National Library
of Canada

Acquisitions and
Bibliographic Services Branch

395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale
du Canada

Direction des acquisitions et
des services bibliographiques

395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file Votre référence

Our file Notre référence

NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

Canada

IMPACT OF THE HEALTHSTYLES HEALTH PROMOTION PROGRAM ON HEALTH CARE UTILIZATION: A secondary analysis of data

by

Margaret Cousins

Thesis submitted to
the School of Graduate Studies and Research
in partial fulfilment of the requirements for the
M.Sc. degree in Epidemiology

University of Ottawa

ISBN 0-315-95902-9

Canada

UNIVERSITÉ D'OTTAWA
UNIVERSITY OF OTTAWA

ABSTRACT

Healthstyles was a community based health promotion program that ran in Ottawa from 1982 to 1984. This program provided an opportunity to assess the effects of health promotion on health behaviours and use of health care services through a quasi-experimental design. Primary analyses of questionnaire responses have shown a number of positive behavioural outcomes. For the secondary analysis, data on health care use 6 months before, 18 months during and 6 months after the program were obtained from Ontario's universal health insurance plan (OHIP). These data were available for 520 volunteer participants and 932 matched controls. Participants and controls were matched on sex, age (within 5 years), education level (2 categories), stress (3 levels), smoking status and an aggregate risk score.

Controlling for baseline differences through analysis of covariance, program participants had higher costs and number of visits for ambulatory care during the first year ($p<0.01$) and second year ($p<0.05$) after entry into the program. Within ambulatory care, participants used significantly more diagnostic services than controls during both years of follow-up. Participants were also more likely to use counselling and psychotherapy services in year 1 (RR 1.53, 95% CI 1.28, 1.81) and year 2 (RR 1.57, 95% CI 1.31, 1.89). There were no differences between participants and controls in visits for medical consultations and assessments, preventive services or surgical services.

These results suggest that health promotion may lead to increased costs and use of ambulatory services, at least over the short-term. However, these costs may be off-set over a longer follow-up period if health promotion successfully reduces the incidence of

chronic illness. Since the sample population was relatively young and had positive health behaviours at baseline, compared to the general population, the results may not apply to other, less healthy populations.

ACKNOWLEDGEMENTS

Firstly, I would like to thank my thesis supervisors, Dr. Ian McDowell and Dr. Rama Nair, for their advice throughout the past year. Jack Boulet and David Gilkes also deserve thanks for their help with SAS programming and other computing aspects of this thesis.

Finally, many thanks to my parents and to all my friends for their encouragement and support, especially during times of "thesis blues". Without them this thesis would never have been completed.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

BACKGROUND

Over the past few decades, health care spending by the Canadian government has increased faster than inflation and population growth. In 1988, Canada spent 8.7% of its gross domestic product on health care, compared to 7.7% in 1981 and 7.4% in 1971[12]. With the rapidly escalating costs and shrinking budgets, governments are seeking alternatives to the heavy focus on physician and institutional care, and cost containment has become an important issue.

The use of health care services is influenced by a number of societal and individual determinants. Societal norms, technology, environment, health care resources, organization of the health care system, and health status of the individual are among the factors included in models of health care use[3 4]. In economic terms, use of services is determined by supply and demand.

Efforts to reduce health care costs and use of services may target one or several of these determinants. A greater emphasis on health promotion and disease prevention is one option that is frequently hailed by politicians, employers and others as a cost containment measure[1 5 6 7 8]. J.Epp, a former Canadian Minister of Health, stated that "we believe that the health promotion approach has the potential over the long-term to slow the growth in health care costs"[9]. Russell, on the other hand, suggests that "successful health promotion results in health gains valued by those who benefit rather than cost-containment measures."[10]

Is the view of health promotion as a cost containment measure realistic or merely wishful thinking on the part of politicians and others?

OBJECTIVE

A community-based health promotion program in Ottawa called Healthstyles, which ran from 1982-1984, provided an opportunity to examine the effects of health promotion on health care use. The objective of this thesis was to assess the impact of this program on the use of ambulatory care services, based on health insurance records. The impact of the program on overall ambulatory service use as well as on selected types of ambulatory services (medical consultations, diagnostic services, preventive services, and counselling and psychotherapy) was of interest.

Before outlining the hypotheses some of the concepts of health promotion, the effects of previous health promotion programs on use of health care services, and the sources of data will be described.

HEALTH PROMOTION

1. Definition

In the Ottawa Charter, the World Health Organization (WHO) defined health promotion as "the process of enabling people to increase control over, and to improve, their health"[11], where health is considered to be "a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity". Health promotion is a broad concept and encompasses elements of health care, health education, health administration, public health, risk factor reduction, prevention and a number of other fields[12].

The five key strategies for health promotion identified by the WHO include building healthy public policy, creating supportive environments, strengthening community action, developing personal skills through provision of information and education, and reorientating health services in line with the goals of health promotion[11]. In addition to these five strategies, three main mechanisms for health promotion have been identified in Health and Welfare Canada's Achieving Health for All: self-care, mutual aid and the creation of healthy environments[9].

Self-care refers to the decisions and actions taken by an individual to enhance or maintain health[9] [13]. Programs that promote self-care are directed towards enabling individuals to make informed health-related decisions. These programs also try to improve knowledge and appropriate use of self-treatment, non-medical supports and medical services. Adoption of healthy behaviours (risk factor reduction), good

management of chronic disease, and optimal compliance with prescribed treatments are also among the goals of such programs[13].

Mutual aid implies that people can help each other cope with health issues by working together, sharing ideas, information, and experiences, and providing each other with support[9].

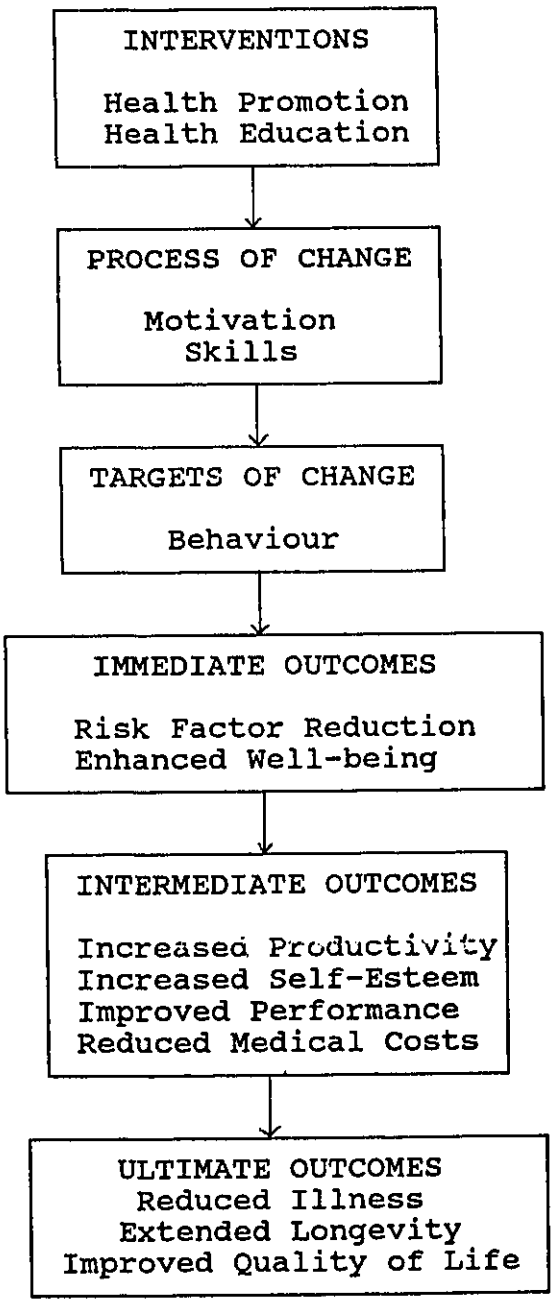
Creation of healthy environments involves modification of social, economic and/or physical environments in order to preserve and enhance health[9].

2. Potential Effects on Health Care Use

According to Sherman et al.'s modification of Green's evaluation model for health promotion (fig. 1), it is expected that health promotion and health education will result in behaviour change which in turn is expected to lead to improved health and reduced illness [14].

The associations between lifestyle and health, and between health promotion and behaviour change, have all been studied extensively. This research has established associations such as that between smoking and lung cancer, as well as the associations between seatbelt use and reduction of alcohol consumption before driving and reduced risk of serious injury or death[15] [16]. In addition, associations between a number of behaviours (smoking, physical activity, alcohol consumption, hours of sleep and eating breakfast) and self-assessed health status, happiness and total mortality rates have been reported [17] [18]. However, the relationships between type A behaviour, exercise, diet and

Fig. 1 Evaluation model for health promotion
(Source: JB Sherman et.al.[14], adapted from Green[21])



weight, and heart disease are less clear[15 16 19 20].

Research also indicates an association between psychological well-being and health. Life stressors, such as recent bereavement, depression, and family conflict have all been found to be associated with increased frequency of medical encounters[22 23 24]. Stress itself has been found to be accompanied by a number of physiological changes that may lead to symptoms such as tachycardia, muscle tension, fatigue, insomnia, headaches and backaches[23 25].

Several large-scale studies, such as the North Karelia project[26] and Johnson and Johnson's Live for Life program[27], have evaluated the impact of health promotion programs on health behaviours and risk factor levels. Overall, these studies indicate that health promotion programs have a positive impact on a number of health behaviours including smoking[5 26 27 28 29 30 31], physical activity[5 27 32], weight[29 31 32 33] and diet[28]. Improvements in risk factor levels, such as blood pressure and cholesterol are also evident [5 26 31 32 34]. The effects of these programs on stress and other measures of mental health are less consistent but suggest that health promotion programs can also have positive effects on type A behaviour, stress and well-being[29 32].

Given the evidence linking behaviours and health, and the ability of health promotion programs to modify a number of health behaviours and stress, it is not unreasonable to expect a resultant reduction in illness and symptoms. Since need (perceived and evaluated by health professionals) accounts for the largest proportion of the variance in health care use, decreases in illness and symptoms should then lead to reduced use of health care services, as outlined in Green's model[3 35 36 37 38].

Since many of the illnesses targeted by health promotion have long latency periods, declines in health care use may not be evident until several years later[8]. However, it is also hypothesized that healthier lifestyles, improved well-being and lower stress levels will lead to fewer symptoms and more immediate decreases in use of services[8]. These reductions are most likely to be evident in use of ambulatory services, although increases in use of safety precautions such as seatbelts may also reduce use of emergency hospital services.

In addition to improving health behaviours, mutual aid and increased social support through health promotion should also contribute to reductions in health care use. Social support appears to improve health and well-being by buffering the effects of stress on mental and physical health[16]. Social support and improvements in coping skills are also expected to lead to some decrease in use of services by frequent users; these individuals have been found to have less effective coping strategies and poor social environments[39 40].

Finally, the promotion of self-care and improved knowledge of appropriate use of health services should lead to a reduction in inappropriate visits for minor symptoms.

The potential for some increases in use of services must also be recognized. Belief in preventive health behaviour has been found to be a significant predictor of both illness-related utilization and general utilization[41 42]. Individuals with greater health knowledge may seek medical care for prevention or early intervention[8]. In addition, initial health screens may lead to overall increases in service use if follow-up care is required[7]. Although the use of these services will lead to some increased costs, such an

increase will hopefully lead to lower costs in the long-term as health problems are detected and treated earlier. On the other hand, some of the visits made by those who have a belief in health maintenance may be unnecessary and represent a "worried well" effect[7 43].

Efforts to deal with psychological distress and life problems are also expected to lead to initial increases in the use of counselling and psychotherapy services. Again, this increase should lead to a subsequent reduction as individuals improve their coping skills and deal with their problems. Reduced distress should also translate into fewer physical symptoms and reductions in use of illness-related services.

Overall, health promotion has the potential to lead to increases in the use of some services and reductions in the use of others. Consequently, health promotion may not lead to an overall decrease in costs. Even if there is a net decrease in health care utilization, economists have argued that health promotion and disease prevention programs are not necessarily cost-effective; program costs may outweigh savings from reductions in health care use[44]. The costs of health promotion programs are immediate and may continue over several years and benefits may not be evident for years [44 19]. The cost-effectiveness will therefore depend critically on the discount rate. In addition, if health status is improved and there are fewer premature deaths indirect costs, such as pensions, may be incurred[20 44]. At the population level health promotion's cost-containment ability will be limited by participation rates, the influence of unmodifiable determinants (e.g. age, genetics) on disease, the ability of individuals to respond to

programs and long-term maintenance of behaviour changes[19 20 44 45 46]. Other social and economic determinants of health care use and the nature of the medical system are also factors affecting the ability of health promotion to lead to a significant reduction in health care costs[20 47]. However, while health promotion may prove to be ineffective as a cost-cutting measure it is recognized that it may still have valuable effects on health and that cost-effectiveness should not be the sole basis for health care decisions[10 20 48 49].

3. Impact on Health Care Use

3.1 Work-site Based Programs

While several evaluations of work-site health promotion have included program effects on health care costs as a component of cost-benefit analyses, only a few have considered health care utilization and measures of health as outcomes.

In the evaluation of the Blue Cross and Blue Shield of Indiana's employee health promotion program participants were found to have higher payments and more claims than controls during the first 6 months of follow-up[7]. However, these group differences were not statistically significant. After 4.75 years of follow-up significant decreases were found in total, hospital inpatient, medical-surgical and major medical but not hospital outpatient costs. The average difference between costs for participants and controls was 24% and savings were estimated to be $143.60/employee. The study included 667 exposed and 892 unexposed individuals.

The Johnson and Johnson Live for Life (LFL) program also had positive effects

on health care costs and absenteeism[8]. Significantly lower inpatient hospital costs and lower rates of increase in hospital admission rates and hospital days were observed for employees at LFL program sites(n=2955) compared to employees at non-LFL sites(n=8461). In addition, a significant reduction in sick days was observed among wage-earning employees at the LFL sites. Outpatient costs and other costs were not significantly affected by the program over the 5 years of follow-up.

In a Canadian based study, Shepard et.al. evaluated the impact of an employee fitness and lifestyle program at the head office of the Canada Life Assurance Company[50]. In this study Ontario health insurance records from 1977 and 1978 for employees at Canada Life were compared to those for North American Life employees. Despite the large sample size of 1125, the difference in total costs between participants and control groups failed to reach statistical significance ($p=0.07$). Comparing one year baseline to one year follow-up, no significant effects were seen for costs in obstetrics and gynaecology, ECGs, orthopaedic or other services.

In another study, a work-site fitness program led to reductions in absences due to illness[51]. The effect was greatest among those having the highest number of absences at baseline and the highest participation rates.

Other studies have also reported reduced absenteeism[5 28], although differences were not significant in all cases[29]. Those studies evaluating number of symptoms and health status found no significant program effects[28 52].

Overall, the results are mixed, with two health promotion programs showing reductions in health care use and another showing no effects. Similarily, three studies

found reductions in absenteeism while another found no effect and two more failed to find decreases in number of symptoms.

3.2 Community Based Programs

Few of the community based health promotion programs have included health care utilization as an outcome variable. In a prospective, randomized study, the effects of a self-care educational intervention were evaluated by comparing use of health care services during the first year following the start of the program to health care use during the baseline year[53]. Comparing 363 control households to 1262 experimental households, reductions in total ambulatory and minor illness visits were significantly greater in the intervention group. Hospital use was not significantly affected by the program. Adjusting for baseline utilization levels, the average participant-control difference in total ambulatory visits was 14% while the average participant-control difference for minor illness visits was 28.3%.

In the evaluation of a self-care program for older adults, the self-reported number of medical visits did not differ significantly between participant and control groups[52]. The program also failed to have any effect on self-reported health status, life quality or social function.

In a study of retirees, the reductions in the number of physician visits and hospital days, comparing participants to controls, also failed to reach statistical significance[15]. The results of this study were based on medical insurance claims.

In terms of large scale studies there was a 10% lower disability payment rate for

cardio-vascular disease reported among the exposed in the North Karelia study by the fifth of the program[26]. These results were based on population data for experimental and control regions. This program also led to significant decreases in emotional problems and psychosomatic symptoms in the experimental region.

Overall, few studies have evaluated health care utilization as an outcome of health promotion. Of the studies that have been done, the majority were based on quasi-experimental designs. However, in most instances non-equivalent comparison groups with pre and post-program measures for both groups were included, thereby controlling a number of potential threats to validity. These studies also tended to have large sample sizes and follow-up periods of at least one year.

