, at an Inaugural Conference in London, the Charter of the International Institute for Applied Systems Analysis was signed. The establishment of the Institute reflected the intent of prestigious scientific organizations from twelve countries to join efforts in combating a number of complex problems engendered by the contemporary stage of scientific and technological evolution.

The scale of scientific and technological problems confronting present-day societies has reached a point where international cooperation becomes a necessity in order to cover all their aspects and to predict the outcome of decisions taken.

The possibility of establishing a research center to deal with these problems was discussed at a Soviet-American meeting in Moscow early in 1967; the USA was represented by Mr. McGeorge Bundy. There followed a period of exploration and multinational negotiations to crystallize the idea. The main participants in these negotiations were:

Monsieur Pierre Aigrain, Government of France;

Prof. Philip Handler, the National Academy of Science, USA;

Dr. O. Leupold, German Democratic Republic;

Signor Aurelio Peccei, Italy;

Dr. Friedrich Schneider, Max Planck Society, of the Federal Republic of Germany;

Prof. D. Smolenski, Polish Academy of Sciences; and

Lord Zuckerman, UK, Chairman.

The Institute commenced operations in Vienna toward the end of 1972, in the favorable environment created by the Austrian Government.

The twelve founding members who inaugurated the Institute in October 1972 expressed the strong desire to welcome further member organizations to IIASA. Accordingly, a gradual increase

in the number of new National Member Organizations in the first phase of development of the Institute was mutually agreed upon.

The host country, Austria, represented by the Austrian Academy of Sciences, was unanimously admitted to IIASA during the Third Council Meeting in November 1973, and thus became the thirteenth member. In 1974, Hungary was admitted as the fourteenth member. The present list of member organizations is as follows:

The Academy of Sciences,
Union of Soviet Socialist Republics;

The Austrian Academy of Sciences;

The Committee for the International Institute for Applied Systems Analysis, Canada;

The Committee for the International Institute for Applied Systems Analysis of the Czechoslovak Socialist Republic;

The French Association for the Development of Systems Analysis;

The Academy of Sciences of the German Democratic Republic;

The Japan Committee for the International Institute for Applied Systems Analysis;

The Max Planck Society for the Advancement of Sciences, Federal Republic of Germany;

The National Centre for Cybernetics and Computer Techniques, People's Republic of Bulgaria;

The National Academy of Sciences, United States of America;

The National Research Council, Italy;

The Polish Academy of Sciences;

The Royal Society, United Kingdom;

Hungarian Committee for Applied Systems Analysis.

It should be noted that IIASA's members are not governments or governmental agencies, but academic organizations.

## Activities at IIASA

IIASA employs applied systems analysis as an approach to

problem solving in broadly diverse contexts. At its two 1973 meetings, the Council identified an initial roster of research areas for the Institute. At present, IIASA has ten separately led but closely cooperating research areas. These are:

1) Methodology of Applied Systems Analysis,

2) Design and Management of Large Organizations,

3) Computer Systems,

4) Integrated Industrial Systems,

5) Urban and Regional Systems,

6) Ecological Systems,

7) Biological and Medical Systems,

8) Energy Systems,

9) Water Resources, and

10) State-of-the-Art Survey Project.

The first three of these areas provide methodological support for the other activity areas. We recognize, however, that we cannot obtain a sufficient number of methodologists organizations experts or computer specialists if they are permanently confined to auxiliary, support roles. For this reason, in each of the first three areas, the specialists devote part of their time to methodological or theoretical research that will enable them to remain at the frontiers of their various disciplines.

The state-of-the-art survey project is preparing a handbook, in an effort to disseminate more widely the methods and effective practice of systems analysis.

Finally, a "general activities" project serves as a home for embryonic activities that do not fit easily into the more focussed projects.

The research tasks are chosen on the basis of a number of criteria. Among these are:

1) Appropriateness to the systems approach. IIASA seeks out real-world problems requiring the inputs of many scientific disciplines. Wherever possible, we intend that purely methodological research will evolve from applied work. There is also strong interest in the managerial aspects of problems; there is often participation in the implementation of decision alternatives selected through analysis.

2) Global or universal nature. As an international institute, IIASA must restrict itself to two classes of problems:

a) global problems where the effects and the reins of policy necessarily involve many nations as, for instance, studies on oceans, the atmosphere, and continental riversystems; and

b) universal problems where activities may be separately controlled by single nations, but similar methodologies are used as, for example, municipal services, health care services, and management of industrial systems.

3) Non-redundancy. IIASA has no desire to duplicate the scientific work most appropriately carried out by individual groups in single nations.

4) Importance to NMO's and to mankind in general.

5) Feasibility. IIASA considers the tractability of all proposed research in the light of its resources.

Cervical Screening

As a specific task, the study of screening for cervical cancer meets all of the criteria. It is of universal interest; indeed, screening programs for cervical cancer are worldwide. It is important; what is more important than saving lives? And research of this kind appears not to be taking place elsewhere.

We believe that the following pages will demonstrate both the feasibility and appropriateness of applying systems analysis to the study of screening for cervical cancer. To solve this problem, the disciplines of medicine, economics, epidemiology, sociology, and managerial science must combine and be coordinated by the powerful and sophisticated techniques of systems analysis. Because of the evident widespread interest in this subject, we have strong hopes that our results may be implemented.

Collaboration

IIASA is a small institute (60 scientists) with large ambitions. Thus we must amplify our efforts through collaboration with national and international institutions. One method of collaboration is to serve as an information clearing house for researchers who are widely separated geographically, but who share research interests. Another method is to gather together people with common interests (but not necessarily common knowledge or skills) at meetings such as this one. A third method is to collaborate directly with other institutions. This meeting, which is co-sponsored by WHO and IIASA, is among the first fruits of such collaboration.

## Conclusion

Thus the Workshop, and the larger study of which it is a part, can serve to exemplify the most important goals and methods of IIASA. Each time a meeting is held or a study is commenced, these methods are once again tested and the goals are again put to question. Each time, this occurs before a new and usually skeptical audience. If the meeting or study succeeds, members of that audience will return and lasting associations can be formed. If the event is a failure, the audience never returns. We can only hope, therefore, that the Workshop will succeed, and that afterwards the participants remain in close association with IIASA, and with applied systems analysis.

## Welcoming Address

Dr. A. Garin

It is a great honour and pleasure for me to convey to the Workshop the greetings of the Director-General of the World Health Organization. The participants understand very well the great importance of integrating activities in the struggle against cancer. The World Health Assembly recognized that work on cancer absorbs a substantial and increasing part of the financial and other resources of member states and of their research institutions.

The complexity of the problems of cancer and the unlikelihood of their being fully elucidated by any one country is the reason why the 140 member states of WHO recommended the development of a long-term global effort of international cooperation in cancer research. The following objectives in the area of cancer for ongoing and future activities are attracting WHO's attention:

- organization of an international dialogue between basic scientists and clinicians to foster the application of fundamental research achievements to cancer control;
- promotion of epidemiological investigations, analysis of their results, and determination of the role of environmental agents in cancer causation in different geographical areas;
- international review of the current status of early detection, diagnosis, treatment and rehabilitation of patients with the more common forms of cancer, e.g. lung, stomach, breast, uterus, bladder, colon, prostate; worldwide dissemination of information on optimal methods;
- standardization of systems for registering cancer cases, and for reporting and evaluating results of treatment;
- promotion of widespread use of WHO standardization nomenclature, methodologies, reagents, etc.;
- determination of the significance of the cancer problem in countries where cancer registration is currently deficient;
- elaboration of organizational principles for the structure of national cancer health services;
- global collection and dissemination of information about cancer control resources; and

- promotion of the development of health manpower for cancer research and control.

I do not need to emphasize to this audience the benefit of attracting mathematicians, statisticians, and information and computer specialists to the problem of cancer. The alliance of mathematical analysis and abstraction with medical experience and impressions seems to us to have a very promising potential for success in oncology.

This Workshop is WHO's first scientific contact with IIASA in the field of cancer. Cervical cancer has been selected as the object of this first collaborative activity. This type of cancer is common in both developing and developed countries. About 35,000 women die annually from cervical cancer in the Common Market countries; cervical cancer is a leading type of tumour among women in Asia and South America.

Owing to the development and implementation of early detection methods and screening programs, there appears to be notable improvement in the results of treatment and prognosis of cervical cancer in many countries, e.g., Canada, UK, USSR, USA.

An important task before us is to analyze and to evaluate these achievements, and to disseminate appropriate information to countries where cervical cancer remains a main killer of women.

Elucidation of the natural history of this disease should help both to select the best screening tactics and to choose the optimal methods of treatment.

Achievements with this model could be applied in the future to research into other types of cancer. Many aspects of the biology and epidemiology of cervical cancer, its pathogenesis, natural history (development characteristics, phases of growth) and cytological evaluation are controversial. We certainly understand that this project will not solve all the problems of cervical cancer, but we hope that it will shorten the time needed to clarify some of these.

Once again I would like to convey to you the best wishes of WHO for a successful Workshop, and for fruitful collaboration in the future.

## IIASA's Interest in Cancer

A. Kiselev

The IIASA Bio-Medical Project is devoting a large proportion of its resources to the study of cancer. Since our resources are limited, it would be absurd to think that we could solve many of the problems through our own direct efforts; IIASA must amplify its effect through collaboration.

One form of collaboration is joint activities with the large organizations that are currently collecting enormous amounts of information on cancer and cancer research, as for example, the World Health Organization. While these organizations have facilities for storing and retrieving this information that IIASA does not have, we believe that we can help them to make better use of these facilities.

Our direct effort in this activity will be to investigate the use of information systems theory. This involves developing new definitions of directions of cancer research, not only as to the site of the cancer and the methods used, but also as to the theoretical bases of a research project. We are now surveying the various "theories" of cancer and carcinogenesis to see if they can provide a useful partial classification of cancer research efforts.

This, in brief, is our direct approach to the problems of cancer and cancer research--the "top down" approach, so called because it addresses the problem initially in its most general form and only later in its specific manifestations. But our approach cannot only be from the top; we must concurrently look at smaller parts of the problem. If not, we will find that our general results do not apply to specific cases.

The cervical screening study is a specific activity. For this study, we intend to look at further research on cervical cancer. For example, let me quote from the research prospectus that I think all of you have received

"...one may ask how much one should pay to improve the prognosis of cases of invasive cancer by a stated amount. If the prognosis is improved, one will be able to reduce the size (and hence the cost) of the screening program while maintaining the total benefit (e.g. reduced mortality) unchanged. The reduction in screening cost is then a measure of the value of improving the prognosis."

So, this study is not only an example of research into cancer, it is also research into research. We need the aid of all

interested parties to carry out both the general and the specific work. The general effort is still in the planning stage; as planning proceeds, we will be asking you and your colleagues for advice and criticism. Later, when specific research tasks have been defined, we may be asking you for other kinds of help. We hope you will treat these requests kindly.

This Workshop is, in essence, a request for assistance to the cervical screening study. Although the study is already planned, you may be aware of other problems that must be overcome, problems that we do not know exist. If this were true, we would change our plans, and allot the necessary resources. You will hear of other ways you can help us, and you may think of ways yourselves during our discussions.

I will close now by welcoming you once again to the Workshop. All of us in the Bio-Medical Project hope that you find the Workshop interesting and worthwhile, and that your stay here is pleasant.

# Introduction

J. H. Bigelow

## Purpose of the Workshop

In hosting this Workshop, we of the Bio-Medical Project had in mind obtaining help for our proposed study on screening for cervical cancer. Three types of help are needed. First, we seek advice on the subject of cervical screening from people knowledgeable in the field. Second, we need access to data from existing cervical cancer screening programs. Third, we wish to recruit a client for our study, someone who has an interest in the results and who might implement these results at the conclusion of the study, at least on a pilot basis.

The first two types of assistance are self explanatory. The practical problems of taking a satisfactory PAP smear, or of inducing a gynecologist to exercise proper care in taking smears, can be known only to those who have had long and extensive experience in a cervical screening program. Such matters must be considered in our proposed study. Equally, without data obtained from actual screening programs, our study will be no more than a schoolroom exercise.

The third purpose, that of recruiting a client, is less obvious. Our proposed study will be a policy study. That is, its ultimate results will be recommendations as to the design and implementation of cervical screening programs. It is not worthwhile to make recommendations unless there is the possibility of testing these recommendations. Moreover, they cannot be tested unless they are specific to a particular situation--i.e., geographical area, subject population, available resources and manpower--and unless someone is willing to implement the recommendations, at least on a pilot basis. A client provides the particular situation, and the possibility for testing the results.

## Study Outline

Since the purpose of the Workshop is to obtain help for our study of cervical screening, it will be worthwhile to give a brief outline of the study that will enable the reader to judge the pertinentcy of the advice that we received.

The study may best be described in terms of the methodology that will result from it. This methodology will consist of a number of models that relate the following three kinds of information:

- design of the screening program, including those who take the smears (e.g. doctor, nurse, or technician); the physical structure in which the smears are taken (e.g. doctor's office, hospital clinic, mobile unit); the procedures (or lack thereof) for recruiting women to be screened (e.g. letter, personal visit, wait until woman sees her doctor), etc.;

- composition of the population to be screened, by age, race, income, and other relevant factors; and

- effects that the screening program will have on the given population, including the program's cost, its manpower needs, its effect on mortality and morbidity, etc.

We will be able to operate this methodology in two directions. First, given the design of the screening program, and the population it is to serve, we will be able to predict the effects that it will have. This is the easier direction. Others, notably Knox [1], [2], have built simulation models that carry out this task.

Second, given the population to be screened and the effects that one wishes to produce, we will design a screening program that will do the job. The desired effects are usually to maximize or to minimize something subject to constraints, as for example, to reduce mortality as much as possible within given budget and manpower requirements. To the best of our knowledge, this kind of model has not been applied to the cervical cancer screening problem.

Required Knowledge

In order to build these models, we will need: a) technical information, and b) sociological and organizational information. The most important technical information needed is knowledge of the natural history of cervical cancer. This involves knowing how rapidly cases of carcinoma in-situ may progress to invasive cancer. Cases that do so in very short time periods are likely to be missed by a screening program. We also wish to know what proportion of preclinical cases of carcinoma in-situ are likely to regress to normal. These cases will be unnecessarily treated if detected by screening.

Other important technical information needed includes a knowledge of how, and with what effect, the disease can be treated at different stages. Possible outcomes of treatment include complete cure, death, or complications arising from treatment. Without this knowledge, we cannot estimate mortality and morbidity. It is also important to know the errors that may occur in the collection and interpretation of the PAP smear. False-positive smears involve the medical community in unnecessary work, and subject the patient to unnecessary anxiety. False-negative smears allow cases of a potentially lethal disease to remain untreated.

Finally, it is important to know who is most likely to have cervical cancer. Factors such as age, hygiene, and sexual behavior are known to influence the incidence of the disease. One should be aware of those parts of the population that are at greater risk and should thus be screened if resources do not permit the whole population to participate in screening. One should also understand the degree to which changes in habits--e.g. improved hygiene due to educational measures--can replace screening as a means for reducing mortality from cervical cancer.

As regards sociological and organizational information our major information need is in the area of participation. Why do the poor, the young, and the old appear less frequently for screening than the well-to-do middle-aged? What can be done to attract these infrequent participants? What additional measures are needed to increase repeat participation?

Some additional information requirements are: how, in practice, have doctors reacted to requests from their patients to take smears; how can the number of unsatisfactory smears be reduced; how can the quality of cytology be ensured?

## Applications of This Research

The purpose of this Workshop is to describe and to refine a study which we believe would aid the formulation of a policy for screening for cervical cancer. The results of this study would be useful for countries that have created screening programs or those in which screening programs have grown up without conscious political decisions. The results would also be useful for countries that are contemplating cervical cancer programs. Finally, this study could serve as prototype for the study of other diseases where screening programs are being contemplated.

Administrations or organizations contemplating the establishment of cancer screening programs need to know what consequences are likely to result from such a decision. How can they design a program that will best meet their objectives, subject to the constraints on manpower and physical resources with which they are faced? Should they introduce a program at all? If so, how quickly should it be introduced? These are questions that can best be answered by testing and evaluating a number of alternatives. For countries where cervical cancer screening programs already exist, the "political" costs and benefits of reducing or expanding the program will probably be evident to medical policy-makers. What they may not know are the medical and economic consequences of such decisions. Models, such as those we hope to build, that trace such consequences should make a vital contribution to policy discussions.

Attempts to model complex policy questions are bound to be hindered by many difficulties of both fact and method. However, we believe that many of these problems can be overcome, and that the

methods for overcoming them will have useful applications elsewhere. This is especially valid if we remember that the design and implementation of screening programs are likely to be of increasing concern to health services worldwide.

## References

[1] Knox, E.G. "A Simulation System for Screening Procedures." In Problems and Progress in Medical Care, Nuffield Provincial Hospital Trust. London, Oxford University Press, 1973.

[2] Knox, E.G. "Computer Simulation of Alternative Population Screening Policies." In Systems Aspects of Health Planning. N.T.J. Bailey and M. Thompson. Eds. Amsterdam, North-Holland Publishing Co., 1975.

# PART I. TECHNICAL ISSUES

## 1.1 The Natural History of Cervical Cancer

J. H. Bigelow

### Purpose

The subject of the discussion in this session is the natural history of cervical cancer. This means, roughly, "how likely is the disease to progress to invasion, and how long is this likely to take?" In this paper, I will show how I expect to estimate the natural history quantitatively from existing data.

### Description of the Model

Let me first tell you precisely what I mean by the natural history. My definition is based upon the description of cervical cancer shown in Figure 1. The disease passes from an uncertain beginning, through dysplasia and carcinoma in-situ, and to an invasive stage. My model is a simplification of this picture. In the model, the disease passes from an initial state, through a development phase, to a terminal state.

I propose to place the initial state of the model at the boundary between dysplasia and carcinoma in-situ. This will effectively focus the attention of this part of our study on the lesion carcinoma in-situ. I feel this is wise, since this is the main point of controversy in the area of cervical screening.

There is less question about where to locate the terminal point of our model of the disease. It is that instant at which the disease would have been detected in the absence of any screening.

Not all cases will have progressed to the same extent when discovered clinically. We illustrate this as a sloping line that "cuts off" some cases only where they have progressed to late-stage invasive cancers, and cuts off others even before invasion--for example, due to a biopsy performed for some reason unrelated to cancer. Thus the development phase of our model does not exactly correspond to the phase carcinoma in-situ.

One cannot expect that every case will pass through the development phase in the same amount of time. Some cases will develop very rapidly while others will take an extremely long

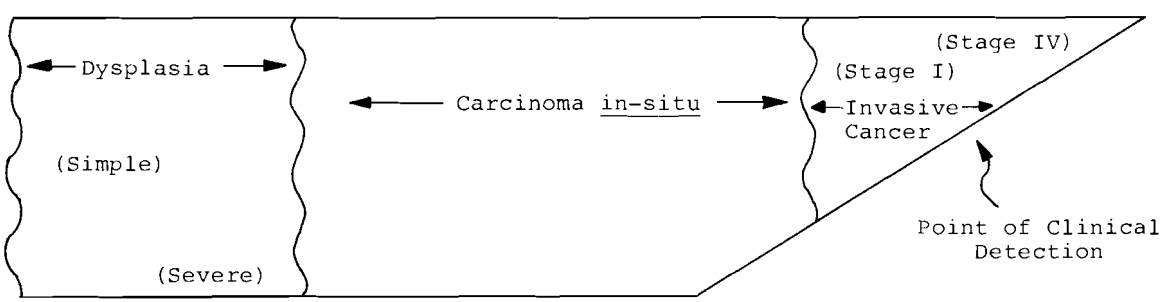



Figure 1. Concept of cervical cancer.

time. We will call the time spent in this phase the dwell time.

We can represent this simple model pictorially, as shown in Figure 2. A case with a particular dwell time, measured on the vertical axis of the figure, must develop from the point of first appearance, across the figure on a horizontal path to its terminal state. Mathematically, we express the model as a function G(T), that gives the number of cases per unit of population per year that have dwell time T.

Because some people believe that a substantial fraction of carcinomas in-situ regress spontaneously to normal, I introduce a second form of the disease. This has the same initial state as the old form, but its terminal state is "regress to normal" rather than "progress to clinical detection." We will require a different dwell time distribution for each of the two kinds of disease, $G_P(T)$ for progressive cases, and $G_R(T)$ for regressive cases.

Figure 3 shows example distributions. The number of cases of each kind, progressive or regressive, is given by the area under the respective curve. For example, the number of progressive cases with dwell time between five years and ten, is given by the area of the shaded region in the figure. Note that there is no need for the two distributions to be the same or even to have the same average dwell times.

These two distributions constitute my definition of the natural history of cervical cancer. Thus to say that I wish to determine the natural history of the disease is to say that I wish to estimate the distributions $G_P(T)$ and $G_R(T)$.

## Estimating the Dwell Time Distributions

My method for estimating these distributions is best illustrated by an admittedly artificial example. I assume, for this example, that the screening test is perfect, never yielding false positive or false negative. I also assume that screening tests are invariably performed at two, four, and eight-year intervals, and that the data include, let us say, one million intervals of each size. I further suppose that dwell times of cases are always either one or three or five years, but the numbers of progressive and regressive cases with each dwell time are not known.

The first type of information we will need is that given in table form in Figure 4. Of the one million screening intervals of two-years duration, I will suppose that 1,100 resulted in the discovery of a carcinoma in-situ. That is, the test that began the screening interval was negative, and the test that ended the interval was positive. In the absence of false negative tests, the carcinoma in-situ must have started sometime during the interval. Similarly, I suppose that 1,900 cases were discovered in the four-year intervals, and 2,200 cases in the eight-year intervals.

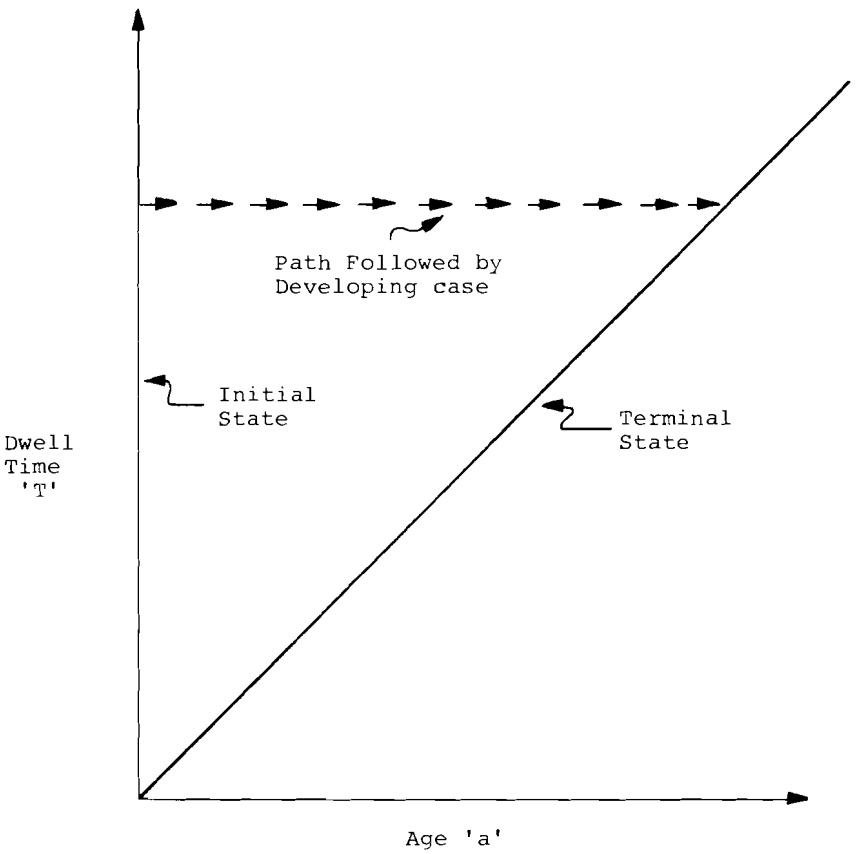



Figure 2. Simple model of cervical cancer.

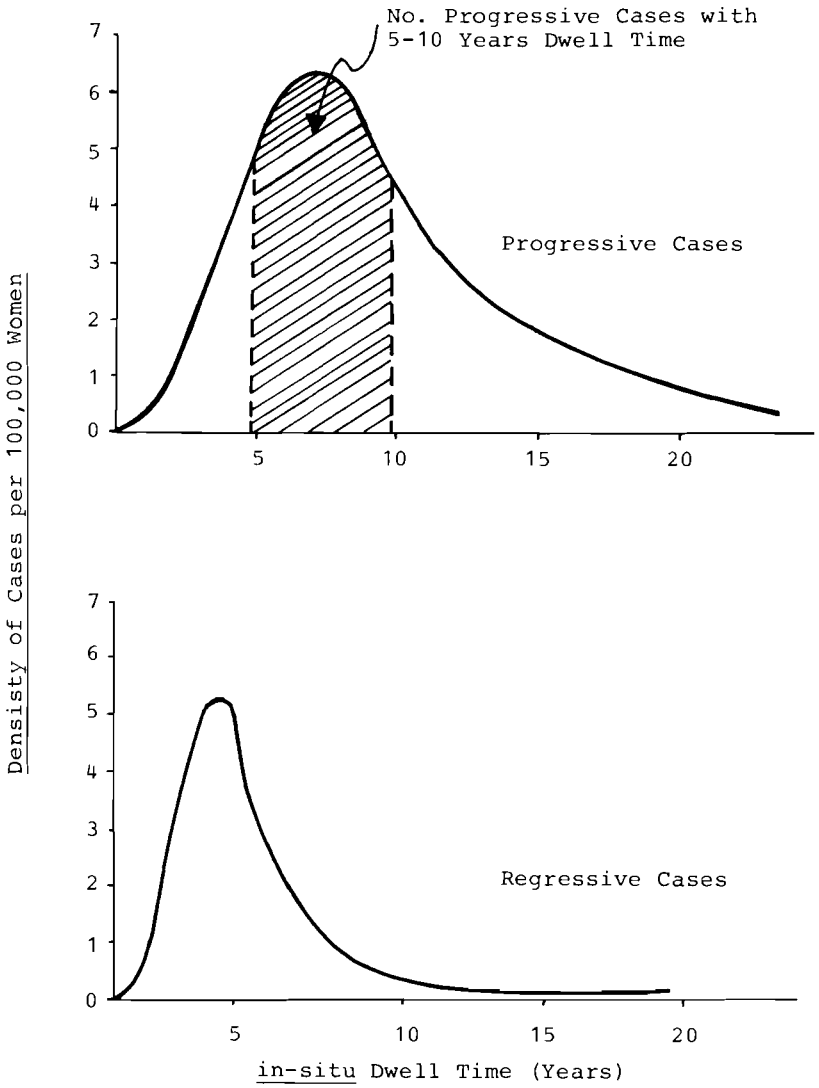



Figure 3. Example of dwell time distribtuions.

| | Length of Interval | | |
|---|---|---|---|
| | 2 Years | 4 Years | 8 Years |
| Number of Intervals | $10^6$ | $10^6$ | $10^6$ |
| Number of Cases Detected | 1,100 | 1,900 | 2,200 |

$x_1$ = Number of cases per woman-year with dwell time i (progressive and regressive combined), i = 1,3,5.

$$2 \times 10^6 \, [1/2x_1 + x_3 + x_5] = 1,100$$
$$4 \times 10^6 \, [1/4x_1 + 3/4x_3 + x_5] = 1,900$$
$$8 \times 10^6 \, [1/8x_1 + 3/8x_3 + 5/8x_5] = 2,200$$

Figure 4. Sum of dwell time distributions.

Let us define $x_i$ to be the number of cases with dwell time i that start in the average woman-year. This quantity includes both progressive and regressive cases. Thus $x_1$ is the rate at which cases with a one-year dwell time are initiated, $x_3$ the rate for three-year cases, and $x_5$ the rate for five-year cases.

The three equations in Figure 4 estimate the number of cases that should be detected in intervals of two, four, and eight years, respectively. Taking the first equation, the leading coefficient of two million is the number of woman-years contained in the one million two-year intervals. The expression in brackets is the expected rate at which cases of each dwell time will be initiated and detected, in intervals of two years. Thus $x_1$ cases, with a dwell time of one year, will be initiated in each year of these intervals; but only one half of these cases will be detected, because to be detected, such a case must be initiated in the last half of the interval. Cases with three- and five-year dwell times will always be detected if they start in a two-year screening interval. Thus the coefficients of $x_3$ and $x_5$ are one. The other two equations are derived in a similar way.

On the right-hand sides of the equations are the observed numbers of the cases detected in screening intervals of each length. These three equations can be solved for $x_1$, $x_3$, and $x_5$, by well-known methods, yielding the result, $x_1 = 1 \times 10^{-4}$, $x_3 = 2 \times 10^{-4}$, $x_5 = 3 \times 10^{-4}$, respectively. That is, each year one expects ten women out of every 100,000 to contract a one-year disease, twenty women more to contract a three-year disease, and another thirty women to contract a five-year disease.

The second sort of information I will use is the number of cases that are detected clinically--i.e. progressive cases that escape detection by screening--at various times after the latest screening test (see Figure 5). In this example, 150 cases were detected in the clinic between one and two years after screening, 400 cases between two and four years, and 1,050 cases between four and eight years. I estimate the number of opportunities for this to occur in the top line of the table. Thus there were three million intervals of at least two years, two million intervals of at least four years, and one million of at least eight years.

Now define $y_i$ to be the number of progressive cases with dwell time i that start in the average woman-year. Taking the second equation for purposes of discussion (see Figure 5), the leading coefficient, two million, is the number of opportunities for the disease to be detected clinically between two and four years after the last screening test. On each of those occasions, a case with a one-year dwell time could have been initiated at any time between one and three years after the test, that is, in a two-year interval. This explains the appearance of $2y_1$ in the

| | Time Since Last Test | | |
|---|---|---|---|
| | 1-2 Years | 2-4 Years | 4-8 Years |
| No. Suff. Long Screening intervals | $3 \times 10^6$ | $2 \times 10^6$ | $1 \times 10^6$ |
| Number of Cases Detected Clinically | 150 | 400 | 1,050 |

$Y_i$ = Number of progressive cases per woman-year with dwell time i, i = 1,3,5.

$$3 \times 10^6 \cdot [y_1 \qquad\qquad\qquad\qquad\qquad\quad] = 150$$

$$2 \times 10^6 \; [2 \cdot y_1 + y_3 \qquad\qquad\qquad\quad] = 400$$

$$1 \times 10^6 \; [4 \cdot y_1 + 4 \cdot y_3 + 3 \cdot y_5] = 1{,}050$$

Figure 5. Progressive dwell time distribution.

expression in brackets. On the other hand, a case whose dwell time is three years has only a single year in which it could be initiated if it is to begin after the most recent test and be discovered clinically within four years of the test. Thus only one $y_3$ appears in the expression. No five-year cases can be clinically detected within four years of the last test, given the assumption of a zero false-negative rate.

The observed numbers of clinically detected cases are also given to the right of these equations, the second line of the table. Again, these equations can be solved for $y_1$, $y_3$, and $y_5$, and the results (the initiation rates of progressive cases) can be substracted from $x_1$, $x_3$, and $x_5$, that were calculated earlier to yield regressive case rates. The results are shown in Figure 6.

I wish to stress that the example is entirely artificial. I do not intend to express an opinion about the progressive or regressive dwell time distributions, nor am I suggesting that cases of cervical cancer require more than four years to develop clinical manifestations. As I said when I first introduced the example, it is entirely artificial.

My purpose in presenting this example is to illustrate that a relationship exists between the progressive and regressive dwell time distributions, and the observed case histories. Moreover, the example shows how this relationship can be used to estimate the two dwell time distributions.

## Extension of the Method

The method illustrated can be extended to deal with more realistic cases. The most important extension, in my view, is to deal with false-negative smears. I have chosen to ignore the false-positive smears, because positive smears are followed up and a false positive will be discovered rapidly. But false-negative smears allow untreated cases of carcinoma in-situ to remain in the population at large, just as though a smear had not been taken at all.

The observation that a false negative smear is equivalent to not having taken a smear serves as the basis for modifying the method. As we scan through the original screening data, we randomly discard some of the negative smears. How many will be discarded will depend on the false-negative error rate. (This will be discussed in Section 1.3.) We carry out the analysis illustrated by the example with only the remaining smears.

This method can also be extended to estimate the dwell time distributions with greater refinement. Instead of assuming that only three dwell times are possible, one could assume that there were twenty, or fifty, or more. To do so, we need information of the kind shown in the example, but richer.

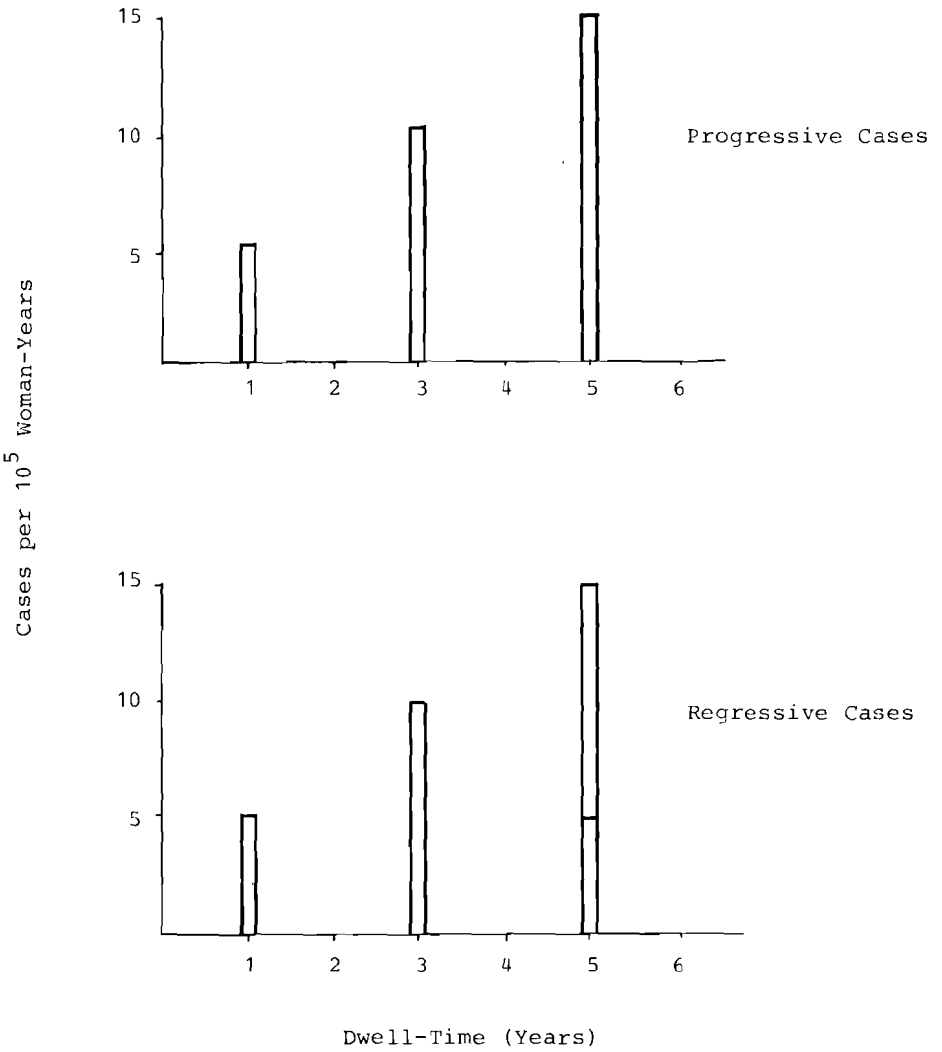



Figure 6. Calculated dwell-time distributions from example.