The assessment of use of health care services was based on either self-reports or health insurance or medical records. While self-reported use is subject to the problems of recall, there are also limitations to the use of records. In the United States insurance plans are not universal nor do they cover all services. Consequently, changes in the use of certain types of services may have been missed and variations in insurance coverage may have led to some bias. Accuracy of insurance and medical records is also a potential problem.

The results of the studies suggest that health promotion may have an effect on health care use as well as on absences due to illness. However, there is little evidence to support the hypothesis that health promotion is associated with fewer symptoms. In addition, there are some inconsistencies in the types of health care services affected by

health promotion. Evaluations of American work-site programs report reduced hospital and total utilization while the Canadian based study found no significant effects. On the other hand, one community based program led to reduced ambulatory and minor illness visits but no reduction in hospital visits. No significant effects were reported in the other community based studies. Unfortunately, the effects of the large community wide health promotion programs on health care utilization have not been evaluated.

The length of follow-up may account for the discrepancies in the effect of health promotion programs on use of health care services. The work-site studies reporting reductions in the use of hospital services had follow-up periods of 5 years while the other studies only had one year follow-up periods. During a short follow-up period initial increases in service use, such as those observed by Gibbs et.al.[7], may counteract the effect of later reductions.

Inconsistencies in the results also suggest that it may not be valid to generalize the findings based on American work-site studies to community based programs where the study populations may be different. Similarly the impact of Canadian programs on health care use may differ from American studies as the health care systems of these two countries differ greatly. There is a need to further evaluate the effects of health promotion on health care use in different populations and settings.

In addition, since most of the studies concentrated on the effects of health promotion on use of hospital services and very broad categories of services, evaluation of effects on ambulatory services and more specific subtypes of services would be beneficial. As program impact on illness may also vary by baseline health status[51]

separate evaluations for frequent and infrequent users of services should be considered. Long-term and short-term evaluations are necessary to further the understanding of the impact of health promotion on health care use.

HEALTHSTYLES

1. Program Description

Healthstyles was a community health promotion program run by the Ottawa Centretown Community Health Centre in the 1980s. This program provided an opportunity to evaluate the impact of health promotion on behaviours and to address some of the questions remaining about the effect of health promotion on use of health care services.

Although Healthstyles was developed before the publication of the WHO's Ottawa Charter[11] and J.Epp's Achieving Health For All[9], this program contained many of the elements of health promotion as described in these documents; healthy behaviour was promoted and personal skills developed through the provision of information, and self-care and mutual aid were emphasized.

The program began with the Healthstyles Basics weekend workshop. During this weekend participants were provided with information on nutrition, physical activity, stress management, smoking and alcohol consumption, and the relationship between lifestyle habits and health. Participants assessed their present lifestyle and were then taught how to set goals, solve problems and implement weekly goals[54 55].

An 18 month social support program followed the weekend workshop. This

program consisted of 5 group sessions, weekly telephone support, monthly calls from a trained volunteer Healthstyles Assistant and a number of optional support programs. Optional support programs included individual sessions for those with particular needs and interests and a weekly lecture series "Taking Charge of Your Health"[56].

2. Study Design for Program Evaluation

Healthstyles was evaluated with a quasi-experimental, non-equivalent control group design[57]. Volunteer participants were attracted to Healthstyles through word of mouth, local media coverage and some were referred by health and social services. Although Healthstyles was run by the Centretown Community Health Centre, enrollment in the program was not restricted to members of the centre. As they enrolled, applicants were placed into 40 groups of 15 to 20 that began the program at two-week intervals.

Potential controls were drawn on a voluntary basis from health clubs, health centre and physicians' offices, starting before the Healthstyles program began. As participants entered the program they were matched with two individuals from the pool of controls. Participants and controls were matched on sex, age (within five years), educational level (2 categories), self-reported stress level (3 categories), smoking status, and an aggregate risk score based on eating breakfast daily, exercise level, alcohol consumption, sleeping habits and body weight (5 categories)[57]. This extensive matching was employed to increase the internal validity of the study.

As a result of the matched design there was a delay between the enrollment of participants and their respective controls. In some rare cases appropriate controls were

not found until over a year after the participant began the program. In addition, in some instances only one matched control was found for a participant. Finally, if a participant dropped out of the study before attending the Healthstyles Basics workshop, the matched control may have been matched to a later participant. In these cases the matched control was enrolled before the participant. A total of 640 participants and 1180 controls were included in the study.

Data were collected at baseline and 3, 12 and 18 months later through self-administered questionnaires. Based on questionnaire responses, summary scores were derived for smoking habits, physical activity, stress, health locus of control, weight, distress, general well-being and overall health. Socio-demographic information and a description of coping behaviours were also obtained. The inclusion of questionnaires for controls as well as participants allowed the program effects to be evaluated while controlling for the potential effects of secular trends and the influence of completing the questionnaire.

3. Behavioural Outcomes

In the primary analysis questionnaire data were used to evaluate changes in health behaviours and risk factor levels from baseline to 18 months of follow-up. Compared to controls, the experimental group displayed significantly greater improvements in health locus of control ($p<0.0001$), self-reported smoking habits ($p<0.02$), exercise ($p<0.002$), body weight ($p<0.0001$), nutrition ($p<0.002$), stress levels ($p<0.001$) and overall health score ($p<0.001$)[57].

HEALTH CARE UTILIZATION DATA

1. Description

The data upon which this thesis is based were drawn from computerized records of health care use, recorded from the universal health care insurance program of Ontario.

In Ontario there are two arrangements through which physicians are paid. Those who are members of Health Service Organizations (HSO) and Community Health Centres (CHC) are usually paid on a capitation basis. The remainder receive fee-for service payments through the Ontario Health Insurance Plan (OHIP)[58]. For the purposes of this study the term HSO will apply to both HSO and CHC as the data from both sources are in the same format.

The data derived from HSO patient encounter forms and OHIP claims are stored on separate files although HSO members may have records in both data bases (encounters with physicians outside of the HSO are billed under OHIP). Together these two data files provide a record of almost all patient-physician encounters (ambulatory and hospital care) in the province made by Ontario residents. However, only a small number of encounters with non-physician health care workers (e.g. psychologists, nutritionists) are included in this data set.

Records for Healthstyles participants and controls were selected from the OHIP and HSO databases at the Ontario Ministry of Health using a matching routine based on the OHIP number, given name, sex, and birthdate. Records were not available for

individuals who lived in Quebec or members of the military. In addition, some people did not provide their OHIP number and some could not be matched to the OHIP and HSO files. The records were accompanied by an index file which provided the identification of all participants and controls found in the OHIP and HSO data files (including those who did not make any visits during the above time period). The final analysis data set represented the health care visits made by 520 (81.3%) of the 640 participants and 932 (79.0%) of the 1180 controls. Two hundred and eighty-nine of the participants had two matched controls, 193 had one matched control, and 38 had no matched controls. One hundred and sixty-one controls had no corresponding participant. All participants and controls for whom health care data were available were retained for the initial analyses; certain analyses used only one matched control for each participant.

Although groups of participants were enrolled at two week intervals, for simplicity OHIP provided records for all study members from July 1, 1982 to November 31, 1986. For each person the time period of OHIP and HSO monitoring covered six months before starting the program (baseline period), 18 months of intervention and an additional six months of follow-up. For the analyses the time periods of interest were the 6 month baseline period, and the first and second years following entry into Healthstyles. Subdividing the first and second years into time periods of six months was also considered. However, it was felt that the year long time periods would provide a more stable and accurate estimate of the use of health care services. In addition, services such as general assessments and annual health exams are restricted to one per year.

As intake into the program was staggered over time, the data used for the analysis

can be viewed as a parallelogram within the rectangle of data provided (fig. 2). July 1, 1982 corresponds to the beginning of the baseline period for group 1. The baseline period began on March 29, 1984 for those in the last group. January 4, 1983 and September 29, 1984 were the dates of the first and last Healthstyles Basics weekends while December 31, 1985 and September 29, 1986 correspond to the end of the two year follow-up period for group 1 and group 40, respectively. The time period between July 1, 1982 and the start of the baseline period has been termed the "pre-before" and the period between the end of the two year follow-up and November 31, 1986 has been termed the "post-after". The records from these two time periods were not used in the analysis.

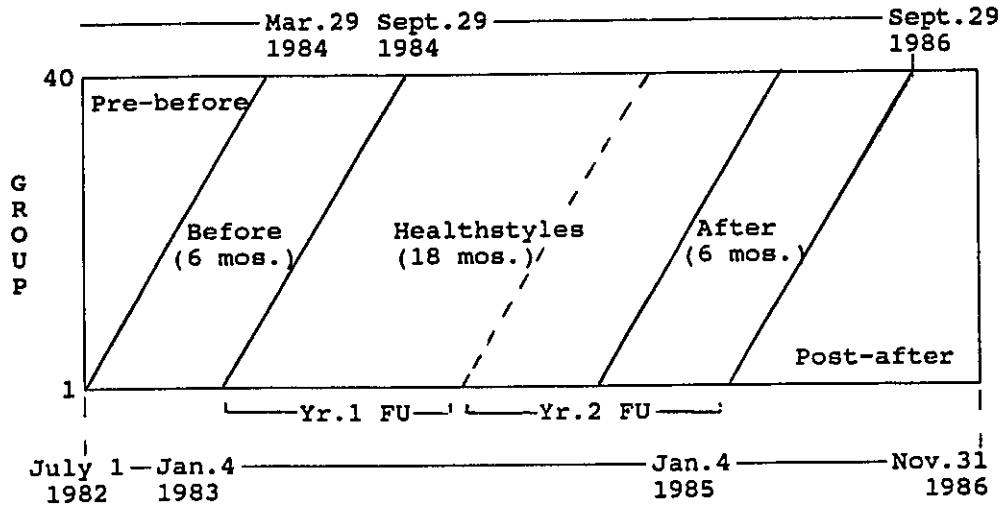


Fig. 2 Diagrammatic representation of the Healthstyles time-frame within the claim files.

July 1 1982: beginning of utilization data, 6 months before first Healthstyles Basics weekend
Jan. 4 1983: first Healthstyles Basics weekend
Mar. 29 1984: beginning of baseline period for group 41
Sept. 29 1984: last Healthstyles Basics weekend
Jan. 4 1985: end 2 year period after program entry for group 1
Sept. 29 1986: end 2 year period after program entry for group 40
Nov. 31 1986: end of utilization data

Note: "pre-before" and "post-after" were not used in the analyses

2. Variables

Each record from the HSO and OHIP data files contained index variables. These included the Healthstyles identification number of the individual, experimental status (participant or control), a unique identifier assigned by OHIP, the date at which they began the baseline period, gender and date of birth. These index variables were followed by information specific to the HSO or OHIP encounter.

The variables in the HSO file include the month and year of visit, and the services and procedures provided at that encounter. Each record in this file is based on information from patient encounter forms and represents one patient encounter. The data set had 6,318 HSO records.

On the other hand, the OHIP records are based on fee claims and there is a separate record for each individual service or procedure. The variables in this file include the day, month and year of the service, fee schedule code for the service provided, and the cost of the service. A total of 135,368 OHIP records were provided by the Ontario Ministry of Health.

A list of all variables is provided in table 1.

Table 1 Variables in HSO and OHIP records

a) INDEX DATA

- unique code number for each person(assigned by OHIP)
- experimental status (1=participant, 2=control)
- ID assigned by Healthstyles
- gender
- birth date (YY/MM/DD)
- first day of baseline period (YY/MM/DD)

b) HSO DATA - one record per patient visit

- INDEX DATA
- service date (YY/MM)
- HSO service codes (maximum of 4 codes/record)
- procedure code (maximum of 4 codes/record)

c) OHIP DATA - one record per service provided

- INDEX DATA
- fee schedule code
- service date (YY/MM/DD)
- fee paid ($)

ESTIMATION OF EFFECT SIZE

To derive testable hypotheses, it was necessary to specify what impact the Healthstyles program was expected to have on use of health care services. Since the effect sizes will vary by type of consultation, effect sizes for total ambulatory care and each type of consultation are presented separately.

Total ambulatory: Previous studies have reported reductions in overall use of services of 14%[53] and 24%[7]. The former value is based on one year follow-up for a community based study while the latter represents the effects of a work-site program after almost 5 years. Since Healthstyles is a community-based study with a follow-up period of two years a conservative estimated effect size of 15% was used for this thesis.

Minor illness visits and diagnostic services: In the study by Vickery et.al.[53] a 34% reduction in minor illness visits was reported after one year of follow-up. However, no reduction was observed in other studies[15 52]. Consequently, for the evaluation of Healthstyles the effect on medical consultations and assessments was estimated to be 25%. Since estimates for the impact of health promotion on use of diagnostic services were not available and diagnostic services are often generated from illness visits, the effect size for these services was also estimated to be 25%.

Psychotherapy and counselling: Studies of primary health care indicate that 29% to 31% of visits are psychosocial in nature[59 60 61]. Although the effect of health promotion on the use of these services has not been evaluated, it may be possible to reduce the use

of psychotherapy and counselling services by 10% (absolute change of 3%) with improved social support and coping strategies.

Preventive services: In terms of use of preventive services, it has been reported that about 20% of patients have annual check-ups[61]. It is expected that this proportion could increase to 30% or 40% of individuals (an increase of 50-100%).

Surgical services: Although surgical and injury related services have been included in the analysis no significant impact is expected on the use of these services during the two year follow-up period of this study.

As previously mentioned, while it is hoped that health promotion will lead to a decrease in use of ambulatory services, increased health awareness and sensitivity to symptoms may lead to significant increases in use of services during the first 6 to 12 months of the program. Therefore, the estimated effect sizes may not be evident until a year or more following the onset of the program.

HYPOTHESES

Based on the study objective (p.10) and the literature review, the following hypotheses were developed for the effects of Healthstyles on use of health care services

I. PRIMARY HYPOTHESES

During the second year of Healthstyles:

1. the use of total ambulatory services by participants will be at least 15% less than that of the controls, controlling for baseline differences.

2. the use of assessments and consultations by participants will be at least 25% less than that of the controls, controlling for baseline differences.

3. the use of diagnostic services by participants will be at least 25% less than that of the controls, controlling for baseline differences.

4. at least 50% more participants than controls will use preventive services, controlling for baseline differences.

5. at least 10% fewer participants than controls will use psychotherapy and counselling services, controlling for baseline differences.

6. there will be no significant difference between the groups for use of injury related and surgical services.

II. SECONDARY HYPOTHESES

1. During the first year of Healthstyles participants, compared to controls, will exhibit a significantly higher use of

   a) total ambulatory services
   b) assessments and consultations
   c) diagnostic services
   d) preventive services
   e) counselling and psychotherapy

2. There will be no significant difference between the groups for use of injury related and surgical services during the first year of Healthstyles.

3. There will be a significant baseline level of health care use x time interaction in assessing changes in health care utilization.

METHODS

DATA PREPARATION

1. HSO-OHIP Coding and Grouping of Services

As mentioned in the description of the claims data, HSO members may have records in both HSO and OHIP files if they have used services outside of the HSO. Consequently, it was necessary to combine the records from these two files to obtain valid measures of use of health care services by HSO members. Since the coding of services and procedures for data collection differs greatly between HSO and OHIP systems it was necessary to identify code equivalents.

Under the HSO system, 21 service and 18 procedure codes define broad categories of the services provided. In contrast, the fee schedule codes used under the OHIP system are numerous and very specific. In order to analyse the OHIP and HSO data together the OHIP codes were collapsed into the broader HSO categories using the 1986 OHIP Schedule of Benefits[62] to select appropriate OHIP code equivalents.

Based on the definitions of the various types of assessments and consultations provided in the 1986 OHIP Schedule of Benefits (table 2), and consultations with physicians, the HSO services and procedures and OHIP equivalents were then collapsed into 5 groups by type of service: medical consultations, diagnostic services, preventive services, counselling and psychotherapy and surgical/injury related services (see table 3). These groups were formed to allow the impact of Healthstyles on different types of ambulatory services to be assessed while avoiding problems of multiple testing and

Table 2 OHIP and HSO terms and definitions[62]

1. Consultation - refers to the situation where a physician, in light of his/her professional knowledge of the patient, requests the opinion of another physician competent to give advice in the field, because of complexity, obscurity, or seriousness of the case, or because another opinion is requested by the patient.