The two kinds of information needed for the general case are illustrated in Figure 7. To calculate the sum of the two distributions--i.e. progressive plus regressive, I need information on the frequency with which screening intervals of each length result in the detection of a case. To obtain the progressive dwell time distribution, I must know, for each of the intervals of time, how many cases are found clinically in the amount of time after the most recent negative smear. I must also know how many screening intervals there are of each length.

The refinement one can achieve in estimating dwell time distributions will be limited by the amount of data available. To estimate the distributions, we must have reliable frequency estimates of the two kinds of disease history. A frequency estimate will not be reliable unless enough data points contribute to it.

On the basis of my statistical experience, I estimate that, for every five cases detected by screening, one point can be estimated on the sum of the two dwell time distributions. Similarly, for every five cases that occur clinically, one point on the progressive dwell time distribution can be estimated. The British Columbia [9] experience gives a good example of the relative numbers of the two kinds of cases. From 1949 until 1966, there were sixty-four clinically detected cases among screened women, and over 3,000 cases detected by screening.

Finally, we know that the population at risk from cervical cancer is not homogeneous. Incidence varies with age, for example [4], as well as with race, social class, and age at first coitus. Different sexual habits [14] or habits of personal hygiene could also make a difference.

To deal with this lack of homogeneity, I expect to split the original screening data into parts, each of which deals with only a homogeneous population of women. Of course, we can split the data only according to the characteristics recorded in the data, and so our study of these factors will be incomplete.

## Summary

I can think of no better way to end this presentation than by saying that I believe the natural history of cervical cancer can be estimated from existing data. Of course, there are difficulties. The data must be adjusted to account for false-negative smears, and possibly split into segments that deal with homogeneous groups of people. There may be errors in the data, in particular misidentifications that will affect the results. (This will be discussed in Section 1.2.) The results will, in any case, contain an inherent probable error. Nevertheless, I believe that the natural history (that is, the progressive and regressive dwell-time distributions) can be estimated from data that have been collected from existing cervical screening programs.

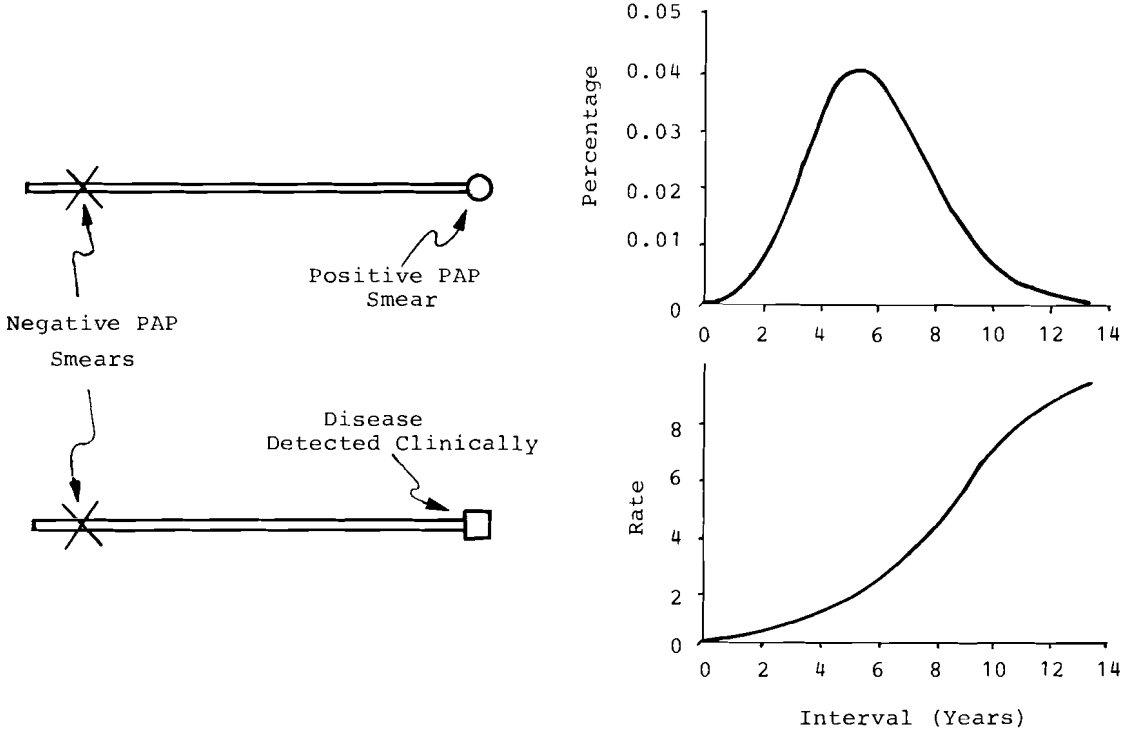



Figure 7. Frequencies of disease histories.

## References


[1] Ashley, D.J.B. "The Biological Status of Carcinoma in-situ of the Uterine Cervix." J. Obst. Gynec. Brit. Cwlth., 73 (1966) 372.

[2] Ashley, D.J.B. "Evidence for the Existence of two Forms of Cervical Carcinoma." J. Obst. Gynec. Brit. Cwlth., 73 (1966) 382.

[3] Barron, B.A. and Richart, R.M. "A Statistical Model of the Natural History of Cervical Carcinoma Based on a Prospective Study of 557 Cases." Journal of the National Cancer Institute, 41, 6 (1968) 1,343.

[4] Beral, V. "Cancer of the Cervix: A Sexually Transmitted Infection?" Lancet (May 25, 1974) 1,037.

[5] Boyd, J.R., Royle, D., Fidler, H.K. and Boyes, D.A. "Conservative Management of in-situ Carcinoma of the Cervix." Am. J. Obst. and Gynec., 85, 3 (1963) 322.

[6] Dunn, J.E., Slate, J.A., Merritt, J.W. and Martin, P.L. "Finding for Uterine Cancer from One or More Cytologic Examinations of 33,750 Women." Journal of the National Cancer Institute, 23, 3 (1959) 507.

[7] Fidler, H.K., Boyes, D.A. and Worth, A.J. "Cervical Cancer Detection in British Columbia." J. Obst. Gynec. Brit. Cwlth., 75 (1968) 392.

[8] Jordan, M.J., Bader, G.M. and Day, E. "Carcinoma in-situ of the Cervix and Related Lesions." Am. J. Obst. and Gynec., 89, 2 (1964) 160.

[9] Knox, E.G. "A Simulation System for Screening Procedures." In Problems and Progress in Medical Care, Nuffield Provincial Hospital Trust. London, Oxford University Press, 1973.

[10] Knox, E.G. "Computer Simulation of Alternative Population Screening Policies." In Systems Aspects of Health Planning. N.T.J. Bailey and M. Thompson. Eds., Amsterdam, North-Holland Publishing Co., 1975.

[11] Krieger, J.S. and McCormack, L.J. "The Indications for Conservative Therapy for Intraepithelial Carcinoma of the Uterine Cervix." Am. J. Obst. and Gynec., 76, 2 (1958) 312.

[12] Nieburgs, H.E. "The Significance of Tissue Cell Changes Preceeding Uterine Cervix Carcinoma." Cancer, 16 (1963) 141.

[13] Petersen, O. "Spontaneous Course of Cervical Precancerous Conditions." Am. J. Obst. and Gynec., 72, 5 (1966) 1,063.

[14] Punnonen, R., Groenroos, M. and Peltonen, R. "Increase of Premalignant Cervical Lesions in Teenagers." Lancet (October 19, 1974) 949.

[15] Spriggs, A.I. "Follow-Up of Untreated Carcinoma in-situ of Cervic Uteri." Lancet (September 11, 1971) 599.

## 1.2 Summary Discussions of the Natural History of Cervical Cancer

The model of cervical cancer (see Figure 1) and the data from which dwell time distributions are estimated (see Figure 7) were discussed at length at the Workshop; a summary of these discussions is given below.

### Initial State of the Model

The initial state of the model is to be placed at the boundary between dysplasia and carcinoma in-situ. This would seem to suggest that there is a relatively clear cytological distinction between these two lesions. However, the facts do not bear out this distinction. A smear that one cytology laboratory rates as carcinoma in-situ may, upon review by another laboratory or by the same laboratory years later, be interpreted as a simple dysplasia.

For example, some years ago, the National Cancer Institute of the United States was receiving reports every six weeks from about fifty cytology programs. There was an enormous variation in the prevalence of carcinoma in-situ found by different programs; some of the findings reported as many as thirty women per thousand with this condition. Ratios of the incidence of carcinoma in-situ to the incidence of invasive cancer ranged from a low of one-to-one, to a high of approximately twenty-to-one. On review, the differences proved to be due to overdiagnosis by the laboratories reporting the highest incidence of carcinoma in-situ.

The grading system for smears that a laboratory uses may also present difficulties. Because it eliminates needless detail, grading a smear with the histological diagnosis expected for the patient in question is the best method. Thus a smear would be labeled, for example, normal, moderate dysplasia, carcinoma in-situ, or invasive cancer. However, many laboratories still use the Papanicolaou grading system, or a modification of it, that does not relate directly to the expected histological diagnosis.

In summary, there is no clear distinction between carcinoma in-situ and dysplasia; a distinction may be made differently by different laboratories, and even by the same laboratory at different times.

## The Terminal State of the Model

The terminal state of the model was to be placed at the point where the disease would have been clinically detected in the absence of a cytology program. However, this end point will depend on the skill of the clinician. Many good gynecologists will fail to recognize cervical cancer before Stage 3 invasion. Others can consistently recognize Stage 1A lesions. Thus the dwell time of a case might be as much as four years longer if the woman happens to see the first type of gynecologist than if she sees the second type.

In addition, the frequency with which a woman sees a gynecologist will influence the time at which her cancer is detected. The woman who visits her clinic regularly will probably have her case discovered in an early stage. The woman who visits a clinic only when symptoms occur will probably have a late-stage cancer. Dwell times observed in the former sort of woman will generally be shorter than dwell times in the latter.

The gynecologist who suspects that his patient has cervical cancer will want histological confirmation before he begins treatment. This will occur if the lesion is a very early one. A late lesion will be obvious even without histological tests. However, the tissue sample sent to be studied may be inadequate for diagnosis, or the laboratory may mishandle the sample, usually examining too few sections. As many as 30% of the tissue samples that contain evidence of early lesions may be passed as normal.

Finally, the age of the woman will influence how early clinical diagnosis can occur. After menopause, the neck of the uterus closes down, and tissue that was once exposed is now hidden. This hidden area is frequently a site of cervical cancer that would be detected later in postmenopausal than in premenopausal women.

In summary, the end point of the model of cervical cancer is not a physiologically well-defined point. It depends upon the general level of medical skill, the age of the woman involved, and on the frequency of contact between the woman and the medical system.

## Regressive Cases

It is well known that cases of dysplasia will frequently regress to normal without treatment. Whether this is also true of carcinoma in-situ is a hotly debated matter.

The opinion expressed at this Workshop was that carcinoma in-situ virtually never regresses spontaneously. Cases reported to have regressed were taken to be overdiagnosed cases--i.e., cases diagnosed as carcinoma in-situ that were in fact only dysplasias, metaplasias, or even mild atypias. This opinion rests largely upon unpublished data. Studies by Kottmeier [1] and

Peterson [2] showed that about one-third of carcinomas in-situ progressed to invasive cancers within approximately fifteen years. Since publication of these results, more cases have progressed, bringing the fraction of progressions in these series to above 80%.

In summary, then, the apparent incidence of regressive cases will depend on the policy of diagnosis adopted by a cytology laboratory.

## Effects on Dwell Time Distributions

The factors discussed above influence the dwell time distributions of progressive and regressive cases. We deliberately say that the distributions themselves will be influenced and not merely the estimates of the distributions. This is because we are not estimating an event with a precise, widely accepted definition. Rather, we are measuring the time it takes for the disease to progress from the earliest point at which a certain laboratory calls it carcinoma in-situ, until the point at which the disease would have been detected, using the medical system of the given locale.

The effects of the different points of disagreement on the dwell time distributions are as follows. If laboratory A diagnoses smears as carcinoma in-situ that laboratory B would call dysplasia, then we would estimate longer dwell times and more regressive cases in the population served by laboratory A. If women in population I visit their gynecologist less frequently than do women in population II, or if the gynecologists serving population I are less skilled at detecting cervical cancer clinically than gynecologists for population II, then we would estimate longer dwell times in population I. We also note, however, that clinically discovered cases were more severe in population I.

## Deficiencies and Errors of the Data

It is evident that if a screening program has been in existence for only five years, the longest possible interval between successive screenings of the same woman cannot exceed five years. Further, no cases can be detected clinically more than five years after a negative smear among women who have participated in the program. This means that one cannot estimate the number of progressive cases with dwell times of more than five years. Nor can one distinguish the number of cases (both progressive and regressive) with a seven-year dwell time from the number with a ten-year dwell time. One can only estimate the total number of cases with dwell times longer than five years.

The problem is aggravated by the probability that the laboratory's criteria for interpreting smears have changed. Thus it is unlikely that more than five or, at most, ten years of screening data will be self-consistent. Considering that the

average dwell time of carcinoma in-situ is probably fifteen years, it is impossible to estimate the entire dwell time distribution.

However, it is possible to estimate the number of cases with short dwell times. These cases are the most difficult to screen because to detect them screenings must occur very frequently. By contrast, cases with long dwell times--i.e. ten years or more--will be detected even if screenings are infrequent. Thus estimates can be made of those cases that will most strongly influence the design and performance of a screening program.

A second deficiency in the data arises because of the difficulty of determining whether a woman who is in a screening program has died for reasons unrelated to cervical (or other uterine) cancer, or has had a hysterectomy for reasons not related to cancer. In most countries, these data are not integrated with screening data; rather they are kept separately by hospitals or physicians. An enormous, time-consuming effort would be necessary to discover the follow-up information on each woman in a screening program.

Estimates of the dwell time distributions should be adjusted to account for the removal of these women from the population at risk. Otherwise, the incidence rates of cases with long dwell times will be underestimated. Since it is cases with short dwell times that matter, this data deficiency assumes a diminished importance.

Misdiagnosis at autopsy is a type of error that could distort the results. For example, a woman who had been earlier treated for carcinoma in-situ and later died, might be declared to have died of cervical cancer. Or, a woman with cancer and also a bad heart could be declared dead due to coronary problems, the cancer never having been reported. An uncritical acceptance of mortality data could distort the results in unpredictable ways.

Nevertheless, it is necessary to use these data. A woman who is found to have had cervical cancer, but who has died of some other cause, may be said in a sense to have had a clinically discovered cancer. In any case, the cancer was not found by screening. To leave these cases out of the process of estimating the progressive dwell time distribution would cause one to underestimate the incidence of progressive cases.

A final type of data error is mis-identification. A woman may be screened twice, but may appear in the records as two different women, each woman having been screened once. This could be due to a change of name at marriage, or a change of doctors, or to a misspelling by the doctor of the patient's name. This error will cause an error in the observed frequency distribution of screening intervals. That is, if a woman had two screenings four years apart, but was identified as a different woman on each occasion, the number of four-year screening intervals would be underestimated. Thus the incidences of cases with different

dwell times will be overestimated. However, the number of intervals cannot be underestimated by more than 5 - 10%, and the incidences should be overestimated by approximately the same amount.

## References

[1] Kottmeier, H.L. Am. J. Obst. and Gynec., 76 (1958) 243.

[2] Peterson, O. "Spontaneous Course of Cervical Precancerous Conditions." Am. J. Obst. and Gynec., 72, 5 (1956) 1,063.

## 1.3 The False-Negative Rate in Screening for Cervical Cancer

J. H. Bigelow and D. B. Ellis

This paper deals with the error rates for the Papanicolaou smear test when it is used as a screening test (as opposed to diagnostic test), and discusses how these rates can be reduced.

The success or failure of a cancer screening test is measured in terms of the percentage of false-positive and false-negative results. The false-positive smears do not present a real problem, since they will most likely be picked up in subsequent diagnostic procedures and be returned to the negative pool. It is the false-negative rate that is the critical error rate for any screening procedure, an error that can lead to a false sense of security, and worse, to delay in or omission of treatment. The false-negative rate can be estimated under a variety of conditions and by a variety of methods that are discussed in this paper.

The estimates of the false-negative rate will depend on the material used and on the method employed. A useful catagorization of materials is shown in Figure 1. The most direct, relevant source of data is Pap-smears obtained from a screening program, either taken alone or in the course of a more extensive examination. The latter is the more usual procedure.

Second, old Pap smears taken during a screening effort may be re-examined upon the later discovery of disease. These re-classified smears will provide a lower error estimate than would the smears as originally interpreted. However, the estimate would be of interest as a reasonable limit on the accuracy of the smear as a screening device.

Finally, data are obtained from clinical trials. These data are likely to provide the lowest error estimates, since they usually include a small number of smears, each performed and interpreted with the utmost care.

Possible methods for making estimates are shown in Figure 2. Direct methods (that is, methods that attempt to pinpoint the particular smears that are in error) include waiting for a suggestion of error to occur subsequent to a negative smear, or comparing the Pap smear with another test carried out at the same time. There is also an indirect, statistical method that we will describe later.

There are several instances of direct methods applied to clinical trial data. For example, the Mayo Clinic took nearly

— PAP SMEARS FROM SCREENING EFFORT
  - ALONE
  - COMBINED WITH OTHER TEST (E.G. COLPOSCOPY) OR GYN. EXAM.

— OLD PAP SMEARS RE-EXAMINED.

— DATA FROM CLINICAL TRIALS.

Figure 1. Material for false-negative estimation.

— WAIT FOR SUGGESTION OF ERROR (E.G. POSITIVE SMEAR FOLLOWING NEGATIVE).

— COMPARE PAP SMEARS WITH OTHER TEST.

— STATISTICAL METHOD.

Figure 2. Methods for false-negative estimation.

140,000 smears in the course of twelve years (Soule and Dahlin [17]), finding a false-negative rate of 2.4%. In this study, every smear--negative as well as positive--was followed up and verified or falsified by a biopsy. This procedure changed what would have been a modest screening effort into an enormous clinical trial. The rate is only of interest as the lowest possible bound, and is unrealistic when the Pap smear is used to screen large populations of symptomless subjects.

Another method of estimating the false-negative rate is to combine the Pap smear with another painless, safe and simple screening procedure, such as routine colposcopy, each method serving as a control of the other. For example, Limburg [10], using the smear technique alone, had a false-negative rate of 11.1% (based on cases found by colposcopy); using colposcopy alone, he had a rate of 3% (based on cases found by cytology). Combining both techniques, he had a false-negative rate of only 0.68% (based on two cases found by incidental curettage of the canal).

Richart [15] attained a low false-negative rate of 1.4% scoring a false-negative smear. The negative smear was obtained where dysplasia or carcinoma in-situ was found by colpomicroscopy on the day that this and/or the negative smear was bracketed by positive smears. Navratil et al. [11], as a result of screening 18,112 patients with both of these methods, found only 1.3% of all the cases that turned up had remained undetected in the screening trial.

These rates are undoubtedly low because the number of cases missed by both of these methods is surely underestimated. These cases are not sought out in these studies. Rather, the investigators must wait passively for them to come to light, a process that may require years. Peterson [13], for example, found clinical cases developing from known in-situ lesions--biopsied lesions at that--as long as fifteen years after discovery. It must be borne in mind that colposcopty can be used only for discovering carcinomas in the ectocervix, and will entirely miss carcinomas of the endocervix.

When the material used comes exclusively from Pap smears obtained in screening programs, the method for determining the false-negative rate is to wait for a later positive smear, or to wait for clinical symptoms or tissue studies to be done for a reason other than a suspicion of cervical cancer. Then the estimates of the false-negative rate vary from about 10% to as high as 30.6%. One commonly assumes that all cases of carcinomia in-situ that turn up subsequent to a negative smear make that smear falsely negative. Cases that may have begun in the interval between screenings, for instance, also fall into this group.

Often, when a case of carcinomia in-situ is discovered or suspected, the original smears are re-examined. This results in a "true" false-negative rate that is not due to observer error, this rate is lower than the rate that includes the observer error,

as well as cases that did not yield positive cells for the initial smear and new cases which developed in the interval. Cuyler [2] could reduce his 11.3% rate to 8.6% after re-examination of initial smears; Friedell, Hertig and Younge [5] revised their figure from 30.1% to 19% upon diagnosis of malignancy. Here such factors as prior knowledge of the histological findings can also contribute to further reduction of the error upon re-examination, as was the case in [5]. These revised rates are of interest only as limits on the accuracy of the Pap smear as a screening test, and not as estimates of the actual error rates to be expected.

Figure 3 summarizes the direct estimates of the false-negative rate. (Direct methods are those that attempt to identify those particular smears that are in error.) These methods can be criticized because when the data used are from screening programs (the most relevant source of data), there is no clear way to establish those smears that are wrong; thus the estimating procedure is itself subject to considerable error.

The third method, the statistical method, avoids this problem. It is based on the idea, illustrated in Figure 4, that successive screenings will exhaust the cases prevalent in the screened population at a rate that depends on the false-negative rate of the smear.

A group of women, prior to a first screening, will contain a backlog of cases that have yet to progress to invasive cancer. The first screening will detect a fraction of those equal to one minus the false-negative rate. The second screening will detect the same fraction of the cases which appear between the two screenings, plus that fraction of the remaining backlog. After many screenings, each new screening will detect the same number of cases as appear between successive screenings. (Some cases that appear between screenings will be missed, but their number will be made up from cases that were missed earlier and are detected by the present screening.) The change in the number of cases detected will be slow if the false-negative rate is high, because the initial backlog will not be depleted quickly. Conversely, a low false-negative rate implies a rapid change in the number of cases detected.

The model, first developed by Knox [8], is simple.

Let:

$P_i$ = the fraction of women screened who have pre-clinical disease just prior to their $i^{th}$ screening;

and

$D_i$ = the fraction of women screened whose pre-clinical disease is discovered by the $i^{th}$ screening.

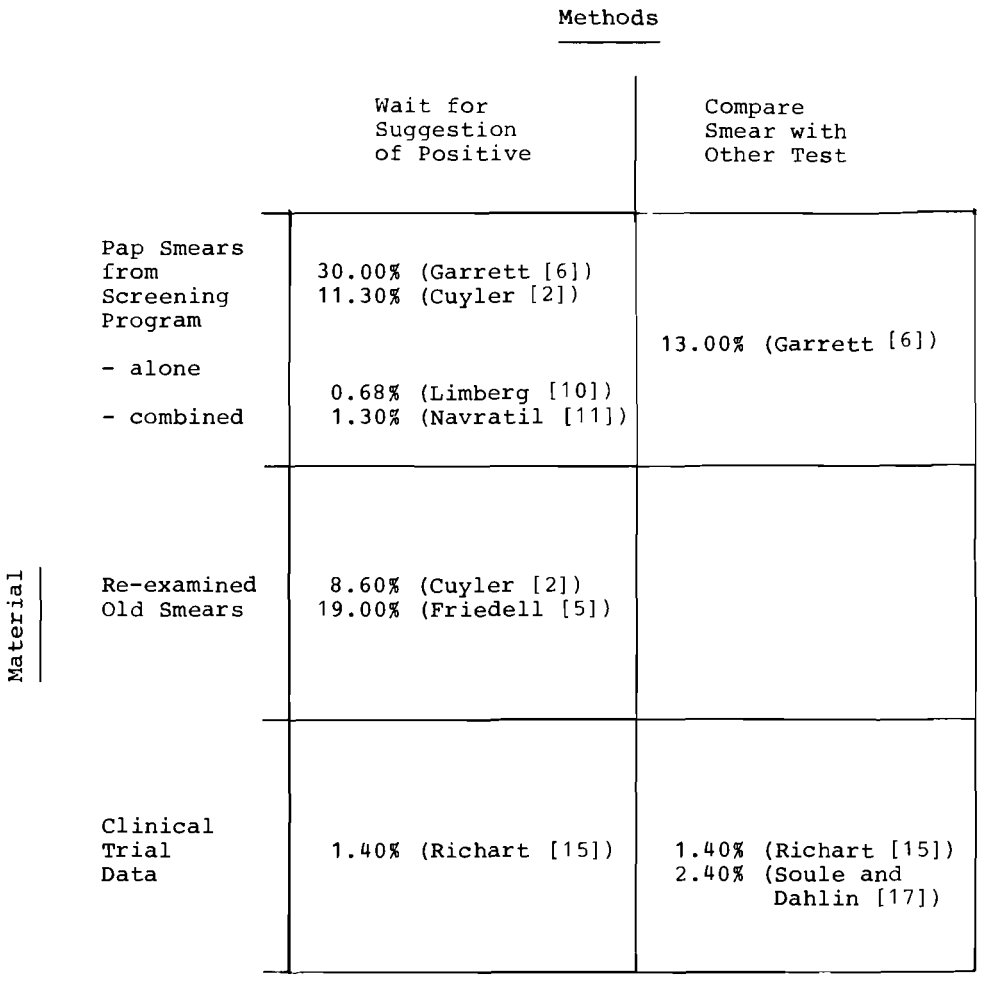



Figure 3. False-negative rates from direct methods.

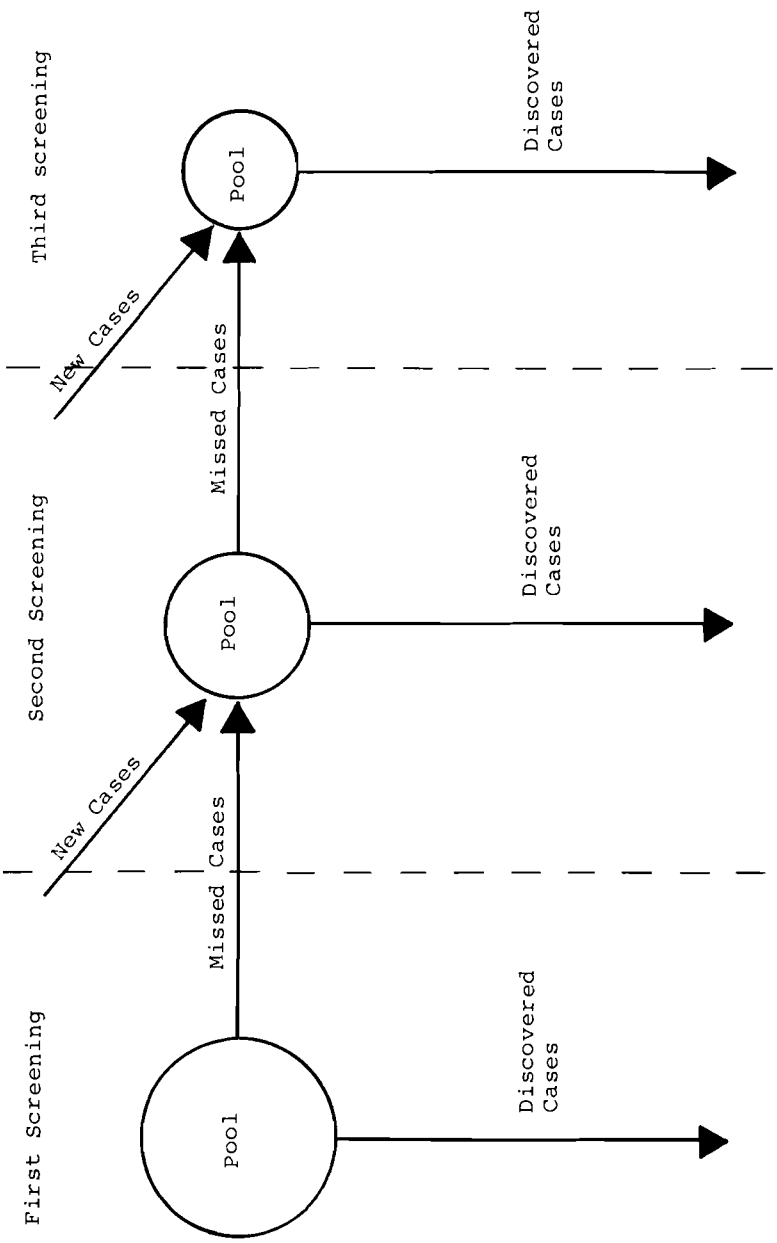



Figure 4. Basis of indirect estimation method.

Clearly, if $\rho$ is the false-negative rate, these quantities are related by:

$$D_i = (1 - \rho)P_i \quad . \tag{1}$$

Further, if:

$n_i$ = the average amount of time that elapses between the $i^{th}$ smear and the $i + 1^{st}$;

and

$\sigma$ = the incidence rate of the disease;

then:

$$P_{i+1} = (P_i - D_i) + \sigma n_i \quad . \tag{2}$$

In this model, the quantities that are observed are $D_i$ and $n_i$; the quantities we wish to estimate are $\rho$ and $\sigma$. The size of the prevalence pool, $P_i$, is an intermediate variable of only marginal interest. Ideally, both $D_i$ and $n_i$ should be calculated on an age-adjusted basis. However, insufficient information is given in our chosen source (Dunn et al. [3]) to "age-correct" the $n_i$; thus we have not bothered to adjust the number of detections, $D_i$.

The data for statistical estimation are summarized in Figure 5. From these data, and using this simple model, one can estimate a false-negative rate of 24%, and an incidence of thirty-six cases per 100,000 women-years.

The source of the error lies in the failures of cytology, observer errors, or in unrepresentative smears--"true" false-negatives. Both types of false-negatives can arise in the taking of the specimen and in associated clinical work. Much literature exists about the site from which to obtain the specimen with the greatest probability of finding diagnostic cells, as well as about the manner of collection (see Reagan [14]; Friedell, Hertig and Younge [5]; Ayre [1]; Seibels [16]; Parker et al. [12]; Gusberg [7]; Kulcsar [9]; Foote and Li [4]).

The rate of false-negatives is reduced when more than a single set of smears is used. It is established that carcinoma does not grow at a constant rate, nor does it grow constantly. It seems

| Screening Number | Number of Women Screened | Number of Cases Detected | Average Month to Next Screening |
|---|---|---|---|
| 1 | 33,746 | 350 | 16.93 |
| 2 | 9,109 | 26 | 14.16 |
| 3 | 3,995 | 4 | — |
| 4 | 1,586 | 0 | — |

Source: [3]

Figure 5. Data for statistical estimation.

logical to assume that the tumor does not constantly exfoliate cells and that it is possible to take a set of smears from a patient with a malignant tumor without finding malignant cells. Thus, the number of cell smears taken in each case, especially if more are obtained at interval examinations, is important to the efficiency of screening techniques, as are dilution, quantity and care in obtaining material and in preparing the slide. The representative cellular sample must include the squamocolumnar junction, wherever it may lie, and all cells must be fixed promptly to avoid distortion in drying.

Reading and interpretation of slides account for observer error. Accuracy standards play an important role for the cytologist in making the decision, for which no quantitative test exists. The problem is whether women with 3a or 3b dysplasia, the "suspicious" group, have carcinomas, and whether the smear is to be scored positive.

Since, from the viewpoint of the patient, there is little choice between "suspicions of malignancy" and "positive," it appears logical to place both grades in a single category. One can go further and include the significantly abnormal, dyskaryotic smears, in this category, with only "normal" and "atypical" smears being called negative. In this context, a false-negative error may be defined as a "normal" report on a smear containing cells that subsequent histology grades dyskaryotic, suspicious or positive.

Accuracy standards, problems in scoring smears and those arising in cytology itself (e.g. training programs, differences dependent on the site of the specimen, use of complementary colposcopy) are all factors to be considered with respect to false-negative rates. Each of these has an important impact on the resultant false-negative rate. However, two additional factors are the method and material used for estimating false-negative rate. Using the indirect, statistical method for estimating the false-negative rate described here, we can avoid problems arising when we try to determine those smears that are falsely negative.

## References

[1] Ayre, J.E. "Selective Cytology Smears for Diagnosis of Cancer." Am. J. Obst. and Gynec., 53 (1947) 609.

[2] Cuyler, W.K. Kaufmann, L.A., Carter, B., Ross, R.A., Thomas, W.L. and Palumbo, L. "Genital Cytology in Obstetric and Gynecologic Patients. A Four Year Study." Am. J. Obst. and Gynec., 62 (1951) 262.

[3] Dunn, J.E., Slate, T.A., Merrit, J.W. and Martin, P.L. "Finding for Uterine Cancer from One or More Cytologic Examinations of 33,750 Women." Journal of the National Cancer Institute, 23, 3 (1959) 507.

[4] Foote, F.W. and Li, K. "Smear Diagnosis of in-situ Carcinoma of the Cervix." Am. J. Obst. and Gynec., 50 (1948) 335.

[5] Friedell, G., Hertig, A.T. and Younge, P. Carcinoma *in-situ* of the Uterine Cervix. Springfield, Illinois, Charles C. Thomas, 1960.

[6] Garrett, W.J. "Symptomless Carcinoma of the Cervix." J. Obst. and Gynec. Brit. Cwlth., 71 (1964) 517.

[7] Gusberg, S.B. "The Relative Efficiency of Diagnostic Techniques in the Detection of Early Cervical Cancer: A Comparative Study with a Survey of 1,000 Normal Women." Am. J. Obst. and Gynec., 65 (1953) 1,073.