2. Repeat consultation - meets the criteria of consultation and implies interval care by the primary physician.

3. Minor assessment - visit with direct doctor patient contact at which either or both of the following are provided and appropriate record made of the visit
a) brief assessment which includes the necessary history and examination of the affected part or region or mental or emotional disorder
b) brief advice or information regarding health maintenance, diagnosis, treatment and/or prognosis

4. Intermediate assessment - more extensive assessment than the minor assessment

5. General assessment - full history and an enquiry into, and an examination of all parts or systems, an appropriate record and advice to the patient (normally limited to one per year per patient)

6. General re-assessment - same as above except that the history need not include all the detail already obtained in the original general assessment

7. Annual health exam - all the elements of a general assessment as it pertains to an individual after the second birthday who presents and reveals no apparent physical or mental illness - limited to one per patient per year

8. Psychotherapy - any form of treatment for mental illness, behavioural maladaptations, and/or other problems that are assumed to be of an emotional nature, in which the physician deliberately establishes a professional relationship with a patient for the purposes of removing, modifying, or retarding existing symptoms, or attenuating or reversing disturbed patterns of behaviour, and of promoting positive personality growth and development.

9. Counselling - activity in which physician engages in an educational dialogue with the patient(s), where the goal is to become aware of the patients' problems or situation and of modalities for prevention and/or treatment.

10.Supportive care - care rendered in the hospital by the referring physician, who is not actively treating the patient, at the desire of patient or family, for purposes of liaison or reassurance.

insufficient numbers of visits which would have arisen from analysis of each service separately.

Ideally, service and procedure codes representing health maintenance practices would have been separated from illness care services. Unfortunately, with the exception of the annual health exam and PAP smears, it was very difficult to differentiate preventive or health maintenance visits from those which are illness related. For example, while minor and intermediate assessments may be made in response to illness symptoms, monitoring chronic illness and blood pressure checks are also reported under these codes. Similarly, lab tests such as urinalysis, anoscopy and ECGs may be done as part of a routine annual check-up or in the effort to diagnose the cause of symptoms. Finally, while general assessments may be done in response to illness, a large proportion of annual health exams are billed as general assessments. Thus, although efforts were made to separate those services most likely to be preventive in nature from those related to illness, the service groupings were not homogeneous.

The reasons for the exclusions were as follows:

1) allergy injections, provider supervision, supportive care and obstetrics were not expected to be affected by health promotion.

2) well child, well baby, injections and immunizations are related primarily services for children rather than the Healthstyles participants.

3) case conferences, certificates, reports and "other" are not benefits under OHIP and thus have no OHIP code equivalents.

Table 3 Types of services: HSO and OHIP codes

	HSO SERVICE	HSO PROCEDURE	OHIP FEE SCHEDULE
MEDICAL CONSULTATIONS			
1. minor assessment	01	--	A001
2. intermediate assessment	02	--	A007
3. consultation	05	--	A005 C005
4. repeat consultation	06	--	A006 C006
5. monitoring chronic illness	12	--	A001 or A007
DIAGNOSTIC SERVICES			
1. lab work	--	01	L001-252 L255-418 L422-799 L900-999 G001-008 G011-012
2. x-ray	--	02	X001-299 X400-999
3. venipuncture	--	04	G480 G482 G489
4. urinalysis	--	05	L253 L254 G489
5. haemoglobin	--	06	L419-421 G481
6. ECG	--	11	G181 G284 G308 G310 G311 G315 G316 G651 G652 G654 G657 G658
7. anoscopy	--	15	Z543
PREVENTIVE SERVICES			
1. general assessment	03	--	A003
2. annual health exam	07	--	K009
3. PAP smear	--	03	G365
COUNSELLING AND PSYCHOTHERAPY			
1. individual counselling	10	--	K013
2. birth control/ family planning	14	--	K013
3. individual psychotherapy	11	--	K007
4. group counselling	16	--	K013

* Group counselling, and birth control/family planning are both billed as individual counselling in OHIP so all three of these HSO codes were included in this category

	HSO SERVICE	HSO PROCEDURE	OHIP FEE SCHEDULE
TOTAL AMBULATORY			
1. MEDICAL CONSULTATIONS			
2. DIAGNOSTIC SERVICES			
3. PREVENTIVE SERVICES			
4. PSYCHOLOGICAL SERVICES			
5. supportive care	15	--	C010
6. general reassessment	04	--	A004
SURGICAL			
1. suturing	--	12	Z154 Z175 Z176 Z179
2. plaster application	--	16	Z198 Z199 Z201-Z203 Z207 Z208 Z211 Z213
3. minor surgery	--	13	Z101-Z108 Z114 Z115 Z118-Z120 Z156-Z164 Z166-Z174 Z153 Z227 Z245
4. surgical procedures	--	--	all codes under surgical procedures in OHIP fee schedule (excluding those listed above)
EXCLUSIONS			
1. well baby	08	--	
2. well child	09	--	
3. obstetrics	13	--	
4. case conference	17	--	
5. provider supervision	18	--	
6. forms completion	19	--	
7. referral letter	20	--	
8. other	21	--	
9. allergy injection	--	07	
10. IM/SC injection	--	08	
11. immunization	--	09	
12. dressings	--	10	
13. IUD insertion/removal	--	14	
14. deliveries	--	17	
15. other	--	18	

2. Assignment of Costs to HSO Records

As HSO records do not include costs of services provided it was necessary to provide a cost estimate for these services based on the costs for the equivalent OHIP codes. For those HSO service and procedure codes that correspond to one OHIP claim code, the mean fee paid by OHIP for that service or procedure was calculated for each year (1982-1986) and applied to the HSO records. For procedures such as labwork and x-rays, in which many OHIP codes correspond to the one HSO code, the fees paid for the OHIP codes in these categories were summed for each encounter and the mean fee paid per encounter applied to the HSO records (see table 4 for costs applied to the HSO services and procedures).

Table 4 Mean service fees($) by year based on OHIP claims

Service	Year				
	1982	1983	1984	1985	1986
minor assessment	10.69	11.57	12.58	13.12	13.68
intermediate assessment	14.11	15.69	17.13	18.71	20.23
general assessment	31.00	33.66	36.58	38.72	40.96
general reassessment	18.20	19.66	21.29	22.71	24.00
consultation	34.40	36.67	39.84	41.52	44.21
repeat consultation	25.45	28.30	30.50	32.30	34.30
individual counselling	33.07	36.27	38.56	39.95	43.70
individual psychotherapy	51.76	59.72	69.32	69.39	74.49
annual health exam	26.80	29.64	32.08	34.11	36.34
PAP smear	3.10	3.20	3.38	3.41	3.61
labwork	21.85	22.84	24.40	27.17	28.89
venipuncture	1.39	1.47	1.84	1.75	1.83
haemoglobin	1.80	1.73	1.84	1.90	1.90
anoscopy	3.80	3.99	4.42	4.73	5.08
urinalysis	2.80	2.41	2.51	2.75	2.63
xrays	31.14	33.40	36.78	37.94	39.33
ECG	14.54	12.42	13.90	17.94	21.54
plaster application	15.50	14.63	18.07	20.03	15.96
sutures	8.90	9.87	16.03	13.46	19.58
minor surgery	12.96	11.91	18.95	25.44	16.10

DISTRIBUTION OF DATA

Since the probability distribution of the data plays an important role in the determination of the statistical methods that can be used to test the hypotheses, frequency distributions for dependent variables during the second year of the program were generated for participants and controls.

The frequency plots of number of visits for total ambulatory care (fig. 3), medical consultations and diagnostic services all approximated Poisson distributions (see appendix 1 for all distributions of number of visits for controls, medical consultations and diagnostic services). From table 5 it is evident that the maximum number and cost of visits for the latter two types of consultations was much lower than for total ambulatory care.

While 13% to 15% of individuals did not use any ambulatory services, over 65% of individuals did not use counselling or psychotherapy during the second year of follow-up. Among those that did use these services, the frequency was highest for 1 visit then declined as the number of visits increased (fig. 4). Compared to medical consultations and diagnostic services the maximum costs and number of visits for psychological services were very high.

As expected, preventive services had the narrowest ranges for number and cost of visits. Fifty-six percent of participants and 61% of controls did not use these services while approximately one third of individuals made one visit (fig. 5).

The distributions for use of surgical services were similar to those for counselling and psychotherapy in that a large percentage of participants (77%) and controls (79%) did not use these services and the maximum costs were high. However, the maximum number of visits was only 18 (fig. 6).

Table 5 Data distributions by type of consultation (year 2 of follow-up)

Type of Consultation			Mean	Median	S.D.	Maximum	% no visits
Total	# visits	P	6.76	4	12.18	192	13.08
		C	4.95	3	6.90	136	15.34
	cost($)	P	277.75	120.08	925.52	14869.84	
		C	149.01	83.88	391.94	9947.27	
Medical Consultations	# visits	P	2.71	2	3.56	31	26.73
		C	2.53	1	3.18	30	28.54
	cost($)	P	47.03	27.56	62.69	599.26	
		C	43.91	20.40	55.75	468.16	
Diagnostic Services	# visits	P	3.89	3	4.52	35	26.15
		C	2.76	1	3.95	30	35.52
	cost($)	P	80.51	39.59	114.92	821.98	
		C	50.43	18.69	77.58	614.29	
Psychotherapy and counselling	# visits	P	1.79	0	10.74	190	69.42
		C	0.68	0	4.76	128	82.08
	cost($)	P	132.55	0	893.59	14430.84	
		C	40.40	0	358.87	9580.86	
Preventive services	# visits	P	0.54	0	0.69	3	55.77
		C	0.44	0	0.60	3	61.27
	cost($)	P	17.39	0	24.58	118.07	
		C	14.20	0	21.17	82.32	
Surgical Services	# visits	P	0.48	0	1.36	18	76.92
		C	0.43	0	1.08	10	78.54
	cost($)	P	32.38	0	116.44	1060.32	
		C	38.77	0	181.85	3539.81	

P=participants C=controls

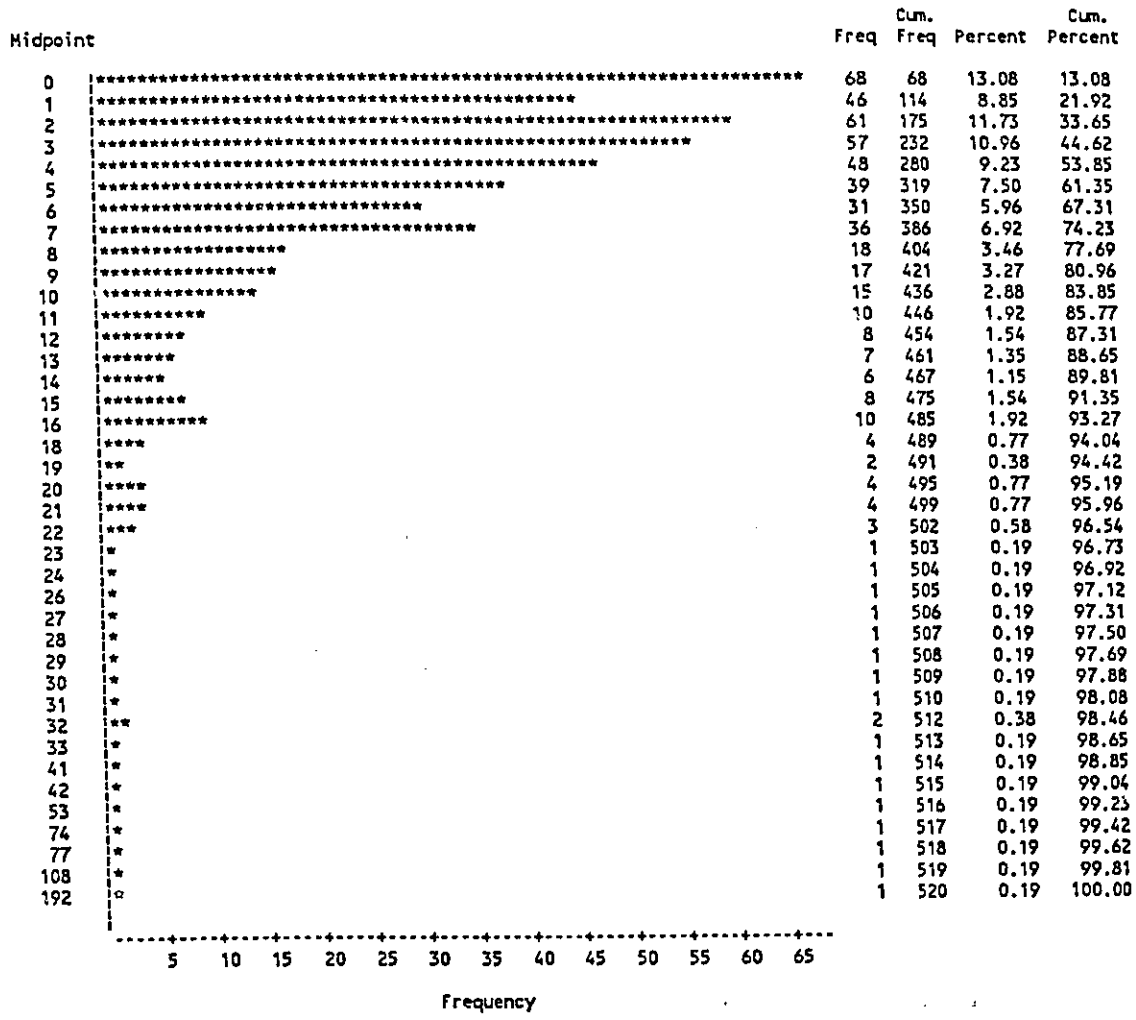


Fig. 3 Visits for ambulatory services in Year 2: participants

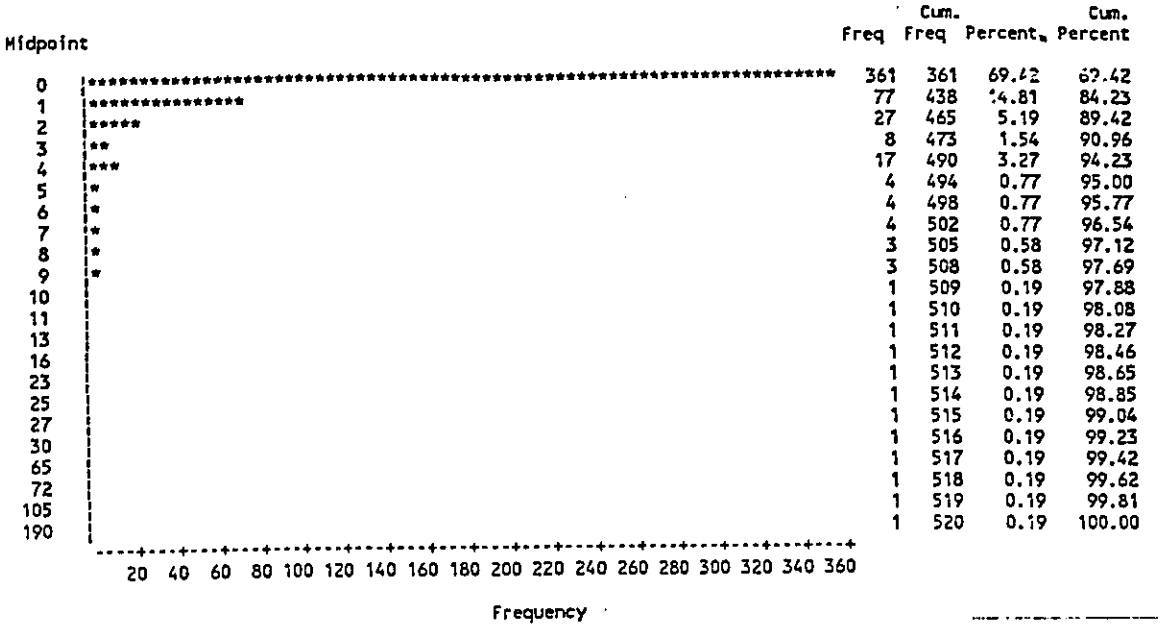


Fig. 4 Visits for counselling and psychotherapy in year 2: participants

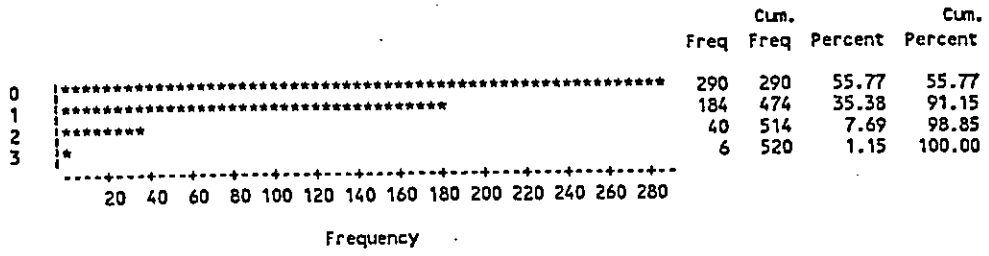


Fig. 5 Visits for preventive services in year 2: participants

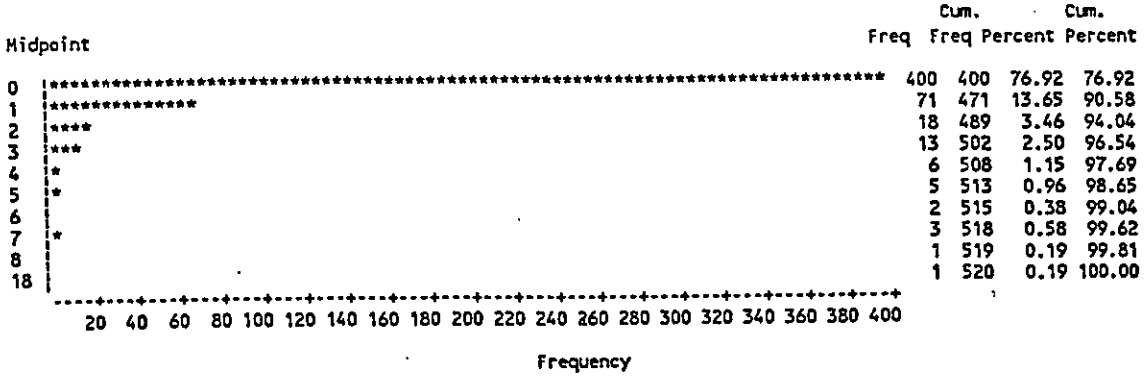


Fig. 6 Visits for surgical services in year 2: participants

DEPENDENT VARIABLES

Several indicators were considered for the evaluation of the impact of Healthstyles on use of health care services.