[8] Knox, E.G. "Cervical Cytology: A Scrutiny of the Evidence." In Problems and Progress in Medical Care (2nd Series). London, Oxford University Press, 1966, 279.

[9] Kulcsar, D.D. "A Comparative Study of Vaginal Smears and Cervical Cell Scrapings in Intra-Epithelial Carcinomas." Am. J. Clin. Path., 20 (1950) 958.

[10] Limburg, H. "Comparison Between Cytology and Colposcopy in the Diagnosis of Early Cervical Carcinoma." Am. J. Obst. and Gynec., 75 (1958) 1,298.

[11] Navratil, E., Burghardt, E., Bajardi, F. and Nash, W. "Simultaneous Colposcopy and Cytology Used in Screening for Carcinoma of the Cervix." Am. J. Obst. and Gynec., 75 (1958) 1,292.

[12] Parker, R.T. et al., "Intraepithelial (Stage 0) Cancer of the Cervix." Am. J. Obst. and Gynec., 80 (1960) 693.

[13] Peterson, O. "Spontaneous Course of Cervical Precancerous Conditions." Am. J. Obst. and Gynec., 72, 5 (1956) 1,063.

[14] Reagan, J.W. "A Cytologic Study of Incipient Carcinoma." Am. J. Clin. Path., 22 (1952) 231.

[15] Richart, R.M. "Evaluation of the True False Negative Rate in Cytology." Am. J. Obst. and Gynec., 89 (1964) 728.

[16] Seibels. "Cytology: A Diagnostic Method in Early Carcinoma of the Cervix." South. Med. J., 41 (1943) 706.

[17] Soule, E.H. and Dahlin, D.C. "Cytodetection of Preclinical Carcinoma of the Cervix: 12 Years' Experience with Initial Screening and Repeat Cervical Smears." Staff Meetings of the Mayo Clinic, 35, 18 (1960) 508.

## 1.4 Summary of Discussions on Error Rates in Cervical Screening

Most of the Workshop's discussion on cervical screening centered on the sources and remedies for errors. There was some limited discussion of the model used to calculate statistically the false-negative rate. Also, there was some discussion of the costs inherent in a false-positive smear.

### The Statistical Model

The model used to obtain a 24% false-negative rate from the data of Dunn, et al. [1] assumes that the false-negative rate is the same for each smear. This is untrue. The largest source of false-negative smears is unsatisfactory smears. A small percentage of doctors provide the bulk of unsatisfactory smears, simply because they are not careful or are not skilled at taking smears. Their patients will have a much higher false-negative rate than the entire screened population.

Further, it may be difficult to take a smear from a particular woman; this is most often the case for women in the post-menopausal phase. In addition, the probability of obtaining a false-negative smear may be higher if the lesion is very small.

The low ability of some doctors to take smears as well as the difficulty of taking smears from particular women can be dealt with by partitioning the data. For example, all smears taken by a particular doctor could be used to calculate a false-negative rate. Data from women over forty-two years (the approximate age of menopause) could also be used separately.

However, if the false-negative rate changes markedly from the earliest in-situ lesions to microinvasive cases, it is difficult to see what adjustments should be made. This question has been dealt with theoretically [2], [3], but to our knowledge no method has been developed for actually measuring the dependence of the false-negative rate on the extent of the lesion. This issue requires more consideration.

### False-Positive Smears

The cost inherent in a false-negative smear is widely understood. A potentially lethal disease will be allowed to remain undetected. The program, whose purpose is to detect these cases, in this instance will have failed.

A false-positive smear, however, will be followed up. The patient will be biopsied, and the absence of disease verified. Thus it would seem that false-positive smears are acceptable, and that no great effort should be expended to reduce their number.

But false-positive smears are costly. To be biopsied, a woman must travel to her doctor or to a hospital. The tissue sample must be examined. When it proves negative, a second, perhaps more extensive biopsy may be taken, possibly requiring some days' hospitalization. Finally, the false-positive smear will have caused the patient a great deal of unnecessary worry.

To some extent it is possible to trade off the false-positive rate against the false-negative rate. However, because interpreting the PAP smear has such a subjective quality, it is difficult to consciously choose a desired false-negative or false-positive rate. Thus changing both error rates relies largely upon quality control measures.

## Reducing Error Rates

Errors occur in cytological screening either in the collection and preparation of smears, or in the actual reading of them. Thus, if the person who takes the smears does not do so carefully, and does not fix the smear immediately, it is likely to be unreadable. If the cytologist is not well trained and experienced, or if his work is not checked by experienced cytologists, there will by many mis-read slides.

In general, physicians were thought to provide the least satisfactory smears. They regard themselves as overqualified, and tend to be careless in taking and fixing smears. Nurses make better smear-takers, as do midwives, who carry out this activity in Sweden. Technicians trained especially for this activity are the most reliable.

The cytology laboratory can help train the smear takers by the simple expedient of rejecting unsatisfactory smears. (These are the source of most errors, especially in laboratories that attempt a diagnosis from every smear they receive). When a smear is rejected, a repeat smear is usually requested, which entails only a small cost. It is certainly worthwhile to incur this cost in exchange for an improved repeat smear and a slightly better-trained smear taker.

Organization within a cytology laboratory can affect error rates. To be good, a laboratory must read at least 20,000 smears a year, and preferably 50,000 smears a year. A laboratory of this size will have enough people reading slides that each can check his interpretations against others. Five or ten percent of all slides should be re-examined. Every suspicious slide, even a mild atypia, should be viewed by many people. This sort of care and self-checking can do much to reduce errors in the reading of slides.

An important implication of these facts is that a good cytological screening program will take years to set up properly. The cytologists as well as the smear takers must have training and experience. A sufficient number of smears must be provided to justify good quality control measures in the laboratory. Given time, a program can reduce its error rates to very small values.

## References

[1] Dunn, J.E., Slate, T.A., Merritt, J.W., and Martin, P.L. "Finding for Uterine Cancer From One or More Cytologic Examinations of 33,750 Women." Journal of the National Cancer Institute, 23, 3 (1959) 507.

[2] Lincoln, T.L. and Weiss, G.H. "A Statistical Evaluation of Recurrent Medical Examinations." Oper. Res., 12, 2 (1064) 187.

[3] Weiss, G.H. and Lincoln, T.L. "Analysis of Repeated Examinations for the Detection of Occult Disease." Health Services Res., (Winter 1966) 212.

## 1.5 Summary of Discussions on Treatment

No paper was presented to the Workshop on the treatment of cervical cancer. A summary of the discussion of this subject is given below.

### Treatment of Dysplasia

Dysplasia is not a predecessor of cancer. Nevertheless, enough cases of dysplasia progress to carcinoma in-situ and invasive carcinoma that it is worthwhile destroying this lesion. This is especially true because treatment of dysplasia is safe, quick, and easy to carry out.

The oldest treatment for dysplasia is coagulation or cauterization. A newer method is cryosurgery, in which a cryode is inserted in the cervical canal and frozen by adiabatic expansion of $CO_2$ or nitrous oxide within it. Either method can be carried out without hospitalization, at little expense, and with a high probability of destroying the dysplastic tissue. Because these treatments are simple, it is justifiable to carry them out even where there is only a reasonable suspicion of dysplasia. One participant suggested this treatment even where the smear is "suspicious" and no biopsy is required.

As a result of the treatment of dysplasia there has been a reduction in the number of carcinoma in-situ and invasive cancer over the last twenty years. The Workshop suggested that IIASA's proposed study should examine this point.

### Treatment of Carcinoma in-situ

In large screening programs, most in-situ lesions are treated by conization. Properly done, and with a careful examination of the excised tissue to determine how completely the lesion has been removed, this procedure can be considered curative [1]. However, it should not always be the treatment of choice.

Hysterectomy is the treatment most likely to replace conization in some women with carcinoma in-situ. The conditions under which hysterectomy should be chosen are a) the woman has almost no contact with the medical system and would therefore be difficult to follow up, or, b) the woman has too many children and does not wish to bear more.

## Treatment of Invasive Cancer

Typical five-year survival rates for treated invasive cancers depend upon the stage of the cancer at the time of treatment. Campbell [2] reports typical five-year survival rates of 75% for Stage I cancer, 55% for Stage II, 30% for Stage III, and 10% for Stage IV. The opinion of participants at the Workshop was that there is little chance to improve these rates using surgery and radiotherapy. Nor does chemotherapy of advanced cases appear promising [3]. A combination of therapies, such as surgery followed by chemotherapy and immunotherapy, might prove effective.

## References

[1] Boyd, J.R., Royle, D., Fidler, H.K. and Boyes, D.A. "Conservative Management of in-situ Carcinoma of the Cervix." Am. J. Obst. Gynec., 85 (1963) 322.

[2] Campbell, H. "Cancer of the Cervix-Survival. Comparative Study of 5-year Survival Rates from Cancer of the Cervix in fourteen Countries from 1953-1957 Followed to 1962." J. Obst. and Gynec. Brit. Cwlth, 73 (1966) 27.

[3] Solidoro, A.S., Esteves, L., Castellano, C., Valdivia, E., and Barriga, O. "Chemotherapy of Advanced Cancer of the Cervix." Am. J. Obst. and Gynec., 94, 2 (1966) 208.

## 1.6 Summary of Discussions on Factors Associated With or Predisposing to Cervical Cancer

Dr. Stormby presented some information on this topic. Since Dr. Stormly's information was not presented in a workshop paper, a summary of his remarks is given below.

Over the last five years, an increasing number of Swedish teenage girls have been given cytologic examinations, usually because they have requested a prescription for the pill or for another kind of contraceptive (e.g. diaphragm or IUD). There are on file 13,200 smears taken between the years 1970 - 1973, from girls between fifteen and nineteen years of age. Within this group there have been twenty cases of histologically verified carcinoma in-situ, and some hundreds of cases of dysplasia, metaplasia and atypia.

It is interesting to observe the association between abnormal cytology and abnormal microbiology. The examinations revealed that very few cytologically normal girls had any flora besides the normal Döderline organisms found on their smears. A rather high proportion of girls with abnormal cytology exhibit microplasma and cocoid flora on their smears.

These abnormal organisms are the result of sexual activity, often with many partners, and poor sexual hygiene. Thus these data establish an association between early sexual activity and poor hygiene, and cervical cancer. There is also the suggestion that some celluar atypias may be explained by virus infections; this point has been suggested elsewhere [1].

Findings such as these suggest that the incidence of abnormal cytological findings, and possibly cervical cancer, could be reduced if physicians would treat these infections when they are found in their patients. This is not the practice. For example, when trichomonas infection is discovered, it is left without treatment in virtually every case. These findings also suggest that if gynecologists would instruct their young patients in proper sexual hygiene, the incidence of cervical cancer could be reduced considerably.

In the discussion that followed these remarks, the workshop agreed that poor hygiene and early and frequent sexual activity were important determinants of cervical cancer. Their presence lowers the age at which women can expect to contract cervical cancer. In India, for example, where poor hygiene and early marriage have been the rule, the peak incidence in cervical cancer occurs at an age ten years younger than the age of the peak incidence among European women.

Although the age at which women suffer cervical cancer is influenced by other factors, age itself is probably the most important factor. Women under thirty-five rarely have invasive cancer, and its incidence also drops off sharply among women over sixty-five. Other factors that appear to be associated with cervical cancer are probably factors associated with age, or early sexual activity, or poor hygiene. For example, the number of children will relate to age or to early sexual activity, and income or husbands occupation will correlate well with hygiene and the amount of regular medical care.

On the basis of these factors, the participants expected that the incidence of cervical cancer in developing nations would decrease, even if no programs were implemented to deal specifically with the disease. This decrease would result from improved hygiene (because of improved sanitation and education), later marriages and fewer children (to be expected as development proceeds), and more widely available, regular medical care. To some extent, general development would take the place of cervical screening programs in these countries.

The workshop observed that, since some of the factors predisposing to cervical cancer were behavioral, the groups at risk could change. The risk to young girls appears to be increasing due to their freer sexual activities. Thus a periodic reexamination of those segments of the population that are at greater risk is essential.

## Reference

[1] Beral, V. "Cancer of the Cervix: A Sexually Transmitted Infection?" Lancet (May 25, 1974) 1037.

# PART II. SOCIOLOGICAL AND ORGANIZATIONAL ISSUES

## 2.1 Factors Affecting Participation

H. L. Brown

### Purpose

The purpose of this paper is to establish why certain groups of women, namely the young, the old, and the poor, are least likely to be screened for cervical cancer, and subsequently to suggest certain policy recommendations that might be implemented to improve their low attendance rates. The problem to be faced is best illustrated in the form of graphs shown in Figures 1-4 [3]; [7]; [9].

### Two Types of Screening

The term participation itself says nothing about the necessary frequency of participation in a cervical cancer screening program. In order to be able to offer continuous protection from this type of cancer, the program must be organized as to ensure that the smears are repeated at regular intervals. Otherwise, the whole program may have no effect on the incidence and mortality rates of cervical cancer. In other words, a program can claim to have achieved a 100% attendance mark, in that every woman has been screened at least once, but may fail in not providing these women with the essential follow-up and repeat services. A differentation between two kinds of participation should therefore be made:

- regular screening at predetermined intervals including a coordinated follow-up system; and
- irregular, infrequent screening where a woman may have been screened but still runs the risk of developing cervical cancer and not having it detected in the earliest phases.

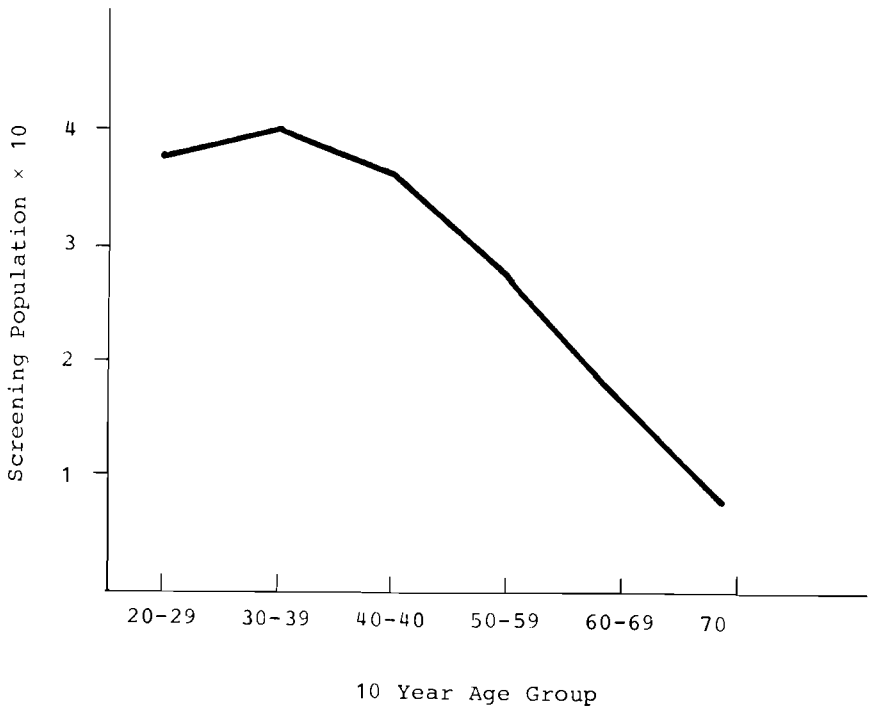
### Taking of Smears

An analysis of where and under what circumstances smears are taken indicates how embedded a cervical cancer screening program is in the existing health care system. In the US and in Canada, about 80-90% [7], [1], [4], of all smears are taken by some type of private physician. Significantly, the Pap test is often done in the course of examining some other complaint

| Age | 1966 Population | 1966 Total Screenings | Screening Population × 10% |
|---|---|---|---|
| 20-29 | 119,900 | 46,474 | 3.88 |
| 30-39 | 112,800 | 45,742 | 4.66 |
| 40-49 | 117,400 | 43,857 | 3.74 |
| 50-59 | 91,600 | 25,955 | 2.83 |
| 60-69 | 60,400 | 10,431 | 1.73 |
| 70 | 64,400 | 4,976 | .77 |

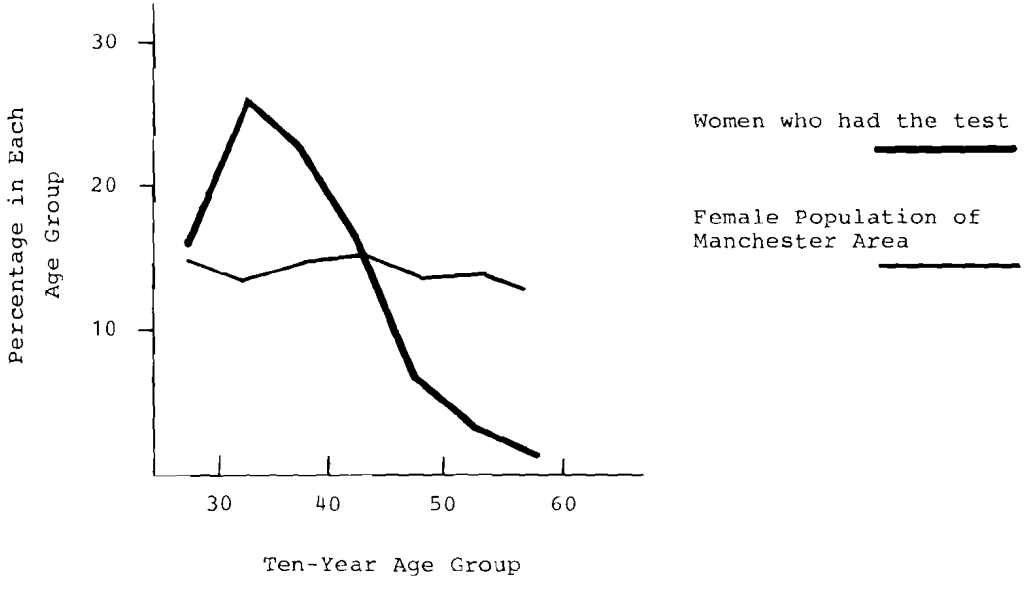
Source: [3]

Figure 1. Age dependence of screening rates.




Source: [3]

Figure 2. Age dependence of screening rates in British Columbia.




Source: [7]

Figure 3. Age dependence of screening rates in Manchester.

| 1,469 Respondents | | |
|---|---|---|
| Household Income (in Canadian Dollars) | Total Number of Respondents | Percentage of Respondents |
| under $3,000 | 295 | 47 |
| $3,000 - $6,000 | 802 | 67 |
| over $6,000 | 399 | 72 |

Source: [9]

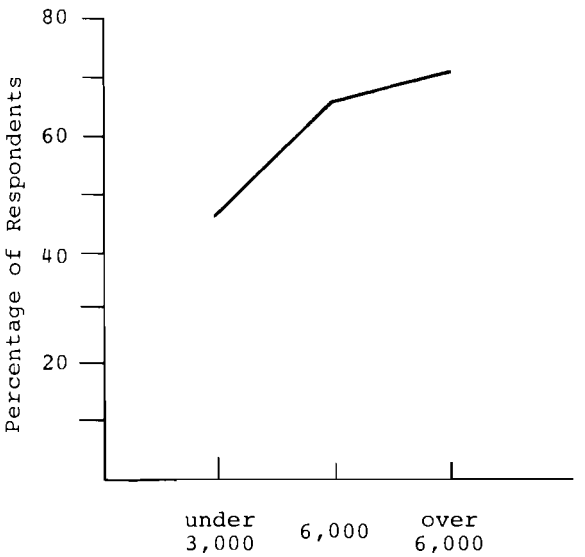



Figure 4. Percentage of women in each income group who had a test.

unrelated to cancer detection, and a woman agreed to have a smear taken mostly on the advice of her doctor. There are three drawbacks to the use of this method for detecting cervical cancer. First, it can only be effective, if the physicians' offices are visited routinely for examinations, secondly, if physicians who are visited routinely also take cervical tests, and thirdly, if the entire population of women is willing to visit physicians' offices for the tests.

Utilization of a private doctor's services, in the US at least, is a function of income; the wealthier can better afford it while the less well off cannot. Poor people depend to a very large extent (in some cases up to 80% of a given population) on their neighborhood clinic, the emergency ward, or outpatient department of a city hospital as their "general practitioner". This type of patient receives even less advice on preventive health care because of the more impersonal nature of most large clinics and because of the demand on time of the house officers. Yet it is precisely these people who are in the greatest need of such preventive care. This is reflected in the number of positive smears and the death ratios among these groups, shown in Figures 5 and 6 [2].

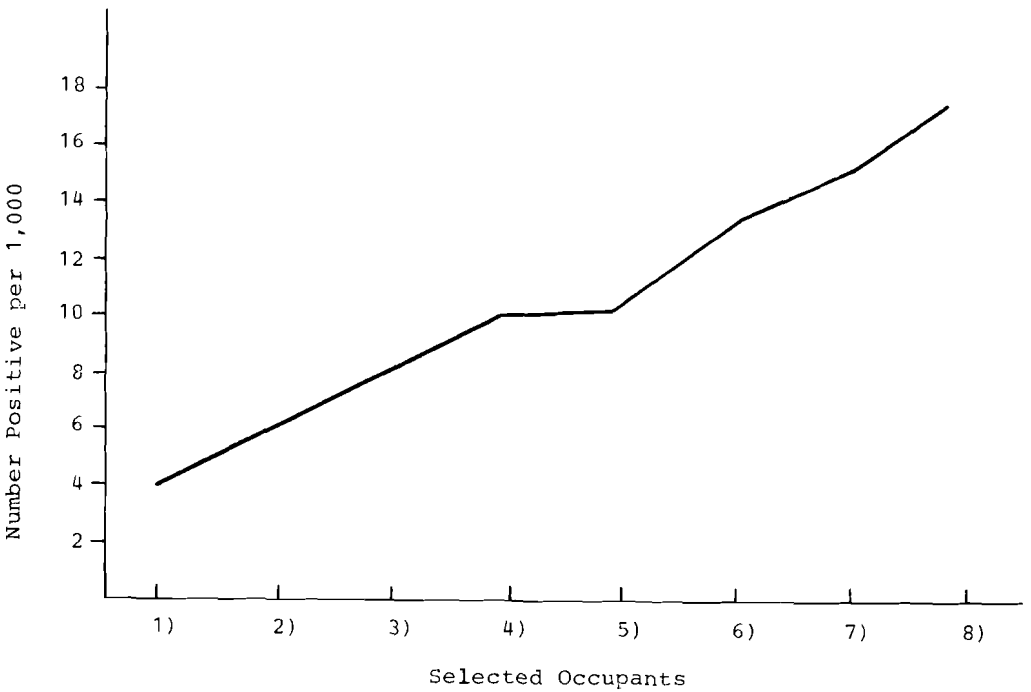
Hospital-based screening programs have achieved success [2]. These are, at times, bedevilled by external circumstances, such as the discontented attitude of the people meant to be screened toward the hospital. McGowan [5] has described a hospital based cervical cancer screening program in the US, and found that most people believed it was of importance that the hospital be of good reputation before they were willing to attend. But hospitals cannot satisfy everyone's expectations, and people are reluctant to enter institutions of mediocre or worse standards for any sort of care.

In the UK, the cervical cancer screening program is carried out on a national basis and intended for all women over thirty five. Doctors subsequently receive a payment from the National Health Service for cytotests performed on women of this age group. Despite the doctors being paid, there is not only poor penetration of screening in certain groups of the population, but more smears are taken in family planning or local health authority clinics than by family doctors.[1] (Figure 7) [7]. It is often alleged that the reason for this is attributable

---

[1]The data are from 1965 and there is a distinct trend toward more general practitioners taking cervical smears. Whether this trend has continued, I cannot say. Further investigation is required in this area.

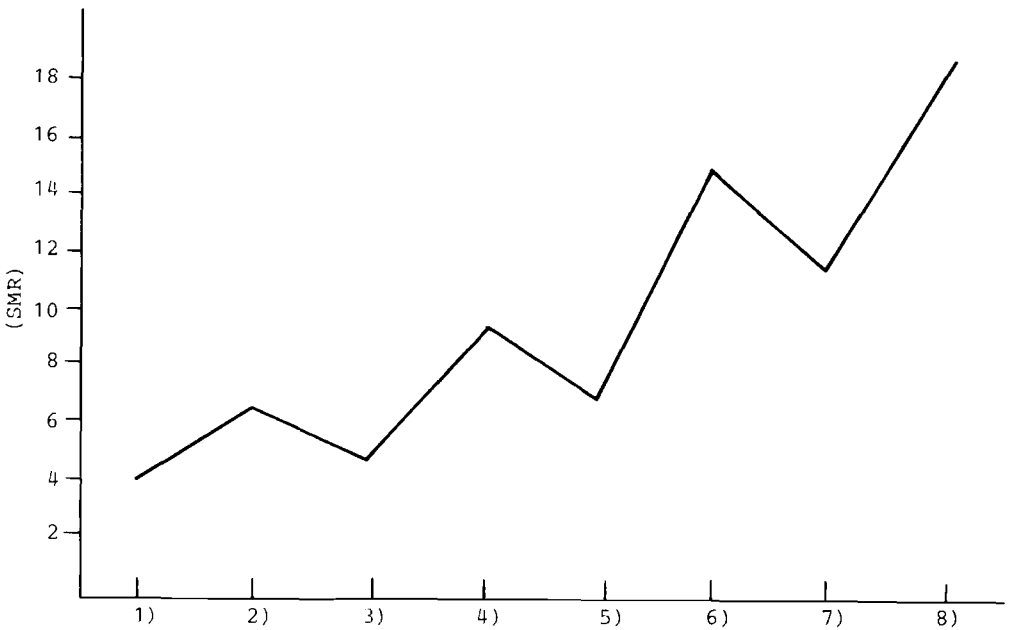
1) Professional workers
2) Clerical workers
3) Administrators and managers
4) Engineering and allied trades workers
5) Sales workers
6) Transport and communication workers
7) Consutruction workers
8) Laborers not elsewhere classified




Source: [8]

Figure 5. Number of positive findings by selected occupants.

1) Professional workers
2) Clerical workers
3) Administrators and managers
4) Engineering and allied trade workers
5) Sales workers
6) Transportation and communication workers
7) Consutruction workers
8) laborers and not elsewhere classified




Source: [8]

Figure 6. Standardized mortality ratios (SMR).

| Date | Total Number of Smears | GP | Source (Percent) | |
|---|---|---|---|---|
| | | | Clinics | |
| | | | FPA | LHA |
| 1963-1964 | 18,446 | 36.5 | 42.9 | 20.6 |
| May-July 1964 | 3,685 | 41.3 | 37.8 | 20.9 |
| Aug.-Oct.1964 | 3,799 | 44.9 | 34.4 | 20.7 |
| Nov.1964-Jan.1965 | 3,688 | 47.5 | 30.6 | 21.9 |

FIGURE 7. Percentage distribution of smears taken, Wakefield and Barić [7], comparing the pattern of the data for the period 1963-64 with those of three recent three-monthly periods.

to the smallness of the fee doctors can claim. The consequences of this particular weakness are revealing. A pilot study conducted in Manchester, UK, and aimed at women between the ages of twenty-five and sixty attempted to find out how women came to have a cervical test in the first place. Most of the women (60%) had the test as part of another examination, only 10% were advised by their doctor, and 30% asked their doctor personally to do the test [6]. 30% is a fairly high percentage in comparison to the Alameda County data [4] collected by Breslow where only 10% voluntarily and individually presented themselves for screening. The reason for this 30% figure may be not so much a greater awareness of English women of the benefits of cervical cancer screening but more the physicians' personal reluctance to suggest the test since the financial reimbursement appears unrewarding.

Follow-up and Repeat of Smears

In a cervical cancer screening program, it is imperative to make both doctor and patient aware of the need of repeat smears and follow-up. When exactly a repeat smear is recommended varies from program to program but it can almost universally be said that most women, across all ages and social groups, fail to return on the anniversary date of their previous negative smear. The reasons for this are multiple and complex, but certain interesting differences are revealed when knowledge of the need for periodic examinations is related to the way in which women come to have their first test. In the Manchester data [6], women who had the test at the family doctor's suggestion (although not as part of a routine vaginal examination) were best informed about the need for repeat tests (91.8%). Those who asked their doctor to do the test were also well informed (89.8%). Unfortunately, over one third of the women who had their first smear as part of another examination did not know of the need for further tests. Furthermore, of the women who were unaware of the existence of the test until they had it done, over three quarters were unaware of the need for later repeat smears. Some women, usually of lower income, were never told that they had had the test and others received no report about their results. These women are unlikely to further participate in the screening program.

The question arises who should have responsibility for recall? Leaving responsibility to the women themselves is impractical for the reasons just mentioned. A centralized recall system that would send out invitations for repeat smears is not the ideal solution since a letter alone has little influence on a woman making an action-oriented decision. The family doctor seems best placed to undertake recall although he would need help and cooperation from other screening services such as hospitals and clinics.

Main Causes for Low Attendance

The reasons why specific demographic groups do not attend cervical cancer screening programs lie not so much in attitudes and cognitive make-up toward cervical cancer in particular, but rather in the inherent disorganization of such screening programs and the restrictive manner in which they are integrated into the medical care system. A program almost totally based on the work of private physicians will never be able to reach people who only irregularly and infrequently visit a private physician's office.

The unavailability of adequate screening facilities for these population groups is striking. Specifically, the old and the poor are simply not offered a realistic opportunity to be screened. The costs involved including finances, time wasted on travel, and sitting in overcrowded waiting rooms, makes screening a tedious and bothersome exercise for these people.

Admittedly, the problem for these nonparticipating women is also, to a degree, one of inexperience and unawareness of the need to take preventive action. Once someone of the medical profession has advised them to have a smear and has clearly explained the procedure, the probabilities of their accepting or refusing the test are similar to those of any social or income group. At this point, very few women refuse the test.

Economics of Screening

Up to now I have happily ignored the problem of the economics of large population screening. A later speaker will have more to say about this than I can. But given certain unavoidable economic constraints in any cervical cancer screening program, the question must be raised whether it is essential that sufficient financial resources be found so that the program be geared toward the screening of all women of all ages and all income levels. Although it is quite unlikely that all women will present themselves for screening at a given time, if by some strange miracle this should occur, the whole screening program would be greatly overburdened and as a result collapse. It might be the unstated policy in some screening programs to, so to speak, "ignore" certain groups of women because the chances of their getting cervical cancer are small or the difficulties in having them tested are too great and expensive.

Policy Measures

The problem remains of what procedures can be altered or newly introduced to increase the overall attendance rates. In the United States, more hospital-based screening

programs could be formed since this is one of the major points of contact for the poor with the medical system. The difficulty becomes one of keeping these people in the program by getting them to return at a later date for a repeat smear; in general, the outpatient department of a city or county hospital rarely provides any sort of continuous medical care. Perhaps to compensate for this particular weakness, a trial program could be given a chance to establish itself in places or areas of employment where women are predominantly employed. This again would probably affect women of a lower income level more than any other and facilitate their entering the screening program by eliminating any inconvenience caused by travel, time, and additional expenses.

The screening of young women presents us with different sorts of problems. First, it is still uncertain at what age it is best to start taking cervical smears. Christopherson, for example, found that the screening of women under twenty is unrewarding; in his study, only a single carcinoma in-situ occurred in this age group while almost 20% of the total effort went into their screening [2]. Nevertheless it was thought to be a good educational procedure to initiate preventive medicine at such an early age. This proved not to be the case since these girls very rarely returned the next year for a repeat smear.

Women in the twenty-to-thirty age group are also not participating in the program. This same age group has the highest incidence of pregnancies and these women invariably seek medical aid. It is apparent, therefore, that if an antenatal and/or postnatal smear could be obtained from every patient, attendance rates would be improved. To an extent, this is the case in Manchester, UK, where about half of the smears of women under thirty five were taken during postnatal care [5]. As a comparison, in British Columbia in 1967 only 20% of the women in this age group knew they had a smear taken during their last pregnancy [9].

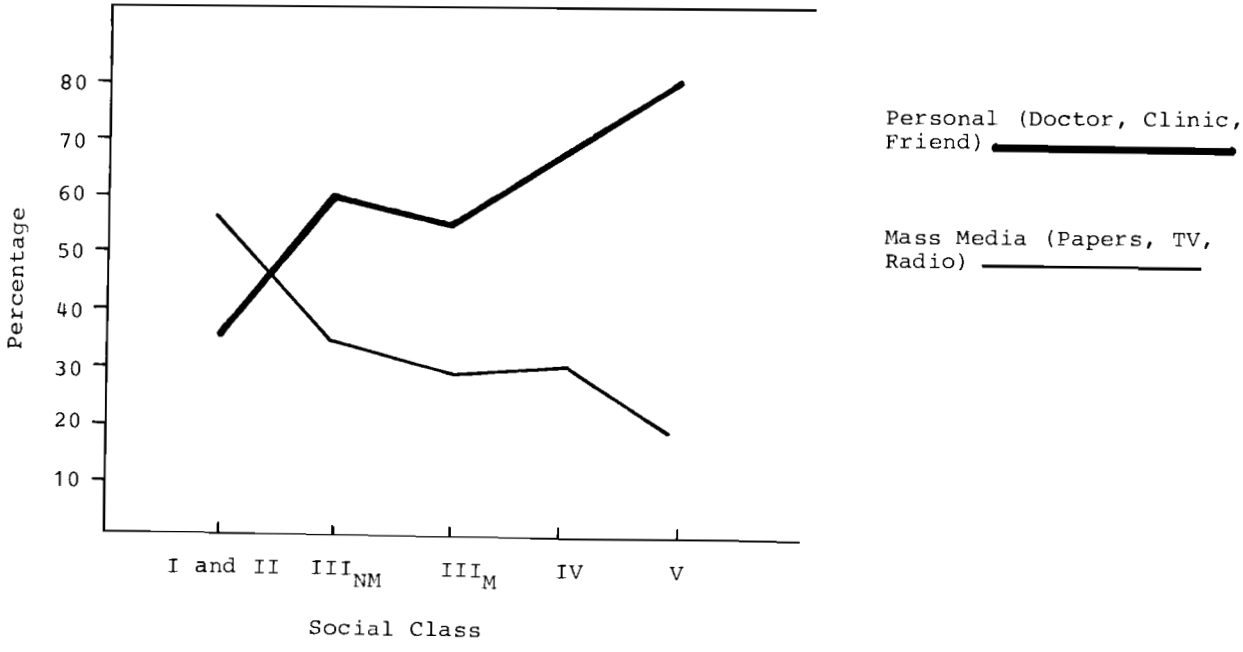
Older women represent a hard core resistance group. However the problem of their nonattendance is often brushed aside in the knowledge that they will eventually be displaced by the middle-aged group who will, by then, have been screened to a much larger extent.