1. Proportion of individuals making at least one visit: If the goal of health promotion is to eliminate all visits for specific types of consultations then the proportion of individuals making visits would be the primary outcome. However, in most instances the goal is to reduce, rather than eliminate, use of services.

The proportion of individuals using services is also the most appropriate outcome for evaluating program effects on services that are dichotomous in nature (e.g. annual health exam).

2. Mean number of visits: The mean number of visits is one of the main outcomes for assessing the ability of health promotion to lead to a reduction in use of services. However, it must be acknowledged that although this measure is related to the nature and number of illnesses suffered, it is also influenced by differences in physician practices and individual differences in perception of illness.

3. Cost of services: Cost of services consumed allows the economic effects of a program to be evaluated and is also expected to reflect the amount of resources consumed in terms of personnel time and technology. As with number of visits, cost of services consumed

may be affected by the number and nature of the illnesses experienced, as well as physician response to the symptoms presented.

4. Illness Episode: One illness episode includes all the encounters related to the same illness. This measure indicates the number of illnesses experienced by individual that received medical attention and provides a measure of their health status. Since health promotion is hypothesized to improve well-being one would expect a reduction in the number of illnesses experienced by participants in health promotion programs.

Unfortunately, it is very difficult, on the basis of utilization records alone, to determine the number of illnesses experienced by each individual.

5. Deviation from the optimal number of visits: For some types of consultations and illnesses there may be an optimal number of visits (e.g. one health exam/year). Deviation from this optimal number of visits, either below or above, would represent an undesirable outcome. For most consultations and conditions the optimal number of visits would be dependent on the health status of the individual and potentially a number of other characteristics such as age, gender and family history. Even the optimum number of annual health exams and other health maintenance and preventive services may vary across individuals. Consequently, although conceptually useful, this measure is difficult to apply in a practical situation.

6. Proportion making excessive visits: Since one of the goals of health promotion

programs is to reduce inappropriate use of services, the proportion of individuals making excessive or inappropriate visits could be used to assess program effectiveness. Unfortunately, the determination of appropriate visits is a difficult task that requires knowledge of the patient's symptoms, the number of visits appropriate for a given problem etc. The available data did not include this information.

Due to insufficient information, illness episodes, deviation from the optimal number of visits and the proportion making excessive visits were not used as outcomes for the Healthstyles study. Based on the data distributions it was decided to use the mean number and cost of visits as outcomes for assessment of the impact of Healthstyles on total ambulatory services, medical consultations and diagnostic services. The proportion of individuals making visits was used as the outcome measure for preventive services, psychotherapy and counselling, and surgical services.

POTENTIAL CONFOUNDERS

Models of health care use suggest that many factors influence the decision to consult medical professionals and to use health care services[3 4 63]. As previously mentioned, "need" accounts for the largest proportion of variance in health care use [3 35 36 37 38]. Worry about health, income and health insurance are also significant predictors [35 64]. Gender, education, race, and living alone have been reported to influence use directly, and indirectly through need variables[35 64]. Use of services is also expected to

be influenced by characteristics of the health care system (e.g. access), time of year (seasonality) and inflation or secular trends.

Participation in health promotion programs is also expected to be affected by several factors. Many of these factors, such as age, gender, education, health status and health behaviours, are among the predictors of health care use and are therefore potential confounders when evaluating the effects of health promotion on health care use.

As mentioned in the description of the study design, as a result of the matching process, controls were enrolled after participants. Since use of health care services is subject to both seasonal and time trends it was also desirable to assess and control for any biases resulting from group differences in month and year of entry into the program.

Finally, group differences in baseline use were also assessed as the impact of health promotion was expected to differ by baseline levels of use. In addition, since there is a floor effect in the measurement of health care use (it is not possible to have less than zero visits) the effects of health promotion may be difficult to detect among infrequent users. Therefore group differences in the proportion of infrequent users may lead to bias. Similarly, if baseline use was higher in one of the groups, changes over time may be the result of regression to the mean rather than program effects [65].

The following three sections describe how group differences in personal characteristics, seasonal and time factors, and baseline use of services were evaluated. The results of these group comparisons and control of potential confounders are also discussed since they influenced the methods selected for testing the hypotheses.

1. Personal Characteristics

Participant and control groups were compared with respect to a number of baseline demographic and health related measures. These comparisons were first made based on all participants and controls, ignoring the matching (unmatched analysis). The analysis was then repeated for matched pairs using the corresponding matched analyses. The matched and unmatched comparisons were done to determine the effects of the matched design on group comparisons. Before presenting the baseline comparisons the scales used to measure internal health locus of control and happiness will be described.

1.1 Description of Measurement Scales

a) Wallston's Multi-dimensional Health Locus of Control Scale[66]

Wallston's multi-dimensional health locus of control scale reflects three dimensions of health locus of control beliefs: internal, external and powerful others. The internal dimension was measured in Healthstyles and indicates the extent to which an individual perceives health outcomes to be dependent on his/her own actions. The internal score is based on responses to six 6 point Likert response items. Scores of one to six are assigned to each item and summed for a maximum internal score of 36 (indicating high internal health locus of control).

b) Goldberg's General Health Questionnaire[67]

The general health questionnaire (GHQ) is a screening instrument to detect current diagnosable psychiatric disorders including depression, anxiety, social impairment and

hypochondriasis. This scale is designed for use in general population surveys in primary medical settings and among general outpatients. The shorter 12 item version was used in Healthstyles to measure mental distress. The GHQ score indicates the probability of being a psychiatric case. A low score is desirable.

c) Bradburn's Affect Balance Scale[68]

The Bradburn Affect Balance scale is used to measure the psychological reactions of members of the general population to happenings in their everyday lives. This scale consists of ten 3 point response items. The items measure positive and negative affect and a summary score of well-being is derived by subtracting the negative affect score from the positive affect score. A high score is desirable. In Healthstyles this scale was used to place people on a scale of happiness.

1.2. Unmatched Comparisons (based on all participants and controls)

Based on the distributions of the variables, group differences were tested as follows:

a) Chi-square test (nominal variables)

marital status

b) Mann Whitney signed rank tests (ordinal variables)

income, willingness to change, self-rated health, effort to improve health, Bradburn summary score, and physical activity

c) T-test (continuous variables, and ordinal variables with several levels)

alcohol consumption, health locus of control and general health questionnaire scores

For each variable observations with missing variables were omitted from the analysis. The results of the comparisons are summarized in tables 6 & 7.

No significant differences were found between the participants and unmatched controls in terms of education or income. The two groups did differ in terms of marital status with the control group having a higher percentage of married individuals.

In terms of health and health behaviours, level of physical activity did not differ significantly by group. However, a larger proportion of participants were "not too happy", and participants tended to be more distressed (General Health Questionnaire) and less likely to rate their health as excellent than controls. On the other hand, controls had a higher mean alcohol consumption and were more likely to actively try to improve their health but they were less willing to change. While statistically significant group differences were also found for internal health locus of control (IHLC), the differences were small.

Table 6 Baseline characteristics: by study group

Characteristic	Unmatched analysis					Matched	
	P	n	C	n	p	p	# pairs
married	46.0%	520	63.3%	932	<0.01	<0.01	481
income >$25000	65.3%	481	64.9%	427	0.79	0.72	196
physical activity 1=sedentary	59.9%	494	59.2%	865	0.10	0.17	481
3=adequate	13.8%		17.8%				
median	0.86		0.84				
self-rated health 1=excellent	12.2%	518	20.1%	925	<0.01	<0.01	477
4=poor	5.8%		1.2%				
median	1.7		1.4				
actively improve health 1=most of the time	35.4%	520	47.1%	930	<0.01	<0.01	481
4=never	0.4%		0.5%				
median	1.3		1.1				
willing to change 1=not very	0%	520	0.90%	924	<0.01	<0.01	479
7=very	51.4%		14.8%				
median	6.0		4.2				
Bradburn score 1=very happy	10.0%	518	19.7%	927	<0.01	<0.01	478
2=somewhat happy	60.6%		67.3%				
3=not too happy	29.4%		13.0%				
median	1.7		1.5				

P=particpant C=control

Table 7 Baseline characteristics: by study group

Characteristic	Unmatched analysis					Matched	
	P mean	n	C mean	n	p	p	# pairs
internal health locus of control	29.4	520	27.9	926	<0.01	<0.01	478
alcohol consumption (drinks/14 d)	5.3	520	6.2	932	0.02	0.13	481
General Health Questionnaire	18.5	519	16.4	927	<0.01	<0.01	477

P=participant C=control

1.3 Matched Comparisons

For each of the participants with at least one control, matched analyses were performed. In cases where a participant had two controls the first control was used in the analysis.

Based on the distributions of the variables, group differences were tested as follows:

a) McNemar Chi-square test (nominal variables)

marital status

b) Wilcoxon rank test on difference scores (ordinal continuous variables)

income, willingness to change, self-rated health, effort to improve health, Bradburn summary score and physical activity

c) Matched t-test (continuous variables, and ordinal variables with several levels)

alcohol consumption, health locus of control and general health questionnaire scores

For each variable, matched pairs with missing observations were excluded from the analysis.

From Tables 6 & 7 it can be seen that the results from the matched analyses were very similar to the unmatched analyses.

Compared to the unmatched analyses, the matched analysis gave higher p values for group differences in education, drinking habits, and physical activity. Lower p values were obtained for income. The discrepancies in significance levels between the matched and unmatched analyses may be partially explained by the fact that some

individuals who were included in the unmatched analyses were excluded from the matched analyses; the matched analysis included 481 matched pairs while the unmatched analysis was based on all 520 participants and 932 controls. The discrepancies may also be due to the different samples sizes.

In summary, the participant and control groups were comparable in terms of all the demographic variables, with the exception of marital status. On the other hand, many of the health and health behaviour variables differed significantly by study group. Participants tended to be less happy (Bradburn scale), more distressed and have poorer self-rated health than controls. Although fewer participants than controls actively tried to improve their health, more were willing to change.

A number of these group differences, especially those that reflect psycho-social problems and distress, may be related to use of health care services and are thus potential confounders in this study[22 23 24 69]. It is expected that group differences in predictors of health care use will be reflected in group differences in use of services during the baseline period. Consequently, differences in baseline health care use were controlled in the analyses, rather than controlling for each of the potential confounders separately. Baseline use of services will reflect not only differences in mental well-being but also any other group differences in predictors of health care use.

2. Seasonal and Time Factors

2.1 Lag time between participant and control enrollment

To assess the potential biases resulting from the lag between participant and control enrollment, the frequency of lag times between the entrance of participants and their matched controls was plotted (fig. 7). From this frequency plot it is evident that approximately 75% of controls were enrolled within 3 months of the participant and 95% within 6 months. A small percentage were enrolled before the participants as after their first matched participant dropped out (before attending the Healthstyles Basics weekend workshop) they were rematched. There were also some controls enrolled as late as 14 or 19 months after their matched participants.

Chi-squared analyses were used to determine whether there were significant differences between the study groups with respect to calendar month and year of entry into the program. In both instances the chi-square values reached statistical significance ($p < 0.001$). Consequently it was necessary to adjust for the time-lag of entry of

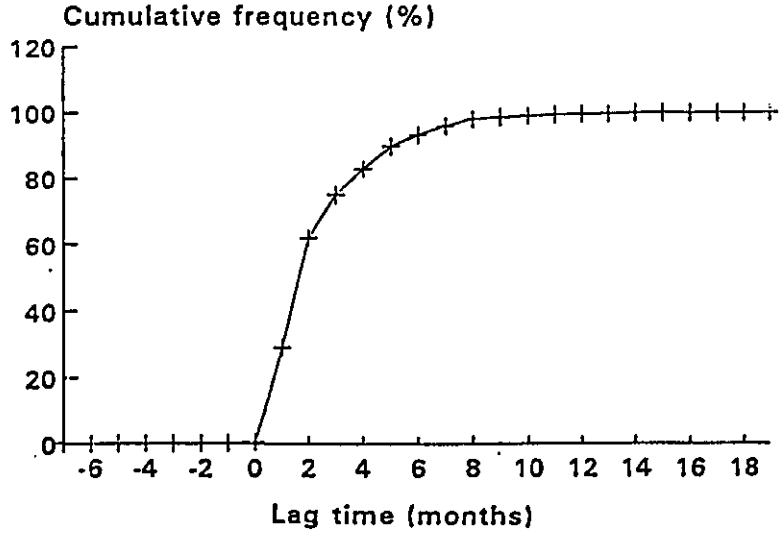


Fig. 7 Frequency of Lag Times Between Participant and Control Entry into Healthstyles

controls, as this would be associated with inflation in health care costs over time. It was also necessary to assess whether significant variations in utilization occurred by calendar month and/or year.

2.2 Adjustment for Inflation

Two indices were considered in selecting a rate to adjust for inflation in health care costs over the period of the program: 1. percentage increases in mean fees for each HSO code, based on OHIP equivalents (data driven) and 2. Statistics Canada average annual consumer price indices (CPI) for 1982 to 1985, using 1986 as the base year.

From table 8 it can be seen that the increase in fees varied widely across the service groupings and by year within the groups (minimum=-13%, maximum=29%, mean=7.63%/year). In contrast, the inflation rates based on the CPIs (table 9) are lower and have a narrow range (minimum=3.9%, maximum=5.7%, mean=4.55%/yr.).

While the percentage increases in fees should provide a better estimate of inflation in health care costs than the CPI, the estimate is dependent on which services were used. Although the CPI rates appear more stable they are based on inflation in a number of consumer goods and services. Therefore changes in CPI do not represent inflation in health care costs alone and may be either lower, or higher, than the true inflation in these costs.