Personal attitudes and opinions about cancer are undeniably important factors in a woman's willingness to participate in a cancer screening program. In the Manchester study, knowledge about the curability of cancer showed a wide variation between the social classes in the population of screened women; 90% of the women in the highest social class thought that cancer was usually or sometimes curable as against 53% of the women in the lowest social class [6]. The belief that cancer is incurable may deter people from taking a preventive measure such as a cervical smear, because it may show that they do have cancer.

Even revealing the presence of pre-cancerous conditions in the neck of the womb is taken to mean the same thing, because the cytotest has been too much publicized as a "cancer test". Given this, a belief in the seriousness of cancer, without a concomitant belief in the possibility of cure, is likely to deter women from having a cytotest.

The obvious conclusion is that some way must be found to change people's beliefs and attitudes. Ideally, a health education scheme could provide the relevant information. The objective of public health education in any preventive health program is to obtain individual, voluntary, preventive health action. As regards cervical cancer screening, it can be stated that, at present, this type of behavior seldom occurs. Therefore, what has to be determined is in what manner and form people should be made aware of the benefits of receiving a cervical smear. For example, if one considers the source of first time information about cervical cancer with social class, a sharply defined gradient appears (Figure 8). More women in the higher social classes hear about the test from the mass media than women of lower social classes who are more likely to acquire their information from so-called "personal sources". The lower the social class, the less likely women are to have first heard about the test from the impersonal mass media [6]. I do not think this is surprising, but it does indicate that a health education program increases its funding for more coverage of cervical cancer screening in newspapers, on TV or radio, will bypass exactly those women who it is trying to attract.

One fact that sociological studies have found consistently is that, among all the possible ways of influencing people to engage in any sort of behavior, the most effective way is through personal contact. The weight of evidence would suggest that, if one really wishes to increase the proportion of people seeking a preventive cancer check-up, this could best be done by addressing direct, personal requests to the individuals concerned.




Source: [6]

Figure 8. Sources of information analyzed by social class.

## References


[1] Breslow, L. and Hochstim, J.R. "Socicultural Aspects of Cervical Cytology in Alameda County, Calif." Publ. Hlth. Reps., 79, 2 (1964) 107-112.

[2] Christopherson, W.M. and Parker, J.E. "Control of Cervix Cancer in Women of Low Income in a Community." Cancer, 24 (July 1969) 64-69.

[3] Fidler, H.K., Boyes, D.A. and Worth, A.J. "Cervical Cancer Detection in British Columbia." J. Obst. and Gynaec. Brit. Cwlth., 75 (April 1968) 392-404.

[4] Kegeles, S.S., Kirscht, J.P., Haefner, D.P. and Rosenstock, I.M. "Survey of Beliefs about Cancer Detection and Taking Pananicolaou Tests." Publ. Hlth. Reps., 80, (Sept. 1965) 815-823.

[5] MacGowan, L. "Hospital-Based Cervical Cancer Detection Program." Obst. and Gynaec., 27 (June 1966) 791-794.

[6] Sansom, C.D., Wakefield, J. and Pinnock, K.M. "Choice or Chance? How women come to Have a Cytotest Done by their Family Doctor." International Journal of Health Education, 14 (1971) 127-138.

[7] Wakefield, J. and Barić, L. "Public and Professionel Attitudes to a Screening Programme for the Prevention of Cancer of the Uterine Cervix; a Preliminary Study." Brit. J. Prev. Soc. Med., 19 (1965) 151-158.

[8] Wakefield, J., Yule, R., Smith, A. and Adelstein, A.M. "Relation of Abnormal Cytological Smears and Carcinoma of Cervix Uteri to Husband's Occupation." B.M.J., 2 (1973) 142-143.

[9] Worth, A.J., Boyes, D.A. and Fidler, H.K. "The Acceptnace of the Cervical Cytology Screening Programme in the Province of British Columbia." J. Obst. and Gynaec. Brit. Cwlth., 74, 4 (Aug. 1967) 479-486.

## 2.2 Economic Aspects of Screening for Cervical Cancer

N. J. Glass

Economic analysis of health service projects is paralysed between two conflicting notions. On the one hand, resources for health service projects are limited, and the fear of the economic consequences of sickness and of early death is important. On the other hand, there is the feeling, widespread among the medical profession and the public at large, that pain and suffering caused by disease cannot, and should not, be measured in money terms.

Much of this ambiguity results from different concepts of the role of cost-benefit analysis. The limited concept of cost-benefit analysis applies this term to a procedure that attempts to trace a large number of the effects attributable to various methods of achieving a specified goal, and to characterise these effects in useful ways some of which will be in monetary terms. The wider conception of cost benefit analysis includes the above procedure, but also defines cost-benefit analysis as a means of choosing between methods of achieving a goal or even between goals on the basis of a cost-effectiveness or cost-benefit index.

There may be areas of public activity where any effects that cannot be valued in monetary terms are adjudged to be trivial. This is certainly not the case with health services.

Economists have insisted over the past ten or fifteen years on the "investment" character of many expenditures formerly regarded as "social" expenditure (e.g. education). They have stressed the contribution of such expenditure to economic growth (Bowman [2], Blaug [1]). However, the deficiencies of this approach are even more apparent with respect to health services. Much of health service expenditure is devoted to activities that have little or nothing to do with economic activity; the relationship between specific health services expenditure and prolonged active life is difficult to establish (Mushkin [6]).

It has been observed that calculations of increased productivity are likely to be of limited guidance to those responsible for the direction and volume of health services expenditure. This observation should not be confused with the notion that decisions on health services expenditure should be made on medical grounds or without any reference to the effect of alternative decisions on the use and availability of resources.

To say that a particular program is essential or that it should be available as a right "no matter what the cost", is to imply that other programs should not be carried out without a prior knowledge of what these other programs are.

The limited concept of cost-benefit analysis as described above is intended to make as explicit as possible the "opportunity costs" of various programs (i.e., the opportunities foregone by adopting certain programs). It does not claim that any form of quantitative analysis can decide those programs that should be adopted.

This paper discusses some of the problems encountered in a cost-benefit analysis of screening. Where possible, examples are given from cost-benefit analyses of cervical cancer screening programs; where necessary these are supplemented by examples from other screening programs.

## Average versus Marginal

A recurring point throughout this paper is that, for decision purposes, it is more important to know about marginal or incremental quantities than to know average quantities. This is best illustrated by the following example.

Let us suppose that a public health authority was thinking of launching a cervical cancer screening program, and that after due consideration, he decided to do this as a "one-shot" operation; and further suppose that the decision rule were: "We shall undertake this screening program provided the expected prevalence of in-situ carcinoma of the cervix is greater than 4 cases per 1,000, since the cost of screening a population with a prevalence of less than 4 per 1,000 is too high in relation to the benefits". Further let us suppose that an epidemiologist could provide (or invent) estimates of prevalence (see Figure 1). It can be clearly seen that the population prevalence of in-situ carcinoma of the cervix is 4.31. Should the public health department screen this population?

Closer examination reveals that although the average prevalence is 4.31, a prevalence greater than 4 is found only in four of the thirteen age groups. Extending the screening program to include women in the 50-54 age group for example would yield a prevalence of 3.92 cases per 1,000, below the threshold level of acceptance stipulated by the public health department.

It may be contended that the "savings" on the high-yield groups can be used to offset the "losses" on the low-yield group. But the decision rule of the department would appear to imply that it has more valuable ways of spending the money necessary to screen 10,000 cases if such expenditure is going

| Age Group | Prevalence Rate (per 1,000) |
|---|---|
| 20-24 | 1.05 |
| 25-29 | 3.23 |
| 30-34 | 5.55 |
| 35-39 | 6.73 |
| 40-44 | 6.11 |
| 45-49 | 5.10 |
| 50-54 | 3.92 |
| 55-59 | 3.98 |
| 60-64 | 3.85 |
| 65-69 | 2.39 |
| 70-74 | 2.98 |
| 75-79 | 1.80 |
| 80 and over | 1.29 |
| All ages 20 and over | 4.31 |

Source: [4]

Figure 1. Age specific prevalence rates of *in-situ* squamous carcinoma of cervix, 1960-66, Canada.

to yield less than forty cases. On this basis, only a program limited to women in the thirty to fifty age group would be justified, despite the fact that the population as a whole has a disease prevalence that exceeds the threshold.

To most of you here I am sure that my highly simplified example is unnecessarily belaboring an obvious point--namely, that a cervical cancer screening program may be justifiable for certain age groups but not necessarily for all. I have only made the point, with some diffidence, because two of the three fully worked-out examples of cost-benefit analyses of cervical cancer screening were done in terms of the whole population rather than discrete segments (Schneider and Twiggs [7], Dickinson [3]). Two other studies appear to have been carried out in average rather than marginal terms (Knox [5], Thorn et al. [9]). The only study that I have been able to find that explicitly mentions the importance of the chosen age groups in affecting costs and benefits was that carried out by an economist (Schweitzer [8]). Since he is an economist, Schweitzer limits himself to noting the importance of the age groups, and does not proceed to any further calculations.

In the example cited above, we held the marginal cost (i.e. the cost of detecting another case) constant and allowed the marginal benefit (i.e. the number of new cases) to vary over age groups. It is possible that the situation might be the opposite of this. The age prevalence of the disease might be constant in all areas of a given geographical region. However, the cost of detecting a case might rise as one moved from urban to rural areas. Thus while the average benefit might be above the average cost for the region as a whole, it might fall below this cost for the rural areas.

This concentration on average quantities can lead to apparently contradictory conclusions. Fidler et al. [4] for example, suggests that, on the basis of mortality figures shown in Figure 2, a cervical cancer screening program is justified in Canada; however, it "would not be worth while" in Israel since the mortality rate in Israel is about one-ninth that of Canada. The figures that they quote on the incidence of <u>in-situ</u> carcinoma in Canada show that for women in the age group of 65-69, the incidence of the disease is less than one-tenth that of the age group 25-29. These data show that: a) a screening program is not worth while for Canadian women aged 65-69; or that b) it is worth while for the Israeli women in the 25-29 age group. This assumes that there is a similar age-prevalence distribution in the two countries.

| Country | Rate |
|---|---|
| USA (non-white) | 181.7 |
| Denmark | 111.5 |
| UK (England and Wales) | 108.4 |
| USA (white) | 95.2 |
| Scotland | 88.6 |
| Canada | 82.2 |
| Australia | 68.5 |
| France | 35.1 |
| Sri Lanka | 19.9 |
| Japan | 14.9 |
| Israel (Jewish) | 8.7 |

Source: World Health Organization, 1959.

Figure 2. Death rates per million, females, carcinoma of cervix, 1952-56.

The Full Costs of Screening

The cost of a screening program does not include only the screening test needed to determine the presumptive existence of an abnormality. It should also include: a) costs of inducing a given population to participate in the program; b) costs of confirmatory tests and subsequent treatment; c) costs to the patient in terms of time away from her normal routine; and d) administrative costs of the program.

The question of participation has been dealt with in previous discussions at the Workshop, and it shall confine my remarks to the costs of inducing participation. Participation, defined to include both the proportion of the population screened and the frequency of participation, might be increased in a number of ways. Methods might include changes in organization so that women would more frequently come in contact with screening agencies; changes in location of screening facilities; the use of tests that are more acceptable to the public; increased health education programs; subsidies to patients and so forth.

The cost of inducing participation is unlikely to remain constant as one tries to move nearer to 100% participation. Changes in organization, education, etc. will almost always involve a cost, either financial or in terms of the efficiency of the test. As one attempts to get all the population to come forward at the appropriate rate, the cost is likely to rise continuously.

Figure 3 shows how the cost curve is likely to look. A program organized around ante-natal and post-natal visits might achieve a slow growth rate of participation at a relatively low cost. As one wishes to increase the rate of participation and to achieve this rate of participation more quickly, the cost of attracting these extra participants is likely to greatly exceed the cost per head of attracting the earlier participants i.e., the marginal cost of participation is likely to increase. Attracting the poor, the old and other recalcitrant groups may turn out to be a relatively expensive business.

The Problem of Time

In an evaluation of screening programs, timing is of the essence. That is, the justification of screening lies in the ability of earlier intervention to avoid or postpone later diseases. In economic terms, screening consists of incurring an earlier cost in order to postpone or avoid a later cost. Thus there must be some way to compare earlier costs with later costs.

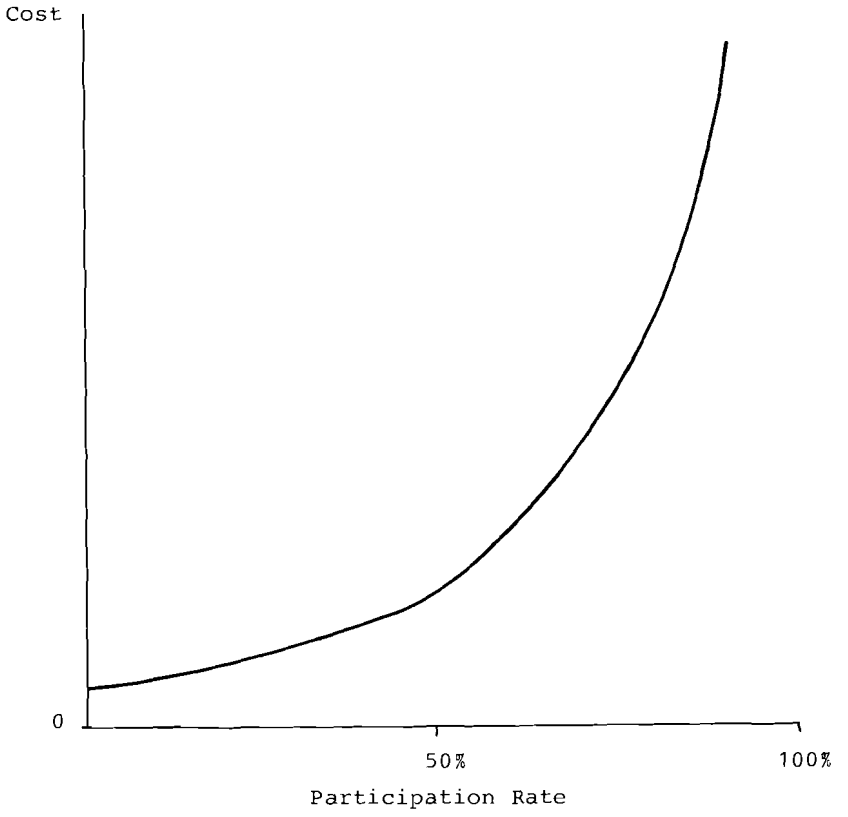



Figure 3. Costs of inducing participation.

Most economic studies of cervical cancer screening programs have cited the avoidance of hospital care for patients with invasive cancer as one of the principal economic benefits to a community to offset against the cost of the screening program. After a cervical cancer screening program had been introduced, and the participation rate gradually increases, the lag time between costs and benefits should gradually increase. This is because, given a lengthy time interval for development of carcinoma in-situ, much of the prevalent stock of cases that would be detected in the early years of the program would have a shorter lag time to invasive cancer than the ten to fifteen year average frequently suggested. As the program approached 100% participation, the costs incurred would be preventing costs that would have been incurred in ten to fifteen years time.

In cost-benefit analysis, costs and benefits occurring at different points in time are normally made commensurate by means of a discount rate. The discount rate is a factor that attributes a lower value to any cost or benefit the longer it is delayed; it takes account of the fact that resources can earn a return over time and that a lower sum is therefore needed to yield a given sum in, say, ten years time. The method by which a discount factor is chosen and its exact mathematical form are not of interest here. However, if a screening program is financed out of public funds, it would seem reasonable that the same factor that is used for other government programs should also be used in order to ensure consistency among programs.

The studies by Schneider and Twiggs [7] and by Dickinson [3] do not use any form of discounting procedure, i.e. they apply implicitly a zero discount rate to all costs and benefits. It is doubtful if in private or business life, responsible persons are prepared to lend sums of money for ten to fifteen years at zero interest! Government investment programs in the UK are supposed to apply a 10% discount rate. The use of such a discount rate significantly reduces the value of costs and benefits that are delayed for ten to fifteen years.

While a zero discount rate is applied to the costs of those persons who contract cervical cancer, an (implicit) infinite discount rate is applied to the hospital costs of those who are prevented from contracting the disease. That is the studies cited do not take account of the fact that terminal hospital care is eventually necessary even for those who are prevented from dying of cervical cancer. These patients do not live forever. If the purpose of an economic analysis is to clarify the costs to the community of preventing cervical cancer, it should be pointed out that a screening program at best postpones terminal hospital care; it does not eliminate the need for it. The discounted cost of each postponed terminal care is greatly reduced by the postponement, but it does not become zero.

## Equity versus Efficiency

A final point I wish to make concerns a possible conflict between the efficiency (narrowly defined) of a screening program (e.g. getting as large a yield as possible for a given expenditure or some other rule of this kind), and the notions of fairness (e.g. that everyone should have an equal chance of being prevented from contracting cervical cancer). Given the fact that certain classes of the population are less likely than others to come forward for a program, the option will always exist that for a given expenditure one might sacrifice a higher yield to ensure that the benefits were more evenly spread. One might, for example, increase the screening interval and use the resources saved to set up mobile clinics in areas of social deprivation.