As neither of the indices considered for estimating inflation was perfect, the average of the mean inflation rate in CPI values and the mean increase in health care costs was used to adjust the costs of services (6%/year). One inflation rate, rather than

Table 8 Percentage increases from previous year in fees: by type of service

Service	1982-83	1983-84	1984-85	1985-86
minor assessment	8.23	8.73	4.29	4.27
intermediate assessment	11.20	9.18	9.22	8.12
general assessment	8.58	8.67	5.85	5.79
general reassessment	8.02	8.29	6.67	5.68
consultation	6.60	8.64	4.22	5.68
repeat consultation	11.20	7.77	5.90	6.19
individual counselling	9.68	6.34	3.60	9.39
individual psychotherapy	15.38	16.08	0.10	7.35
annual health exam	10.60	8.23	6.33	6.54
PAP smear	3.23	5.63	0.89	5.87
labwork	3.54	6.81	11.36	6.32
venipuncture	5.83	24.96	-4.86	4.46
haemoglobin	-3.89	6.36	1.03	0.00
anoscopy	5.00	10.70	7.11	7.19
urinalysis	-13.82	4.33	9.23	-4.04
xrays	7.24	10.13	3.15	3.65
ECG	-14.56	11.90	29.07	20.00
plaster application	-5.81	23.51	10.85	-20.32
sutures	10.90	62.41	-16.03	45.47
minor surgery	-8.10	59.11	34.25	-36.71
Mean	3.95	15.39	6.61	4.55

Table 9 Consumer Price Indices (CPI)

Year	CPI	% increase
1982	83.7	
1983	88.5	5.7
1984	92.4	4.4
1985	96.0	3.9
1986	100.0	4.2
Mean		4.55

separate inflation rates for each service, was selected largely for simplicity. While use of separate rates would have been relatively simple for the adjustment of the costs applied to HSO services, it would have been time consuming to adjust each of the large number of OHIP service codes separately.

2.3 Secular trends and residual inflation effects in health care use

The effects of time on health care use were tested using combined data from 1982-1986 for all participants and controls. Since the results from 1982 included only July-December, and those from 1986 did not include December, the individuals' utilization measures from 1982 were multiplied by 2 and those from 1986 multiplied by 12/11 to make the results for these years comparable to those from 1983-1985. It is recognized that due to the potential seasonal variation in health care use, the adjustment factor for 1982 values may have resulted in estimates that were either higher or lower than the true values.

Repeated measures ANOVAs were used to assess differences in cost (adjusted for inflation) and number of services by year. Scheffe's tests were used to determine which years were significantly different from each other.

Table 10 shows that significant year effects were found for use of total ambulatory services, medical consultations, diagnostic services and preventive services. In all of these groupings health care use in 1982 was significantly lower than use of services in 1984, 1985 and 1986 (see table 11). Significant differences in use were also seen

Table 10 ANOVA results: effect of year on health care use

Type of consultation	Variable	F	Significance of F
Total ambulatory	cost ($)	2.92	0.0199
	# visits	5.15	0.0004
Medical consultations	cost ($)	14.04	0.0001
	# visits	2.51	0.0434
Diagnostic services	cost ($)	2.37	0.0541
	# visits	9.14	0.0001
Counselling and psychotherapy	cost ($)	1.76	0.1346
	# visits	0.73	0.5696
Preventive services	cost ($)	5.87	0.0001
	# visits	3.78	0.0045
Surgical services	cost ($)	0.82	0.5103
	# visits	1.76	0.1339

Table 11 Mean number and cost of services by year

Type of consultation		1982	1983	1984	1985	1986
Total	# visits	4.75	5.25	5.58	5.48	5.60
	cost	82.31	99.09	116.50	113.77	126.15
Medical consultations						
	# visits	2.31	2.54	2.55	2.52	2.53
	cost	35.19	41.12	42.95	44.08	45.63
Diagnostic services						
	# visits	2.55	2.87	3.14	3.19	3.18
	cost	0.39	0.24	0.25	0.09	0.17
Counseling and psychotherapy						
	# visits	0.87	0.95	1.04	1.03	1.15
	cost	46.70	57.72	73.24	69.35	79.64
Preventive services						
	# visits	0.40	0.49	0.44	0.47	0.47
	cost	11.94	14.54	14.10	15.20	15.39
Surgical services						
	# visits	0.37	0.42	0.46	0.45	0.46
	cost	25.06	31.82	34.76	35.40	28.98

between 1983 and 1986 for medical consultations, and diagnostic services.

While the adjustment of 1982 data may have been somewhat inadequate and partially contributed to some of the significant time effects, the significant differences between utilization in 1983 and 1986 suggest that the time effects were not adequately controlled through the adjustments for inflation. In addition, significant year effects were also evident in number of visits, suggesting secular trends in health care use.

Before deciding on an appropriate method to control for these effects the seasonality of health care use was evaluated.

2.4 Seasonality of health care use

The seasonal effects on number and cost of visits in each grouping of services, and overall, were tested for significance using the combined data from 1983-1985 for all participants and controls. Repeated measures ANOVA, testing for the effect of month, was used to assess seasonal differences in number and costs of visits while chi-square analyses were used to assess differences in the number of individuals using services.

The ANOVA and chi-square tests showed that use of all types of services, except surgical procedures, had significant variation by month, both in terms of number and cost of visits and the proportion of individuals making visits ($p < 0.05$)(tables 12 & 13). From figs. 8 to 10 it is evident that use of ambulatory services tended to be lower in the summer months as well as during December. A similar pattern was seen for the different types of consultations (not shown).

Table 12 Effect of month on use of health care: by type of consultation

Type of consultation	Variable	F	Significance of F
Total ambulatory	Cost	5.68	0.0001
	# visits	8.63	0.0001
Medical consultations	Cost	4.45	0.0001
	# visits	5.94	0.0001
Diagnostic services	Cost	1.74	0.0623
	# visits	3.01	0.0005
Preventive services	Cost	3.00	0.0005
	# visits	3.01	0.0005
Counselling	Cost	3.05	0.0004
	# visits	3.72	0.0001
Surgical	Cost	0.91	0.5267
	# visits	0.89	0.5495

Table 13 Monthly variation in proportion of individuals using services

Type of consultation	X^2	p
Total ambulatory	44.97	<0.0001
Medical consultations	44.00	<0.0001
Diagnostic services	29.43	0.002
Preventive services	33.96	<0.0001
Counselling and psychotherapy	13.82	0.243
Surgical services	15.11	0.178

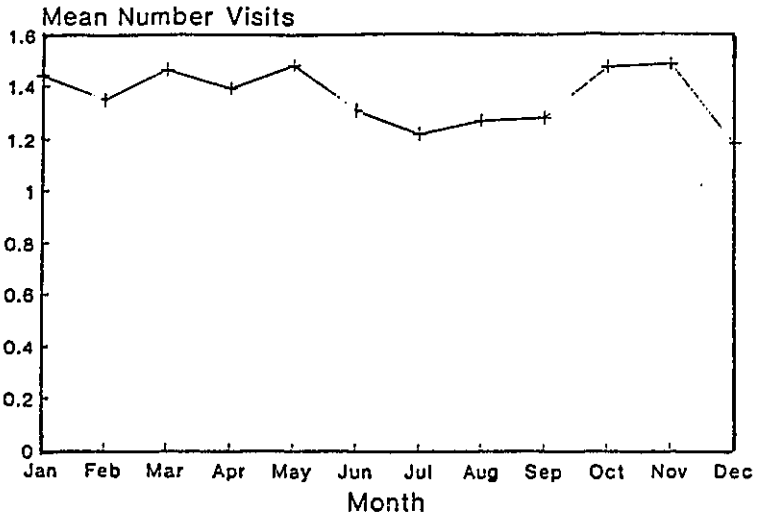


Fig. 8 Mean number of ambulatory visits by month among those making visits

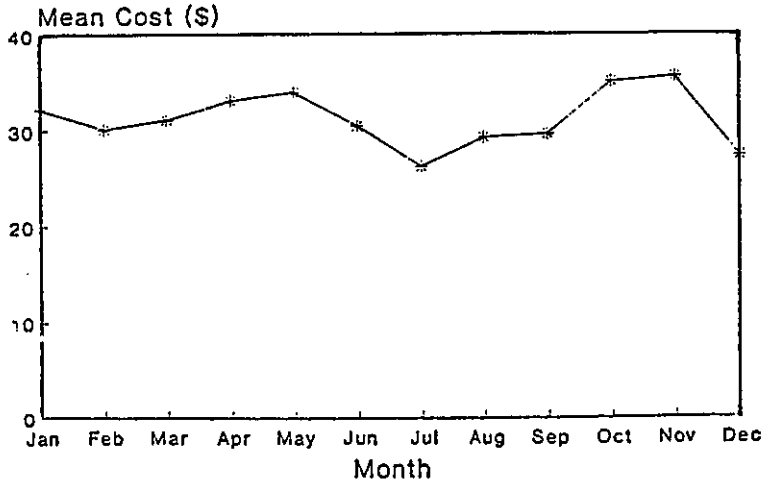


Fig. 9 Mean cost of ambulatory visits by month among those making visits

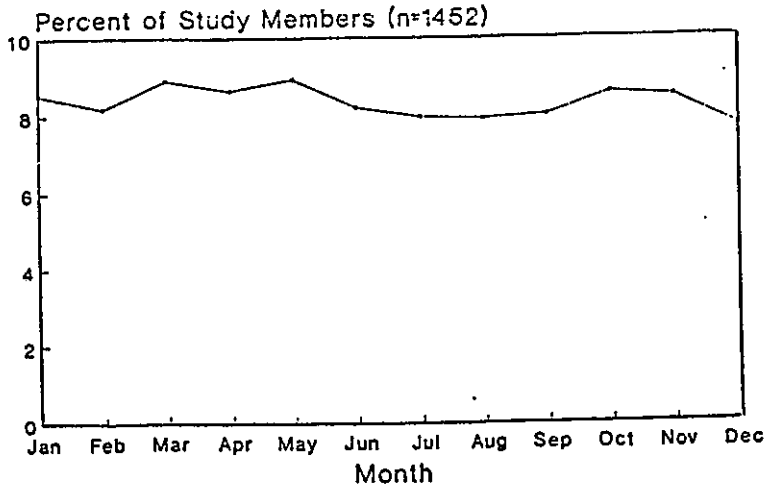


Fig. 10 Proportion of individuals making visits by month

2.5 Control of seasonality effects

As the participant and control groups differed significantly by month and year of entry into the study, and as use of health care services showed significant seasonal and time effects, it was necessary to control for these effects. Two methods were considered to control for time and seasonality effects:

a) **Assignment of the participant join date to its matched control(s).** This method provides exact control for the time and seasonality effects. However, among the controls this method results in the beginning of the follow-up period in health care data to precede the real time of entry into the program and completion of the baseline questionnaire. As a consequence of mismatched questionnaire and outcome time periods any effects on health care use related to questionnaire completion alone will no longer be completely controlled.

b) **Creation of a variable to indicate the number of high demand months during the baseline six months.** This variable, as well as the variable for year of entry into the program could then be included in regression analyses or used as basis for stratification to control for the effects of seasonality and time. While effectively controlling for any questionnaire effects this method only provides an approximate control for seasonality and time effects.

The first option was chosen since effective control of seasonality and time effects

was considered to be of greater importance than potential questionnaire effects. The completion of the Healthstyles questionnaires was expected to have a minimal, if any, effect on health care utilization while seasonality and time have statistically significant effects. In addition, since 95% of the controls were enrolled into the study within six months of their respective participants it was felt that any existing questionnaire effects would be sufficiently controlled. Assignment of participant join dates to controls is also a less complicated method to control for seasonal and time effects.

3. Baseline Use of Services

To assess potential bias arising from group differences in baseline use of services, t-tests were used to test for group differences in the cost and number of visits for total ambulatory, medical consultations and diagnostic services. Chi-square tests were used to compare group differences in the proportions using preventive, counselling and psychotherapy, and surgical services.

From table 14 it is evident that participants used significantly more ambulatory services than controls during the baseline period. In the participant group the mean cost of these services was almost twice as high as in the control group. Use of diagnostic services during this time period was also higher among the participants but no significant differences in use of medical consultations were found.

The proportion of participants using counselling and psychotherapy services was significantly higher than the proportion of controls using these services. No group

differences were found for the proportion of individuals using preventive or surgical services (table 15).

The methods used to control potential bias arising from baseline differences in use of health care services are outlined in the description of the tests used for testing the hypotheses (p.67).

Table 14 Consultations at baseline: by study group

Type of consultation	Variable	Mean		Prob > z
		Participants	Controls	
Total ambulatory	# visits	3.47	2.69	0.0032
	cost($)	136.10	72.96	0.0010
Medical consultations	#visits	1.43	1.42	0.928
	cost($)	23.36	22.71	0.705
Diagnostic services	#visits	1.97	1.38	0.0001
	cost($)	38.97	24.80	0.0001

Table 15 Consultations at baseline: by study group

Type of consultation	Participants (n=520)	Controls (n=932)	X^2	p
Preventive	24.8	22.1	1.38	0.24
Counselling	24.0	16.2	13.3	0.01
Surgical	13.3	11.7	0.77	0.38

MATCHING

As the evaluation of Healthstyles was based on a quasi-experimental design, pair matching was used to ensure that participant and control groups would have similar demographic characteristics and health behaviours, and to control for confounding by these factors. While matched designs are usually coupled with matched analyses of the results, the matching was ignored in the main analyses of this study. The unmatched analyses enabled differences in baseline levels of health care use to be controlled.

Normally unmatched analysis of a matched cohort study design will provide an unbiased estimate of the program effect since the matching is done before the exposure occurs. Consequently, the matching factors are not related to the exposure and the results are not confounded by these factors (in a case-control study crude analysis of a matched design is confounded by the matching factors). However, in this study an increase in bias may result from unmatched analyses since the ratio of participants to controls is not constant.

Compared to pair-matched analysis, the unmatched analysis also reduces the random error thereby increasing statistical efficiency[70 71 72 73].

HYPOTHESIS TESTING

1. Statistical Tests

1.1 Total ambulatory

Analysis of covariance, with baseline utilization as the covariate, was used to assess the overall impact of Healthstyles on number and cost of ambulatory visits. This method controls for the group differences in baseline use of services and provides an overall indication of the program effect.

1.2 Medical Consultations and Diagnostic Services

In order to assess the secondary hypothesis that the effect of health promotion would differ according to baseline level of service use (see Hypotheses p.33), individuals were first classified according to levels of baseline use of medical consultations and diagnostic services. The baseline groupings for cost of visits comprised 50%, 25% and 25% of the sample population (table 16). Similar groups were formed for number of visits at baseline although the discrete nature of this variable prevented the formation of groups that followed the 50%, 25%, 25% divisions exactly (table 17).

Table 16 Baseline levels of health care use: cost

Type of consultation	Baseline cost($)	n participants	n controls
Medical consultations	less than $13.67	246	468
	$13.67 to $33.23	135	238
	at least $33.24	139	226
Diagnostic services	0	240	516
	$0.01 to $38.52	111	222
	at least $38.53	169	194

Table 17 Baseline levels of health care use: number of Visits

Type of consultation	Baseline # visits	n participants	n controls
Medical consultations	0	214	406
	1,2	194	341
	3+	112	185
Diagnostic services	0	240	516
	1,2	114	228
	3+	166	188

Analysis of variance was then used to test first for baseline level x group interaction, then group effects. If the interaction term was significant, group differences were then tested within each stratum using t-tests. An alpha level of 0.05 was used to determine the statistical significance of all tests.

Although analysis of covariance has greater statistical power, the power to detect group differences remained high even after stratification of the sample due to the large sample size (see Power, p.96).

Although the data distributions were all highly skewed, no data transformations were performed and the analyses were based on the raw data. It was felt that the results based on the raw data would be easier to interpret. Furthermore, even the square root transformations on the data only led to marginal improvements towards normality in the distributions (see Appendix II for frequency distributions of transformed data and skewness and kurtosis indices). The potential bias arising from the nonnormality of the data and heterogeneity of variance should have been minimal since this study had a large sample size and the sizes of the study groups were relatively equal.

1.3 Preventive Services, Counselling and psychotherapy, and Surgical Services

2 x 2 tables (study group x visit status) were generated for the first and second years following the start of Healthstyles. In those instances where significant differences in baseline utilization were found, Mantel-Haensel X^2 and adjusted relative risks, controlling for baseline utilization, were calculated. Otherwise only the crude relative risks for each time period were calculated. The relative risks compared the risk of making a visit for participants compared to controls i.e. $RR > 1$ indicates that the participants have a higher risk of visiting.

RESULTS

TOTAL AMBULATORY CARE

Hypotheses:

1. During the second year of Healthstyles cost and number of ambulatory services used by participants will be at least 15% lower than those used by controls, controlling for baseline differences.

2. During the first year of Healthstyles participants will have a higher number and cost of visits compared to the control group, controlling for baseline differences.

The crude comparisons show that in the first and second years of follow-up the cost of visits made by participants were almost double the cost of visits made by controls (table 18). The mean number of visits was also significantly higher among the participants.

Table 18 Mean cost and number of ambulatory visits by year of follow-up

Year of follow-up	Variable	Mean		Prob. > Z
		Participants	Controls	
1	cost ($)	241.56	125.21	0.0001
	# visits	6.46	4.62	0.0001
2	cost ($)	277.75	149.01	0.0026
	# visits	6.76	4.95	0.0019

The analysis of covariance results indicate that even after controlling for differences in baseline levels of health care use, the participants used significantly more services than controls during both years (tables 19 & 20). After adjusting for baseline service use, the mean number of visits for participants in the first year of follow-up was 23% higher than for the controls. The mean cost of these visits were 31% higher among participants, compared to controls. During the second year of follow-up the mean number of visits was 21% higher among participants. The mean cost of visits during this time period was 38% higher for participants, compared to controls.

Conclusion: The first hypothesis was disproved and the second was supported.

Table 19 Ambulatory services: analysis of covariance results

Year of follow-up	Variable	Source	D.F.	Type III SS* MS	F	Prob > F
1	Cost ($)	group	1	711455	9.74	0.0018
		baseline cost	1	138642455	1897.41	0.0001
		error	1449	73069		
	# visits	group	1	409.99	15.81	0.0001
		baseline # visits	1	22972.38	885.77	0.0001
		error	1449	25.94		
2	Cost ($)	group	1	1418729	4.33	0.0375
		baseline cost	1	113181644	345.70	0.0001
		error	1449	327402		
	# visits	group	1	409.59	5.92	0.0150
		baseline # visits	1	21157.07	306.04	0.0001
		error	1449	69.13		

* Type III SS: the sum of squares obtained when the variable is added last to the model.

Table 20 Mean cost and number of ambulatory visits: adjusted for baseline utilization

Year of follow-up	Variable	Mean		Prob. > F (type III SS)
		Participants	Controls	
1	cost ($)	196.69	150.25	0.0018
	# visits	5.99	4.88	0.0001
2	cost ($)	237.21	171.63	0.0375
	# visits	6.31	5.20	0.0150

MEDICAL CONSULTATIONS

Hypotheses:

1. During the second year of Healthstyles, cost and number of medical consultations used by participants will be at least 25% lower than those used by controls, controlling for baseline differences.

2. During the first year of Healthstyles, participants will have a higher number and cost of visits compared to the controls, controlling for differences at baseline.

Although the point estimates of mean cost and number of visits for acute medical assessments and consultations were consistently higher in the participant group, no significant group differences were found in the crude comparisons (table 21). The ANOVA results summarized in table 22 show that there were no significant baseline level x study group interactions and that even after adjusting for the other variables in the model the group effects were still not statistically significant.

In general, these results are comparable to those observed when baseline cost and number of visits were retained in their continuous form and used as covariates in ANCOVA (table 23). However, the ANCOVA results indicate that the mean number of visits made by participants during the first year of follow-up was significantly higher than among the controls. The adjusted means show that in all cases the point estimates of number and costs of visits were higher among the participants (table 24).

Conclusion: Both hypotheses were disproved.

Table 21 Mean cost and number of medical consultations by year of follow-up

Year of follow-up	Variable	Mean		Prob. > Z
		Participants	Controls	
1	cost ($)	45.13	40.22	0.09
	# visits	2.72	2.39	0.05
2	cost ($)	47.04	43.91	0.34
	# visits	2.71	2.53	0.34

Table 22 Medical consultations: analysis of variance results

Year of follow-up	Variable	Source	D.F.	Type III SS		
				MS	F	Prob > F
1	Cost ($)	group	1	2165	0.92	0.3386
		baseline level	2	260554	110.28	0.0001
		group x baseline	2	1849	0.78	0.4575
		error	1448	2362		
	# visits	group	1	10.82	1.30	0.2553
		baseline level	2	965.75	115.57	0.0001
		group x baseline	2	11.54	1.38	0.2517
		error	1448	8.35		
2	Cost ($)	group	1	1698	0.56	0.4560
		baseline level	2	235135	76.96	0.0001
		group x baseline	2	2997	0.98	0.3753
		error	1449	3055		
	# visits	group	1	1.61	0.16	0.6864
		baseline level	2	790.36	80.09	0.0001
		group x baseline	2	20.36	2.06	0.1274
		error	1448	9.87		

Table 23 Medical consultations: analysis of covariance results

Year of follow-up	Variable	Source	D.F.	Type III SS* MS	F	Prob > F
1	Cost ($)	group	1	6468.83	2.99	0.0841
		baseline cost	1	866839.67	400.38	0.0001
		error	1449	2165.04		
	# visits	group	1	34.94	4.46	0.0350
		baseline # visits	1	2946.85	375.76	0.0001
		error	1449	7.84		
2	Cost ($)	group	1	2356.77	0.82	0.3661
		baseline cost	1	756311.16	262.35	0.0001
		error	1449	2882.82		
	# visits	group	1	10.09	1.08	0.2990
		baseline # visits	1	2473.42	264.37	0.0001
		error	1449	9.36		

* Type III SS: the sum of squares obtained when the variable is added last to the model.

Table 24 Mean cost and number of medical consultations: adjusted for baseline utilization

Year of follow-up	Variable	Mean Participants	Controls	Prob. > F (type III SS)
1	cost ($)	44.80	40.40	0.0841
	# visits	2.72	2.39	0.0350
2	cost ($)	46.73	44.08	0.3361
	# visits	2.70	2.54	0.2990

DIAGNOSTIC SERVICES

Hypotheses:

1. During the second year of follow-up the number and cost of visits made by participants for diagnostic services will be at least 25% less than controls, controlling for baseline differences.

2. During the first year of the program participants will have a significantly higher number and cost of visits compared to controls, controlling for baseline differences.

During the first and second years of follow-up, crude comparisons of the mean cost and number of visits by study group show that participants consistently used more diagnostic services than controls (table 25). Once again, the ANOVA results for all of the dependent variables indicate that the baseline level X study group interactions were not statistically significant (table 26). After adjusting for the other variables in the model, participant use of diagnostic services remained significantly greater than use by controls.

These results are comparable to those observed when baseline cost and number of visits were used as covariates in ANCOVA (table 27). The adjusted means derived from the ANCOVA indicate that the cost of services used by participants was over 40% higher than controls costs (table 28). The mean number of visits made by participants was 29% higher than in the control group.

Conclusion: The first hypothesis was disproved and the second was upheld.

Table 25 Diagnostic services: mean cost and number of visits by year of follow-up

Year of follow-up	Variable	Mean		Prob. > Z
		Participants	Controls	
1	cost ($)	72.97	45.94	0.0001
	# visits	3.75	2.68	0.0001
2	cost ($)	80.51	50.43	0.0001
	# visits	3.89	2.77	0.0001

Table 26 Diagnostic services: analysis of variance results

Year of follow-up	Variable	Source	D.F.	Type III SS		
				MS	F	Prob > F
1	Cost ($)	group	1	158537	26.85	0.0001
		baseline level	2	188186	31.87	0.0001
		group x baseline	2	1726	0.29	0.7466
		error	1448	5904		
	# visits	group	1	164.05	12.24	0.0005
		baseline level	2	565.11	42.17	0.0001
		group x baseline	2	18.51	1.38	0.2516
		error	1448	13.40		
2	Cost ($)	group	1	206295	25.01	0.0001
		baseline level	2	259905	31.51	0.0001
		group x baseline	2	2624	0.32	0.7276
		error	1448	8250		
	# visits	group	1	259.31	15.68	0.0001
		baseline level	2	603.07	36.48	0.0001
		group x baseline	2	4.39	0.27	0.7670
		error	1448	16.53		

Table 27 Diagnostic services: analysis of covariance results

Year of follow-up	Variable	Source	D.F.	Type III SS* MS	F	Prob > F
1	Cost ($)	group	1	141360	25.43	0.0001
		baseline cost	1	873893	157.18	0.0001
		error	1449	5559.7		
	# visits	group	1	212.88	16.29	0.0001
		baseline # visits	1	1664.05	127.31	0.0001
		error	1449	13.07		
2	Cost ($)	group	1	195766	24.15	0.0001
		baseline cost	1	711142	87.72	0.0001
		error	1449	8106		
	# visits	group	1	249.63	15.34	0.0001
		baseline # visits	1	1551.54	95.32	0.0001
		error	1449	69.13		

* Type III SS: the sum of squares obtained when the variable is added last to the model.

Table 28 Mean cost and number of visits for diagnostic services: adjusted for baseline utilization

Year of follow-up	Variable	Mean Participants	Controls	Prob. > F (type III SS)
1	cost ($)	68.93	48.19	0.0001
	# visits	3.59	2.78	0.0001
2	cost ($)	76.87	52.47	0.0001
	# visits	3.73	2.86	0.0001

PREVENTIVE SERVICES

Hypothesis: Fifty percent more participants than controls will use preventive services during the first and second years of Healthstyles, controlling for baseline differences.

During the first and second years of follow-up the proportion of controls using preventive services was marginally lower than the proportion of participants using these services (table 29). However, the association between study group and using preventive services was very weak and nonsignificant during both years.

Conclusion: There is no evidence in support of the hypothesis.

Table 29 Use of preventive services: by study group

Year of Follow-up	Proportion visiting		Crude RR	95% C.I.
	Participants (n=520)	Controls (n=932)		
1	39.6	38.4	1.03	0.90, 1.18
2	44.2	38.7	1.14	1.00, 1.28

COUNSELLING AND PSYCHOTHERAPY

Hypotheses:

1. During the second year of Healthstyles at least 10% fewer participants than controls will use psychological services, controlling for baseline differences.

2. Significantly more participants than controls will use these services during Year 1, controlling for baseline differences.

From table 30 it can be seen that the risk of use of these types of services is over 1.5 times higher for the participant group during the first and second years of the program, after controlling for differences at baseline. The associations between study group and use of counselling services were significant at the 95% confidence level.

Conclusion: The first hypothesis was disproved and the second hypothesis was upheld.

Table 30 Use of counselling and psychotherapy: by study group

Year of Follow-up	Proportion visiting		Crude RR	Mantel Haensel RR	95% C.I.
	Participants (n=520)	Controls (n=932)			
1	31.7	18.7	1.70	1.53	1.28, 1.81
2	30.6	17.9	1.71	1.57	1.31, 1.89

SURGICAL

Hypothesis: No significant difference between the proportion of participants and controls using injury related and surgical services is expected, controlling for baseline differences.

No significant associations were observed between study group and use of surgical services in either the first or second year of follow-up (table 31).

Conclusion: The hypothesis was upheld.

Table 31 Use of surgical services: by study group

Year of Follow-up	Proportion visiting		Crude RR	95% C.I.
	Participants (n=520)	Controls (n=932)		
1	25.2	22.5	1.12	0.92, 1.33
2	23.1	21.5	1.08	0.87, 1.30

DISCUSSION

DISCUSSION OF RESULTS

In comparison to other community-based studies Healthstyles had a favourable impact on a number of health behaviours (see Behavioural Outcomes, p.24), a longer follow-up period and a large sample size. In addition, the health care data used for this study were from a universal health insurance plan rather than from less universal and less comprehensive insurance plans, or from self-reports. From this data it was possible to assess the impact of the program on types of ambulatory services not included in other studies. Finally, the use of controls, pre and post program measures in both groups and stratification by baseline use controlled the major threats to the validity of quasi-experimental designs.

During the second year of follow-up the predicted reductions in use of services by participants were not achieved for any of the types of consultations. Unlike the results reported by Vickery et.al.,{130} Healthstyles participants had higher, rather than lower, levels of ambulatory services use, compared to controls. Similarly, although there were no significant group differences in use of medical consultations, the failure to achieve reductions in the use of these services contrasts with the results of the study by Vickery et al.{130}. The significantly higher proportion of participants using psychotherapy and counselling services during the second year of follow-up also contradicts the hypothesized effect of Healthstyles; during both years of follow-up

participants were at least 50% more likely than controls to use these services. Unfortunately, while participants showed higher use than controls for most of the types of consultations, the expected increase in use of preventive services was not observed.

On the other hand, the lack of association between study group and use of surgical services was expected for a follow-up period of two years and these findings are in agreement with those of other community-based studies[15 53]. However, significant decreases in hospital use by participants have been reported in work-site studies[7 8].

The results do show that, as hypothesized (see Secondary Hypotheses, p.33), use of total ambulatory services, diagnostic services, and psychotherapy and counselling services was higher among participants during the first year of Healthstyles, controlling for level of use at baseline. There were no group differences in use of medical consultations during this time period. Again, the hypothesized increases in use of preventive services were not observed and, as expected, use of surgical services was unaffected by the program during the first year of follow-up.

Finally, the hypothesized interactions between baseline level of use and study group were not observed, indicating that the program effect did not differ by level of service use at baseline. For medical consultations there were no significant program effects in any of the strata of baseline utilization while participants had higher use of diagnostic services, compared to controls, for all baseline levels.

A number of factors may have contributed to the failure of Healthstyles to lead to a reduction in health care use and to the discrepancies between the results of this study and those from past studies.

1. The study population

The volunteer sample used in this study was largely female, younger, more highly educated, and had a higher family income compared to the Ottawa regional norms (table 32). In addition, this population seems to represent the worried well with lower proportions of sedentary individuals and smokers but more unhappy individuals. In such a population there is less scope for improvements in health behaviours, general well-being and health status than in a less healthy population with fewer positive health behaviours. Consequently, changes in health care use as a result of health behaviour changes are also less likely and an intervention such as Healthstyles is unlikely to be effective in reducing health care utilization through changes in health behaviours.

In terms of health care use, Healthstyles participants had a higher mean number of visits at baseline compared to a random population sample in Quebec City (3.47 visits/6 months and 3.03 visits/6 months respectively)[74]. Healthstyles controls made slightly fewer visits on average (mean=2.69 visits/6 months). The higher mean value for participants combined with their positive health behaviours once again suggests that they may be "worried well".

The differences between the study population and regional and national populations also restrict the ability to apply the results to the general population. However, similar differences between participants and nonparticipants of health promotion programs have been observed in other studies[75] suggesting that the results of Healthstyles do represent the types of individuals who participate in health promotion.

Table 32 Comparison of baseline characteristics of study population to national and census data

Characteristic	Participants (n=520)	Controls (n=932)	National Data[a]	Census Data[b]
female	76.3%	76.5%	-	51.4%
married/separated*	56.0%	69.8%	-	68.8%
education university or more	51.9%	46.1%	-	25.1%
income > $25000	65.3%	64.9%	-	49.7%
current smoker	19.0%	16.9%	40.0%	-
> 14 drinks/week	8.8%	11.1%	9.4%	-
heavy/excessive weight	49.9%	41.0%	50.9%	-
sedentary	59.9%	59.2%	79.6%	-
not too happy (Bradburn summary score)	29.3%	12.9%	10.4%	-

a Source: Canada Health Survey, adjusted to match age/sex composition of the Healthstyles participant sample as reported by I McDowell et al.[57]

b Source: 1981 Regional Census as reported in the Healthstyles report[56]

* in the Regional Census data, married and separated people were grouped together

2. Program content

Differences in program content between Healthstyles and other programs may also have contributed to the disappointing results. In particular, self-care books were provided in the studies by Vickery et.al. and Leigh et.al.. These books provided algorithms for medical problem solving and when to consult a medical professional, and ". . . emphasized the application of 'home treatment' rather than physician visits for minor complaints" [15 53]. Thus, the positive effects on health care use observed in these studies may have been influenced more by advice offered in the self-care book and reductions in "inappropriate" or "unnecessary" visits than improvements in general well-being.

Although Healthstyles featured seminars, such as "Be Your Own Doctor", which dealt with self-care, attendance at these seminars was optional and appropriate use of

health care services was not part of the main program. The use of preventive services was also not specifically addressed in the program. Emphasis on self treatment of minor illness, appropriate use of services and introduction of algorithms for medical problems may be required for reductions in health care utilization to be achieved, particularly for relatively healthy populations. Unfortunately, on the basis of the health care utilization data alone it was not possible to determine the appropriateness of visits made or the appropriateness of non-visits.

3. Groupings of consultations and types of services assessed

As noted in the methods section (p.37), it was not possible to group services which were strictly preventive or illness related. Consequently, the apparent lack of program effect on use of medical consultations may reflect a combination of reduced illness-related visits and increased numbers of visits for health maintenance procedures, such as blood pressure checks. Similarly, it was not possible to determine whether the higher use of diagnostic and lab services among the participants represented negative program effects or increases in services related to health maintenance efforts. Unfortunately, the lack of increases in use of preventive services suggest that the increase in use of diagnostic services may represent negative program effects.