The question of fairness is a difficult one; the most one can say is that it is unlikely that any screening policy would be concerned only with maximizing yield or minimizing cost. More likely, a screening policy would set some constraint on the distribution of benefits or would attach some value to reductions in the differences in the distributions of benefits between social classes.

The point to be stressed is that, once a decision has been made about efficiency, it will almost certainly imply a decision about the distribution of the benefits or the fairness of the scheme. It is unlikely that these two factors can be separated. It is likely that the price of increasing one of the factors will be a decrease in the other.

## Conclusions

This paper has attempted to set a number of hares running in order to provoke discussion; it is hoped that we do not end up chasing different hares. The measurement of health service benefits in economic terms has provoked disbelief; but this should not provoke the opposite reaction, i.e. that health services should not be discussed in such a way as to frighten the horses. The importance of incremental magnitudes for decisions has been stressed; the implications for choosing the target population, the optimal participation rate and the appropriate manpower have been mentioned. The paper has also noted the importance of time and of discounting, and has pointed to a possible conflict between efficiency and equity in screening programs.

## References

[1] Blaug, M. An Introduction to the Economics of Education. Harmondsworth, UK, Penguin Books, 1972.

[2] Bowman, M.J. "The Human Investment Revolution in Economic Thought." Sociology of Education, 39 (1966) 111-38.

[3] Dickinson, L. "Evaluation of the Effectiveness of Cytologic Screening for Cervical Cancer." Mayo Clinic. Proc., 47 (Aug. 1972) 550-555.

[4] Fidler, H.K., Boyes, D.A., and Worth, A.J. "Cervical Cancer Detection in British Columbia: A Progress Report." J. Obst. and Gynac. Brit. Cwlth., 75 (1968) 392.

[5] Knox, E.G. "Cervical Cancer" in Screening in Medical Care." Nuffield Provincial Hospitals Trust, Oxford, London University Press, 1968.

[6] Mushkin, S. "Health as Investment." Journ. Polit. Econom. Supplement, October (1962).

[7] Schneider, J. and Twiggs, L.B. "The Costs of Carcinoma of the Cervix." Obstetrics and Gynaecology, 40, 6 (Dec. 1972) 851-859.

[8] Schweitzer, S. "Cost Effectiveness of Early Detection of Disease." Health Services Research (Spring 1974), 22-32.

[9] Thorn, J., Swanson, K., MacGregor, J.E. and Russell, E. Unpublished, cited in K. Randall "Cancer Screening by Cytology." Lancet, 30 (Nov. 1974) 1303-4.

## 2.3 Summary of Discussions on Sociological and Organizational Issues

There was much discussion at the Workshop of the causes of the unsatisfactory results often obtained in cervical cancer screening programs. Participants stressed the need to educate doctors, to have laboratories refuse unsatisfactory smears, and to have a laboratory sufficiently large to ensure high standards. A turnover of between 20,000 and 50,000 smears per year was suggested as the optimum range.

It was made clear that for a mass screening campaign to be carried out, the campaign must take place with the cooperation of the existing medical system; withouth such cooperation the organization of a campaign would be prohibitively expensive. The brief flirtation with patient-administered tests has been abandoned due to the unsatisfactory specimens obtained; however one participant reported of a new trial of a patient-administered test that was to begin shortly in Denmark.

Because of difficulties of attracting the old and the poor to participate in a cervical cancer screening program, the Workshop strongly recommended that every woman who came to a hospital either as an inpatient or an outpatient should have a smear. Otherwise, it was noted, one third of women over forty-five would never have a screen.

The participants agreed that although older women were harder to attract, once they were attracted to the program, they tended to return. Younger women, on the other hand, tended to be much less conscientious about returning.

One discussant referred to the difficulties of attracting immigrants to participate in screening programs. There was a lengthy discussion of the difficulties of measuring the benefits of cervical cancer screening programs. Opinions were divided as to the appropriate place to accord measures of loss of productivity due to cervical cancer. All discussants agreed that health ministries were, in general, interested only in the costs of the screening program and in the hospital costs of cancer patients. The majority of discussants believed that ministries of finance were interested in the economic effects of avoiding cancer.

The differential costs of screening different groups of the population was emphasised. One discussant presented figures showing how costs varied between age groups and also between demographic groupings as a result of the different incidence of the disease. The rising cost of inducing very high levels of

participations was also mentioned. Another discussant stressed the importance of having a program as widespread as possible since this would increase participation rates through personal contact.

Dr. Christopherson presented to the Workshop his view of the essentials of a program for the evaluation of the effect of mass screening on morbidity and mortality of cervical cancer. These criteria, as given below, summarize the Workshop's discussions of the organization of a screening program.

## Essentials of a Program for the Evaluation of the Effect of Mass Screening on Morbidity and Mortality of Cervix Cancer

I. General Requirements

1. A population of previously unscreened high risk (forty years of age and older), geographically defined women.

2. Community resources and interest to pursue a study for at least one generation.

3. A population-based uterine cancer registry to establish base line age-dependent incidence rates for a period of 3-5 years prior to initiation of the screening program (this can be done retrospectively).

II. Specific Personnel and Facilities

1. Competent cytopathologist and supportive personnel:

   a) adequate laboratory facilities for both cytology and pathology that can rapidly expand to meet growth needs.

2. Competent gynecologist with an interest in uterine cancer:

   a) outpatient and inpatient facilities to carry out the diagnosis, treatment and follow up of the patients.

3. Competent radiotherapists and supportive personnel:

   a) sophisticated equipment for the delivery of therapeutic irradiation.

4. Epidemiologic statistical and clerical personnel to maintain records, to rapidly retrieve data, to follow up patients, and to operate a population-based uterine cancer registry:

   a) Adequate space and data processing equipment.

III. Collaborative Personnel, Voluntary and Governmental Agencies.

1. The program must have the input and support of the organized medical community.

2. Any existing voluntary organizations with an interest in the problem should be involved from the onset, e.g. cancer societies, women's groups, religious institutions, public service groups.

3. The above (1+2) are necessary for an effective public education program to insure screening of 80-90 percent of the population at risk over a limited period of time, and to maintain interest in the program after the initial thrust has subsided.

4. The governing body of the community and/or country should commit financial support as needed, and give assurance that the long-term study will be completed.

IV. Continuing Professional Education

1. There must be a continuing dialogue among involved personnel to insure good quality of cytologic-pathologic correlation (false-negative and false-positive smears).

2. The adequacy of biopsy material must be continually analysed and controlled.

3. A tumor board, consisting of a cytopathologist and/or pathologist, gynecologist, radiotherapist and epidemiologist should meet regularly to discuss each case in detail, thus providing a forum for their own education and that of students, interns, etc.

4. Presentations at appropriate intervals should be made to the medical community to sustain their interest and to improve their performance.

V. On-Going Evaluation

1. A critical evaluation of the program should be made once a year by outside consultants.

2. A semi-annual print out of results should include:

   a. number of women with biopsy recommended;

   b. number of women with biopsy performed;

   c. a list of women not brought to diagnosis;

   d. rates of proved cases by specific diagnosis and age. Where indicated additional groupings by race, religion etc. could be made;

   e. treatment data; and

   f. data on recurrence and survival.

   The above is the minimal information necessary to evaluate the quality of the program. It must be kept in mind that without rigid control of the quality through all steps of detection, diagnosis, treatment and follow up the program could conceivably result in more harm than good.

3. Other factors to be evaluated at regular intervals:

   a. cost effectiveness ratio;

   b. work load of personnel; and

   c. ability of program to reach those women at high risk.

VI. Final Evaluation

1. Was the program successful in reaching almost the entire female population?

2. Was the morbidity and mortality for cervix cancer significantly reduced in the geographic population area over a 10-20 year period?

3. If so, was the cost effectiveness ratio acceptable to the economy of that particular area?

## Part III. Cervical Screening Research Plan

### 3.1 Design of IIASA's Proposed Cervical Screening Study

J. H. Bigelow

At this last session of the Workshop, I would like to weave the ideas that have been expressed into the plan for IIASA's proposed cervical screening study.

#### Purpose

It is worthwhile recalling that the purpose of our study is to provide guidance for cervical screening policy. Thus we expect to give preliminary answers to questions such as the following:

a) What are the costs (e.g. hospital bed-days, physicians' time, laboratory facilities) and the benefits (e.g. reduced mortality and morbidity, additional years of life) of a screening program? How can they be measured?

b) How does one design a screening program in order to maximize some measure of benefits while keeping the different costs of the program within specified limits? What parts of the population should the program attempt to screen? How frequently? What resources should be expended on recruiting people to be screened? What resources should be expended on following up cases whose screening test shows a positive result?

c) How sensitive is the performance of the screening program to changes in program design? What benefits would result from increasing the resources available for screening? How much of one benefit (e.g. reduced mortality) must be given up in order to obtain a given amount of another benefit (e.g. additional person-years of life)? (This might be accomplished by screening less intensively among the old, and more intensively among the young and middle aged, who, if prevented from dying of the disease, will live longer.) Might some of the resources be better devoted to research, e.g. to make the screening test more sensitive and/or specific, or to improve the treatment for invasive cancer?

d) How shall the program be implemented? What resources are required initially to train needed personnel (e.g. cytologists) to cope with the increased patient load?

How rapidly will the cases that are prevalent in the population be exhausted due to the screening activity? What are the consequences of a gradual increase in the level of screening, as opposed to a rapid increase?

e) What should be the relationship between: a cervical cytology program and the medical care system; and a cervical cytology program and other screening programs; and a cervical cytology program and the health insurance system?

Method

The basic strategy to be applied in this study was described in the Introduction to these proceedings. It is to relate three kinds of information, first by means of an optimization methodology, and second, by means of simulation techniques. The three kinds of information needed are:

a) information about the design of the screening program, including who takes the smears (e.g. doctor, nurse, or technician), where are they taken (e.g. doctor's office, hospital clinic, mobile unit), what procedures are there for recruiting women to be screened (e.g. letter, personal visit, wait until women sees her doctor), etc.;

b) information about the composition of the population to be screened, by age, race, income, and other relevant factors; and

c) information about the effects the screening program will have on the given population, including the program's cost, its manpower needs, its effect on mortality and morbidity, etc.

Simulation techniques will predict the effects of the program from its design and the population it serves. Optimization techniques will calculate the least costly design that will achieve the desired effects while serving the particular population.

In order to optimize or to simulate a screening program, we will need to know the natural history of the disease. Thus three methodologies combine to form the basis of our study: statistical estimation of the natural history; simulation of screening programs; and optimization of screening programs. I have described our estimation methods earlier, and so I need not do so again. But I will outline briefly the other two methodologies.

Optimization

The purpose of the optimization model is to determine the best screening policy to adopt as a function of the population to be served, as well as the resources available for screening. This model will not consider problems of time-phasing, such as the capital investment needed in training facilities or the preparation necessary to convince the population to participate. Rather, it will be assumed that the program has been in operation for many years, and that the composition of the population, and the prevalence and incidence of the disease, have reached their steady-state values. Thus, this model will chose only the best steady-state situation.

We will chose some measure of benefit to maximize, while limiting the amounts of the different resources available. Thus we might maximize the reduction in mortality, requiring that a limited amount of the physician's time be taken by screening activities, or that a patient not be required to travel more than ten kilometers to receive her test, or that the total screening budget not exceed a certain number of dollars.

The optimal design of a screening program will depend on the medical environment in which the screening program is implemented. For example, in a place where people are medically served only by a few large hospitals and clinics, to set up small, neighborhood screening facilities would be very expensive. Where neighborhood clinics already exist, the screening test could be offered there at little additional cost. It might prove optimal in the first case to provide a few mobile screening facilities, housed in large trucks, while, in the second case, it would probably be better to dispense screening tests through the existing clinics.

Some data on resource requirements are available in the published literature (e.g. Dickinson [1], and Schneider and Twiggs [4]. Other information - largely qualitative but still quite valuable - is available in these proceedings. Benefits, on the other hand, will be calculated using our own model of the disease, possibly in a manner similar to Knox [2], [3], but probably more direct. Other considerations involved in calculating costs and benefits have been discussed by Glass (see II.2).

Simulation

Our simulation model will help us to investigate certain questions of implementation of screening programs. These questions include:

a) How, and how quickly, are the necessary resources (e.g. cytologists and cytology facilities) to be mobilized?

b) How quickly are efforts to attract participants in the program to be implemented?

c) What will be the changing needs of the program from the first few years when it is dealing with the backlog of prevalent cases to later years when it is locating only the incident cases?

d) What will be the impact on the program of variations in the incidence of the disease or participation in the program?

Questions of this type are not dealt with in the framework of the optimization problem because to do so would result in the model being too large. Instead, we will identify prefered policies using the optimization model and assuming a steady-state (hence constant) participation, level of screening effort, and disease incidence. In order to explore possible difficulties of arriving at these policies, as well as potential problems in returning to a steady-state following a perturbation, we resort to a simulation approach.

Our simulation method is a fairly common one. We will simulate one case at a time. From the starting point (e.g. age twenty) we will step forward in intervals of some convenient length (e.g. three months) until the subject has died or has reached a maximum age (probably seventy years); or until a case of cervical cancer is detected by screening or in the clinic. In each step we may ask whether a screening test has been done, and if so, whether it is "valid". (That is, if the woman were to have carcinoma of the cervix, this test would detected it.) If a valid test has been performed and the subject is diseased, the simulation of this case terminates with a cancer detected by screening. If the test is not valid or if no test has been performed, and if the subject has cervical cancer, we ask whether the case will be detected clinically in this time interval. If so, the simulation terminates with a clinically detected cancer. Otherwise, the remaining dwell time of this case is reduced by the length of the interval, the patient's age is increased by the length of the interval, and we are prepared for the next step forward.

Finally, we deal with the possibility that the patient does not have cervical cancer at this time. We then ask whether she will contract the disease during the interval. If she will not, we advance her age; we are ready to step forward again. If she will be diseased, we decide if her case will be progressive or regressive, and we select a dwell time; we advance her age, and proceed to the next forward step.

As the simulation proceeds, the number of women screened and the frequency with which they are tested may change. So may the false-negative rate, for example, due to the increasing experience of the cytology laboratory personal. Cases prevalent in the population at the beginning of the program will be exhausted, and the yield will drop to the number of cases that occur in the intervals between screenings. Thus, as the simulation proceeds, we will observe time streams of both costs and benefits, that must be compared to determine the cost-effectiveness of a particular implementation strategy for a screening program.

Data Requirements

The following was given in response to a request by the participants for information or data requirements of the IIASA study.

For estimating the natural history of cervical cancer, we need the following data on a large number of women:

a) birth date, plus any information on race, income, occupation (or husbands occupation), date of marriage(s), number of children;

b) dates and results of each Pap smear. Possible results are: normal, dysplasia, carcinoma in-situ, smear unsatisfactory;

c) if the woman has had carcinoma in-situ or invasive cancer, when and how (screening or clinic) was it discovered and treated. Also, the length of subsequent survival, and, if the woman has died, whether death was for a cause related to uterine cancer;

d) if the woman has been removed from the group at risk from cervical cancer either by hysterectomy or death, the date when this removal occured and whether it was for reasons related to cervical or uterine cancer; and

e) survival rates by age for the population being served.

The data will contain three different kinds of cases. First, there will be cases which, after a succession of normal or unsatisfactory smears, will be found clinically to have cervical cancer. We wish to be given all available data on all of these cases. Second, there will be cases that will have a positive smear, and subsequent histological confirmation, after a succession of normal or unsatisfactory smears. We also wish to have all available data on all of these cases.

Finally, there will be cases that consist of only a succession of normal and unsatisfactory smears. We will need only a sample, consisting of a few thousand women.

Success of the Workshop

It is not amiss to give our evaluation here of whether the Workshop succeeded. The reader will recall that we had three objectives in hosting the Workshop. These were:

a) to obtain advice on cervical screening from experts in the field;

b) to obtain access to data from existing screening programs; and

c) to persuade one or more of the Workshop participants to act as a client for this study.

The Workshop succeeded in its first objective. These proceedings reflect the advice we were given by the invited experts on cervical screening. Further, we expect these experts will review our progress from time to time and offer us corrective advice.

We may tentatively judge the Workshop to have accomplished its second objective. Several participants expressed a willingness to give us access to their data. However, acquiring the data in a useful form may require too much time. At the time of this writing, we are still exploring this problem.

It is not possible to say whether the third objective was achieved. At least one participant expressed an interest in the role of client, but no one made a firm commitment. Negotiations on this point are now in progress.

## References

[1] Dickinson, L. "Evaluation of the Effectiveness of Cytological Screening for Cervical Cancer. III. Cost-Benefit Analysis." Mayo Clinic Proceedings, 47 (1972) 550-550.

[2] Knox, E.G. "A Simulation System for Screening Procedures." In Problems and Progress in Medical Care. Nuffield Provincial Hospital Trust, London, Oxford University Press, 1973.

[3] Knox, E.G. "Computer Simulation of Alternative Population Screening Policies." In Systems Aspects of Health Planning. N.T.J. Bailey and M. Thompson. Eds. Amsterdam, North-Holland Publishing Co., 1975.

[4] Schneider, J. and Twiggs, L.B. "The Costs of Carcinoma of the Cervix." Obst. and Gynec., 40, 6 (Dec. 1972) 851-859