It is also possible that Healthstyles did lead to reductions in use of health care services but not the ones evaluated in this thesis. For example, the Blue Cross and Johnson and Johnson work-site studies both reported significant reductions in hospital inpatient services, but these services were not assessed for Healthstyles[7] [8].

In addition, the data base used for the evaluation of Healthstyles is based primarily on physician visits and does not include use of non-physician health care workers such as nutritionists, exercise counsellors and psychologists. It is expected that the use of these services would also increase over the short-term as a consequence of health promotion. Consequently, the increases in health care costs observed in this study are probably lower than the real increases in costs to the health care system.

4. Length of follow-up period

It is possible that with a longer follow-up period the hypothesized program effects may have occurred. There is some evidence that the participant-control differences in health care use declined from the first to second years of follow-up. For example, there was some decline in the magnitude of group differences in medical consultations from year one to year two. However, for diagnostic services and total ambulatory there was no evidence of the participant-control difference decreasing over the follow-up period. Unfortunately, in those instances where reductions in participant-control differences are evident it is not possible to determine whether service use by participants will remain higher, become equal to control use or decrease such that they achieve significantly lower levels than controls in the long-term.

With respect to counselling and psychotherapy, a marginal, nonsignificant increase in the relative risk was observed (RR values of 1.53 and 1.57 for year 1 and year 2 respectively). It is possible that many mental health problems are difficult to correct and require on-going therapy over a longer period of time before reductions are achieved.

The continued elevation in use of these services is not necessarily negative, particularly given the relative unhappiness of the participants, and it is possible that it could lead to reductions in the use of other services in the future. Again a longer follow-up may show a subsequent decline in the use of these services.

A longer follow-up period may also have led to program effects on use of surgical and hospital services, as seen in work-site studies. While the two year follow-up period is longer than that of other community based programs, it is about three years shorter than the follow-up periods in the Johnson and Johnson and Blue Cross and Blue Shield studies[7 8 15 53]. Consequently, it may be beneficial to obtain more data from OHIP in order to study the effects of Healthstyles over a longer follow-up period.

POTENTIAL SOURCES OF BIAS

1. Quality of data

1.1 Coding errors: The claims data used for the analyses are expected to have been subject to errors in coding of patient encounters and inter-physician variation in coding. However, differential bias resulting from physician knowledge about the program is unlikely since the physicians coding the services were in no way connected to the program. Such nondifferential misclassification may have tended to bias the results towards the null hypothesis.

Errors in coding are unlikely to have affected the results for total number of ambulatory visits as these merge all types of consultations. The number of visits made

for each of the types of services are also unlikely to have been greatly affected as coding incorrectly by broad service type is unlikely. For example, an intermediate assessment may be coded as a minor assessment but will remain in the medical consultation category. Coding of an intermediate assessment as a diagnostic service is much less likely. However, costs of services are more likely to have been affected by coding errors.

1.2 OHIP - HSO differences: Potential differences between HSO and OHIP based practices with respect to cost and number of health services provided may have led to some bias in the results. HSO services and procedure codes such as minor, intermediate and general assessments, that correspond directly to one OHIP fee schedule code and are not priced on the basis of length of visit, do not pose a problem. However, application of the mean OHIP fees to HSO records may be invalid for some types of services. For example, the number and types of diagnostic tests per visits, and time spent on individual counselling and psychotherapy may differ systematically between the two systems leading to cost differences.

If Healthstyles led to a change in the number and/or cost of services for a procedure or service code (e.g. labwork) without altering the number of encounters for this particular service, the change in costs will be missed for HSO members. Since a larger proportion of participants (43%) than controls (25%) had records in the HSO file, the inability to detect changes in costs related to number, or types of services in a broad category may have biased the results either toward or away from the null hypothesis.

Health care practices may also have changed over the period of the study. If HSO and OHIP practices were affected differently by trends and policies in health care this difference, combined with the unequal proportions of HSO members by study group, may also have led to biased results.

In addition, there are some individuals who became members of an HSO at the time of, or after, entry into Healthstyles. For these individuals the baseline use would have been based on OHIP claims while later measures would have been based primarily on HSO records. If utilization patterns and costing differ between these two health care systems and the change in proportion of HSO members differs by group the results will be biased. It is possible that more participants than controls became HSO members after starting Healthstyles.

To control for the potential biases resulting from HSO and OHIP differences and the differing proportions of participants and controls in the HSO file, separate analysis of HSO and OHIP data could have been done. However, HSO members often had records in the claims file as well as the HSO file making it impossible to assess overall use of a given type of consultation through separate analysis of the two files. In addition, while the number of participants and controls represented by the two files together was provided with the data, the denominators for the HSO and claim files separately were not provided. Furthermore, the number of HSO members may have changed over the course of the study leading to changes in the number of people represented by the HSO file.

Overall, the differences between the HSO and OHIP records are expected to have

been minimal, especially with respect to number of encounter dates for a particular service. The effects of any systematic differences between the two coding systems would also have been somewhat diluted as the vast majority of the records are OHIP claims (135,368 OHIP records compared to 6,318 HSO records with a total of 7,834 HSO codes).

2. Internal validity

The nonequivalent control group design, with pre and post-test measures in both treatment and control groups, is the strongest of the quasi-experimental designs and is often recommended where nothing better is available[65]. However, there are a number of potential threats to the internal validity of this and others studies based on quasi-experimental designs[65]. These threats and their potential effects on the results of the Healthstyles study are summarized below:

2.1. History: History is a threat to validity when an event that is not part of the treatment occurs between pre and post-test and has an effect on the outcome measures.

In this study the control group is expected to have removed any effects of history. However, the lag between participant and control enrollment into the study may represent some differences in history by study group (selection-history interaction), especially in those cases where the control was enrolled several months after the participant. Since participants and controls were matched on a number of characteristics, were selected

from the same population and geographical region and most were enrolled within a few months of each other, the effects of this potential interaction should be minimal.

2.2. Testing: Changes due to the number of times responses are measured may be a threat to validity. In this study external records were used rather than self-reports. The potential biases arising from use of this data have already been discussed (Quality of data, p.88).

2.3. Instrumentation: Instrumentation bias is related to changes in the measuring instrument or shifts in metric at different points of the scale (unequal scale intervals) resulting in ceiling and/or basement effects.

As it is not possible to have fewer than zero visits, there was a strong basement effect and the impact of the program may have been difficult to detect among those who have low use of services at baseline. Since participants tended to have higher levels of use at baseline a selection x instrumentation interaction may also have occurred. However, this potential interaction was controlled through stratification of study subjects by baseline use of services.

2.4. Statistical regression: Statistical regression may occur if individuals are assigned to groups on the basis of pretest scores or correlates of these scores. Individuals with high scores at pre-test will tend to score lower at post-test and vice versa. Consequently, if pretest scores differ by group, differential pre-posttest effects may be due to regression

to the mean rather than the intervention.

In this analysis stratification by baseline use should have minimized statistical regression effects and confined them to the highest level baseline users. For total ambulatory use the effects of differences at baseline were controlled through ANCOVA.

2.5. Selection: Differences in the characteristics of individuals in experimental and control groups may account for apparent treatment effects and thus pose a threat to internal validity of the study.

While the matched control design insured that the participant and control groups were similar with respect to gender, age, education level and smoking status it failed to achieve similarities in stress and weight (tables 33 and 34). Baseline comparisons show that more participants than controls were heavy and "not too happy". These variables provide only a crude indication of potentially varied and complex differences between the study groups. The groups may also have differed with respect to other unmeasured characteristics, such as health status and attitudes towards use of health services. As such, there is the potential for differential maturation rates affecting use of services in absence of the program (selection x maturation interaction).

Table 33 Baseline characteristics: by study group

Characteristic	Unmatched analysis					Matched	
	P	n	C	n	p	p	# pairs
female	76.4%	520	76.5%	932	0.95	0.56	481
university education or more	51.9%	520	46.1%	931	0.07	0.37	481
current smokers	19.0%	520	16.9%	929	0.31	0.33	481
body type 1=underweight	2.2%	495	4.6%	862	<0.01	<0.01	426
5=excessive	28.3%		16.7%				
median	3.0		2.6				
Bradburn score 1=very happy	10.0%	518	19.7%	927	<0.01	<0.01	478
2=somewhat happy	60.6%		67.3%				
3=not too happy	29.4%		13.0%				
median	1.7		1.5				

P=particpant C=control

Table 34 Baseline characteristics: by study group

Characteristic	Unmatched analysis					Matched	
	P		C		p	p	# pairs
	mean	n	mean	n			
age (years)	40.4	520	39.8	932	0.36	<0.01	481
self reported weight (lbs.)	152.5	520	143.6	929	<0.01	<0.01	478
General Health Questionnaire	18.5	519	16.4	927	<0.01	<0.01	477

P=participant C=control

2.6. Mortality: Losses of study subjects over the course of a study may pose a threat to validity as effects may be due to the different types of people dropping out.

From the utilization data it was not possible to determine whether all subjects were available to use health care services in Ontario throughout the entire study period; individuals who were out of the country or who died over the course of the study could not be differentiated from those who were in province at all times.

One might hypothesize that controls would be more likely than participants to be out of the country at any given time as they were unwilling to commit to the Healthstyles program and were not involved in regular group support sessions. As a result, service use by controls may have been underestimated. However, unless this effect changed or cumulated over time it should not have led to serious bias. Furthermore, high response rates for the questionnaires at baseline, 3, 12 and 18 months (table 35) among participants and controls suggest that most study subjects were in-province at these times and biases due to death or using health care services elsewhere should be minimal. The response rates were actually higher among the controls than among the participants suggesting that the estimates for participant use of services during the follow-up period may be low compared to control values. If so, Healthstyles may have led to even greater increases in health care use than found in the analysis.

Differential mortality rates are not expected to have been a major threat to the validity of this study as participants and controls were matched on several demographic and health-related variables.

Table 35 Questionnaire response rates: by study group

Study group	Baseline	Month 3	Month 12	Month 18
Participants	640 (100%)	584 (91.1%)	518 (81.0%)	535 (83.5%)
Controls	1180 (100%)	1137 (96.3%)	1100 (93.0%)	1114 (94.4%)

2.7. Diffusion or imitation of treatments: Some diffusion or imitation of treatments may have occurred resulting in slightly conservative results (biased towards accepting the null hypothesis).

3. Statistical issues

3.1 Power

From the power curves it is apparent that this study had sufficient power to detect most of the differences of interest, when performing unstratified analyses (see Appendix III for power calculations). The power to detect an RR of 1.5 for use of preventive services and for use of counselling services was 98.6% (fig. 11) while a group difference of 25% for use of assessments and consultations, and diagnostic services could be detected with a power of close to 100% (fig. 12). A difference of 15% between participants and controls use of these services could be detected with a power of 76%. Although the latter power estimate is somewhat low, the point estimates for total ambulatory number and costs of visits indicate that the direction of the effects were not even in the hypothesized direction.

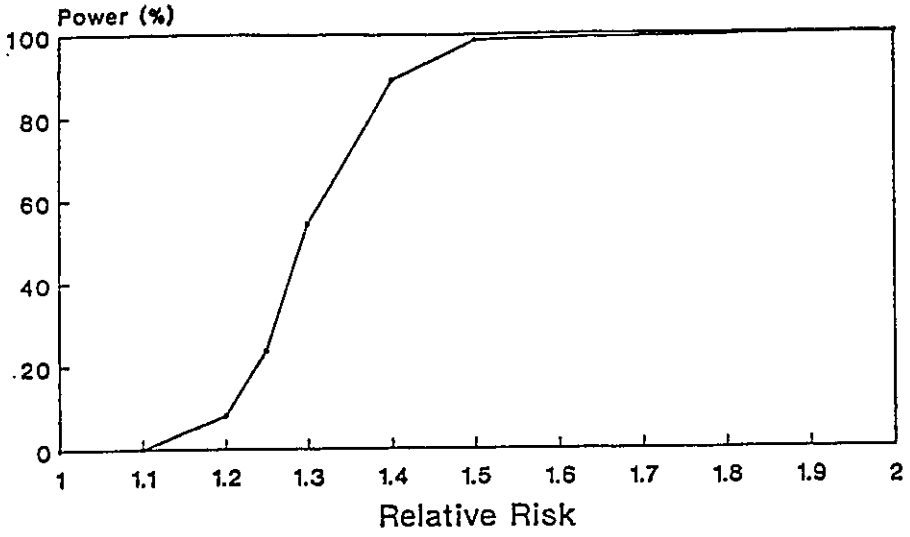


Fig. 11 Power curve for preventive services, and counselling

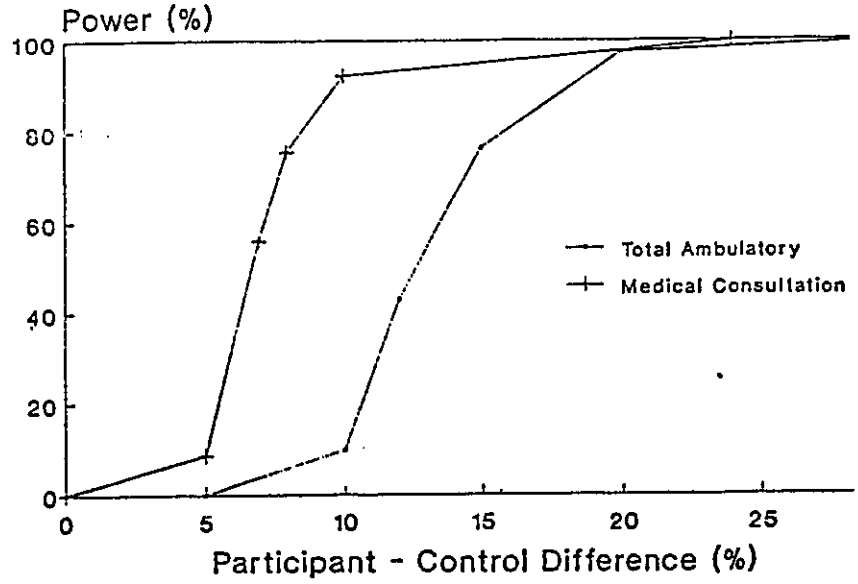


Fig. 12 Power curves for total ambulatory services and medical consultations

Although the stratification by baseline level of health care use led to smaller sample sizes, the power to detect the hypothesized group differences in diagnostic services and medical consultations remained high (99%).

3.2 Assumptions of statistical tests

Analysis of variance and t-tests are both based on a number of assumptions. The raw data in this analysis violated the homogeneity of variance assumption and the data were not normally distributed. However, the large sample size in the study compensates for the nonnormality (central limit theorem). In addition, since the sample sizes of the study groups were relatively equal, the tests should not have been greatly affected by the heterogeneity of variance. The F statistic is robust to heterogeneity of variance when the sizes of the study groups are approximately equal.

The statistical tests used are also based on the assumption on independent random samples. In this study the samples are neither independent nor random since participants and controls were matched and were volunteers. However, the sample might be considered a random sample of volunteers.

As previously mentioned the use of unmatched rather than matched analyses would have led to a reduction in random error and increased potential for bias (see Matching, p.66).

RECOMMENDATIONS FOR FUTURE STUDY

From the discussion of the results it is evident that evaluations of health promotion programs with longer follow-up periods and based on different target populations are needed. In order to advance the understanding of the effects of health promotion on health care and the potential of health promotion as a cost-containment measure, more research with respect to predictors of success in health promotion programs, predictors of health care use and the relationship between behaviours and health is also required.

Further analysis of data from the Healthstyles questionnaires and claims data may help to identify types of individuals for whom health promotion is successful. Although use was higher among the participant group as a whole, it is possible that the program led to reduced health care use for specific subgroups of the sample. The impact on use of health care services may differ by gender, age, education level, health locus of control, self-efficacy and other personal characteristics. This may form a topic for future study. Identification of subgroups that are more susceptible to the effects of health promotion would improve targeting and tailoring of similar programs in the future, ideally improving cost-effectiveness.

It would also be beneficial to assess the association between health behaviour changes and utilization. While the results from analysis of the behavioural outcomes indicate that Healthstyles had a positive impact on a number of health behaviours, it is possible that the changes were not great enough to lead to a modification of susceptibility

to illness and/or reduction of psychosomatic symptoms. In addition, certain health behaviours may be more highly related to health status and illness, and adequate changes may not have occurred with respect to these behaviours. Alternatively, particular threshold levels, rather than specific magnitudes of change in behaviours, may be required.

Finally, while it is believed that improvements in health behaviour should lead to improved general well-being and thus reduced use of health care services, multivariate studies of health care utilization illustrate that there are many factors that influence the need for, and use of services by the individual [3 4 35 36 38 63 64]. These other determinants may have a much greater role in prediction of health care use than health behaviours. The relative importance of these factors and how they are related as well as the determinants of health need to be understood in order to reduce use of health care services.

While reductions in health care use may be viewed as positive outcomes, it is also important to evaluate the appropriateness and timeliness of visits. One would expect that successful health promotion would lead to reduced use of services as a result of improved health , as well as reductions in inappropriate visits and increased self-treatment for minor illnesses. However, in some instances, individuals may fail to consult medical professionals when such a consultation would be advisable.

RECOMMENDATIONS FOR MORE EFFECTIVE PROGRAMS

The failure of the Healthstyles program to lead to reductions in use of health care services despite positive changes in health behaviours suggests a need to include education on appropriate use of services (physician and non-physician, preventive and illness related) in future health promotion programs. Given the relative unhappiness of the participant group, future programs could also place greater emphasis on improving mental well-being (self-esteem, stress management etc.). Finally, Healthstyles was a low-intensity program and it is possible that with a more intensive program that the results would have been more positive.

CONCLUSION

During the first year of follow-up in the Healthstyles study, participants had significantly higher levels of utilization than controls in the overall use of ambulatory care services while, as expected, there was no significant effect on use of surgical services. Within total ambulatory care, increases were limited primarily to use of psychotherapy and counselling services, and diagnostic services. There was little effect on use of medical consultations and no effect on use of preventive services.

Use of ambulatory services during the second year was also higher among participants. The program was not successful in increasing use of preventive services, reducing use of assessments and consultations, or reducing use of diagnostic services. Greater emphasis on appropriate use of services and longer follow-up periods may be required to observe significant reductions in the use of services by participants.

In terms of the on-going cost-effectiveness debate, the results do not support the argument that health promotion leads to reduced ambulatory costs, at least in the short-term. On the contrary, once the cost of the program is considered, Healthstyles in fact led to increased costs to the health care system. However, it is possible that some of these costs will be off-set by reduced health care use in the long-term. Evaluation of the long-term effects of health promotion on chronic disease incidence and accident prevention and the use of health care services related to these conditions, is required.

Finally, while participants in this study did not show significantly lower use of services than controls, the sample population was fairly young and had positive health

behaviours compared to the general population. The short-term effects on general well-being, occurrence of minor illnesses and use of the related health care services, need to be assessed for other populations which are more likely to have poor health and poor health behaviours at baseline.

REFERENCES

1. Deber RB, Hastings JEF, Thompson GG. Health care in Canada: current trends and issues. J Public Health Policy 1991;12(1):72-81.

2. Evans RG, Lomas J, Barer ML, Labelle RJ. Controlling health expenditures -the Canadian reality. N Engl J Med 1989;320(9):571-577.

3. Andersen R, Newman JF. Societal and individual determinants of medical care utilization in the United States. Milbank Mem Fund Q 1973;51:95-124.

4. Antonovsky A. A model to explain visits to the doctor: with specific reference to the case of Israel. J Health Soc Behav 1972;13:446-454.

5. Wood EA, Olmstead GW, Craig JL. An evaluation of lifestyle and absenteeism after two years in a worksite health promotion program. Am J Health Promotion 1989;4(2):128-133.

6. Gebhardt DL, Crump CE. Employee fitness and wellness programs in the workplace. Special Issue: Organizational psychology. American Psychologist 1990;45:262-272.

7. Gibbs JO, Mulvaney D, Henes C, Reed RW. Work-site health promotion. J Occup Med 1985;27(11):826-830.

8. Bly JL, Jones RC, Richardson JE. Impact of worksite health promotion on health care costs and utilization. JAMA 1986;256(23):3235-3240.

9. Epp J. Achieving health for all: a framework for health promotion. Ottawa: Minister of Supply and Services, Catalogue No.H39-102, 1986.

10. Russell L. Is prevention better than cure? Washington DC: The Brookings Institution, 1986.

11. World Health Organization. Ottawa charter for health promotion. 1980.

12. Health and Welfare Canada. Knowledge development for health promotion: a call for action. Ottawa: Minister of Supply and Services, Catalogue No. H39-147/1989E, 1989.

13. Perreault R, Malo R. Self-care: a review of the literature. In: Knowledge development for health promotion:a call for action. Ottawa, Canada: Minister of Supply and Services, 1989:

14. Sherman JB, Clark L, McEwen MM. Evaluation of a worksite wellness program: impact on exercise, weight, smoking, and stress. Public Health Nurs 1989;6:114-119.

15. Leigh JP, Richardson N, Beck R, et al. Randomized controlled study of a retiree health promotion program. Arch Intern Med 1992;152:1201-1206.

16. Matarazzo JD, Leckliter IN. Behavioral health:the role of good and bad habits in health and illness. In: Maes S, Spielberger CD, Defares PB, Sarason IG, eds. Topics in psychology. New York: John Wiley and Sons Ltd., 1988:3-18.

17. Stephens T. Health behaviour and health status. In: Rootman I, Warren R, Stephens T, Peters L, eds. Canada's health promotion survey technical report. Ottawa: Minister of Supply and Services, 1988:171-179.

18. Breslow L, Enstrom JE. Persistence of health habits and their relationship to mortality. Prev Med 1980;9:469-483.

19. Kaplan RM. Behavioral epidemiology, health promotion, and health services. Med Care 1985;23(5):564-583.

20. Barry PZ, DeFriese GH. Cost-benefit and cost-effectiveness analysis for health promotion programs. Am J Health Promotion 1990;4(6):448-452.

21. Green LW. Research agenda: building a consensus on research questions. Am J Health Promotion 1986;1(2):70-72.

22. Barsky AJ, Wyshak G, Klerman GL. Medical and psychiatric determinants of outpatient medical utilization. Med Care 1986;24:548-560.

23. Barsky AJ. Hidden reasons why some patients visit doctors. Ann Intern Med 1981;94:492-498.

24. Godkin MA, Rice CA. Psychosocial stress and its relationship to illness behaviour and illness encountered commonly by family practitioners. Soc Sci Med 1981;15E:155-159.

25. Berkanovic EB, Hurwicz M, Landsverk J. Psychological distress and the decision to seek medical care. Soc Sci Med 1988;27(11):1215-1221.

26. Puska P. Community-based prevention of cardiovascular disease: the North Karelia project. In: Matarazzo JD, Weiss SM, Herd JA, Miller NE, eds. Behavioral health: a handbook of health enhancement and disease prevention. USA: John Wiley and Sons Inc., 1984:1140-1147.

27. Breslow L, Fielding J, Herrman AA, Wilbur CS. Worksite health promotion: its evolution and the Johnson and Johnson experience. Prev Med 1990;19:13-21.

28. Puska P, Niemensivu H, Puhakka P, et al. Results of a one-year worksite and mass media based intervention on health behaviour and chronic disease risk factors. Scand J Soc Med 1988;16:241-250.

29. Spilman MA, Goetz A, Schulz J, Bellingham R, Johnson D. Effects of a corporate health promotion program. J Occup Med 1986;28(4):285-289.

30. Perry CL, Kelder SH, Murray DM, Klepp K. Communitywide smoking prevention: long-term outcomes of the Minnesota heart health program and the class of 1989 study. Am J Public Health 1992;82(9):1210-1216.

31. Farquhar JW, Fortmann SP, Flora JA, et al. Effects of communitywide education on cardiovascular disease risk factors. JAMA 1990;264(3):359-365.

32. Clifford PA, Tan S, Gorsuch RL. Efficacy of a self-directed behavioral health change program: weight, body composition, cardiovascular fitness, blood pressure, health risk, and psychosocial mediating variables. J Behav Med 1991;14(3):303-323.

33. Barr Taylor C, Fortmann SP, Flora J, et al. Effect of long-term community health education program on body mass index: the Stanford five-city project. Am J Epi 1991;134(3):235-249.

34. Murray DM, Kurth C, Mullis R, Jeffery RW. Cholesterol reduction through low-intensity interventions: results from the Minnesota heart health program. Prev Med 1990;19:181-189.

35. Andersen R, Aday LA. Access to medical care in the US. Med Care 1978;16(7):533-546.

36. Wolinsky FD, Johnson RJ. The use of health services by older adults. J Gerontol 1991;46(6):s345-s357.

37. Wolinsky FD. Assessing the effects of predisposing, enabling, and illness-morbidity characteristics on health service utilization. J Health Soc Behav 1978;19:384-396.

38. Briscoe ME. Why do people go to the doctor? sex differences in the correlates of GP consultation. Soc Sci Med 1987;5:507-513.

39. Robinson JO, Granfield AJ. The frequent consulter in primary care. J Psychosom Res 1986;30(5):589-600.

40. Katon W, Dengerink HA. Somatization in primary health care. In: Carr JE, Dengerink HA, eds. Behavioral science in the practice of medicine. New York: Elsevier Science Publishing Co., 1983:

41. Leavitt F. The health belief model and utilization of ambulatory care services. Soc Sci Med 1979;13A:105-112.

42. Strain LA. Uses of health services in later life: the influences of health beliefs. J Gerontol 1980;46(3):s143-s150.

43. Philips BU. Editorial:cost containment and health promotion. Health Values 1988;12(5):30-31.

44. Higgins CW. The economics of health promotion. Health Values 1988;12(5):39-45.

45. Guidotti TL. Health promotion in perspective. Canadian Journal of Public Health - Revue Canadienne De Santé Publique 1989;80:400-405.

46. Philips BU. Epidemiological issues in health promotion and cost containment. Health Values 1988;12(5):32-38.

47. Rentmeester RL. The economics of wellness promotion: values versus economics. Health Values 1984;8(5):6-9.

48. Shepard DS, Thompson MS. First principles of cost-effectiveness analysis in health. Public Health Rep 1980;94(6):535-543.

49. Genest M, Genest S. Psychology and health. Champaign Illinois: Research Press, 1987.

50. Shephard RJ, Corey P, Renzland P, Cox M. The influence of an employee fitness and lifestyle modification program upon medical care costs. Can J Public Health 1982;73:259-263.

51. Lynch WD, Golaszewski TJ, Clearie AF, Snow D, Vickery DM. Impact of a facility-based corporate fitness program on the number of absences due to illness. J Occup Med 1990;32(1):9-12.

52. Benson L, Nelson EC, Napps SE, Roberts E, Kane-Williams E, Salisbury ZT. Evaluation of the staying healthy after fifty educational program: impact on course participants. Health Educ Q 1989;16(4):485-508.

53. Vickery DM, Kalmer H, Lowry D, Constantine M, Wright E, Loren W. Effect of a self-care education program on medical visits. JAMA 1983;250(21):2952-2956.

54. Black A, McDowell I. Healthstyles:moving beyond disease prevention. Can Nurse 1984;80(4):18-20.

55. O'Hagan M. Healthstyles basics: lifestyle and behaviour change. Can Nurse 1984;80(4):21-23.

56. Centretown Community Health Centre , Health Care Research Unit University of Ottawa. Healthstyles, a report of a community health promotion demonstration project. Ottawa: Centretown Community Health Centre, 1986.

57. McDowell I, Black A, Collishaw N. The healthstyles health promotion program: description and behavioural outcomes. Can J Public Health 1988;79:447-454.

58. Peachey D, Linton A. What you should know about HSOs. Can Med Assoc J 1988;138:352-357.

59. Lamberts H. Problem behaviour in primary health care. J Roy Coll Gen Pract 1979;29:331-335.

60. Garfield SR, Collen MF, Feldman R, Soghikian K, Richart RH, Duncan JH. Evaluation of an ambulatory medical-care delivery system. N Engl J Med 1976;294:426-431.

61. Brown JW, Robertson LS, Kosa J, Alpert JJ. A study of general practice in Massachusetts. JAMA 1980;216(2):301-306.

62. Ontario Ministry of Health . Schedule of benefits. Toronto: Ontario ministry of health, 1986.

63. Mechanic D. Correlates of Physician utilization: why do major multivariate studies of physician utilization find trivial psychosocial and organizational effects. J Health Soc Behav 1980;

64. Rosner TT, Kevan HN, Wykle ML. Physician use among the old-old. Med Care 1988;26(10):982-991.

65. Cook TD, Campbell DT. Quasi-experimentation: design and analysis issues for field settings. USA: Rand McNally College Publishing Company, 1979.

66. Wallston KA, Wallston BS. Development of the multidimensional health locus of control (MHLC) scales. Health Educ Monogr 1978;6:160-170.

67. Goldberg DP. The detection of psychiatric illness by questionnaire. London: Oxford University Press, 1972.

68. Bradburn NM. The structure of psychological well-being. Chicago: Aldine Publishing Company, 1969.

69. Simon GE. Inroduction: concepts of somatization. In: Kirmayer LJ, Robbins JM, eds. Progress in Psychiatry number 31, Current concepts of somatization: research and clinical perspectives. Washington D.C.: American Psychiatric Press Inc., 1991:1-19.

70. Rothman RJ. Modern epidemiology. US: Little, Brown and Company, 1986.

71. Schlesselman JJ. Case-control studies. New York: Oxford University Press, 1982.

72. Kleinbaum DG, Kupper LL, Morgenstern H. Epidemiologic research. New York: Van Nostrand Reinhold, 1982.

73. Hennekens CH, Buring JE. Epidemiology in medicine. US: Little, Brown and Company, 1987.

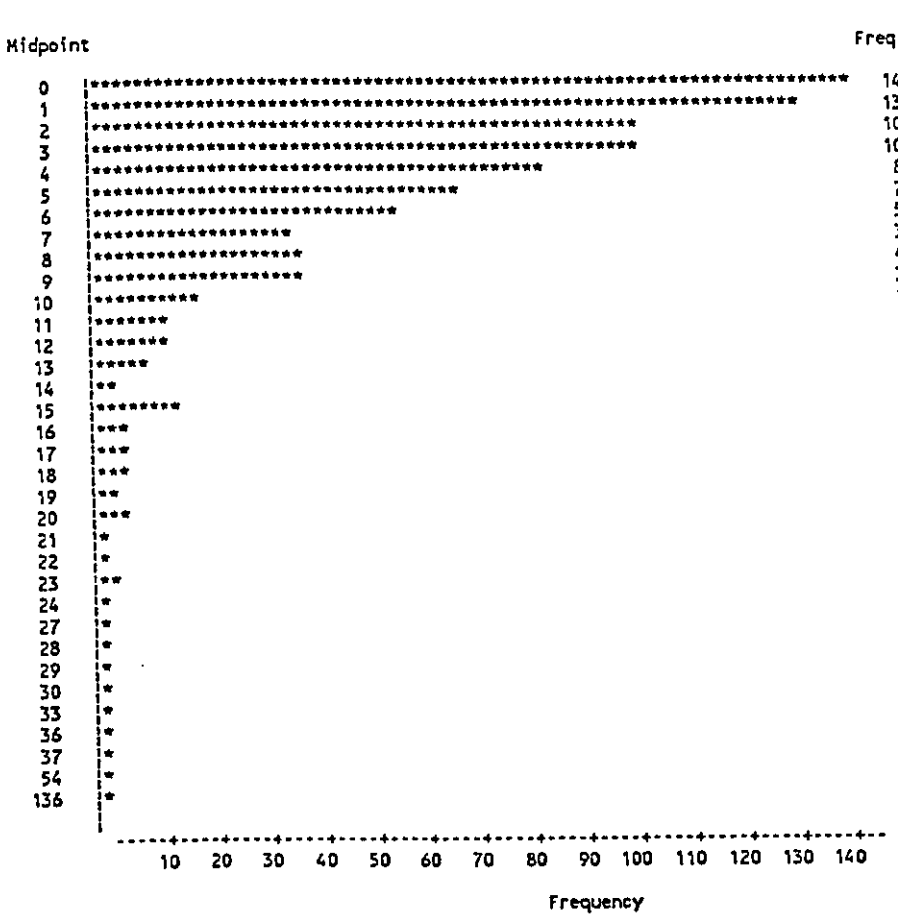
74. Beland F. Les biais systematiques dans la mesure de l'utilisation des services medicaux ambulatoires. Can J Public Health 1989;80(1):58-62.

75. Alexy BB. Factors associated with participation or nonparticipation in a workplace wellness center. Research In Nursing & Health 1991;14:33-40.

76. Kelsey JL, Thompson WD, Evans AS. Methods in observational epidemiology. New York: Oxford University Press, 1986.

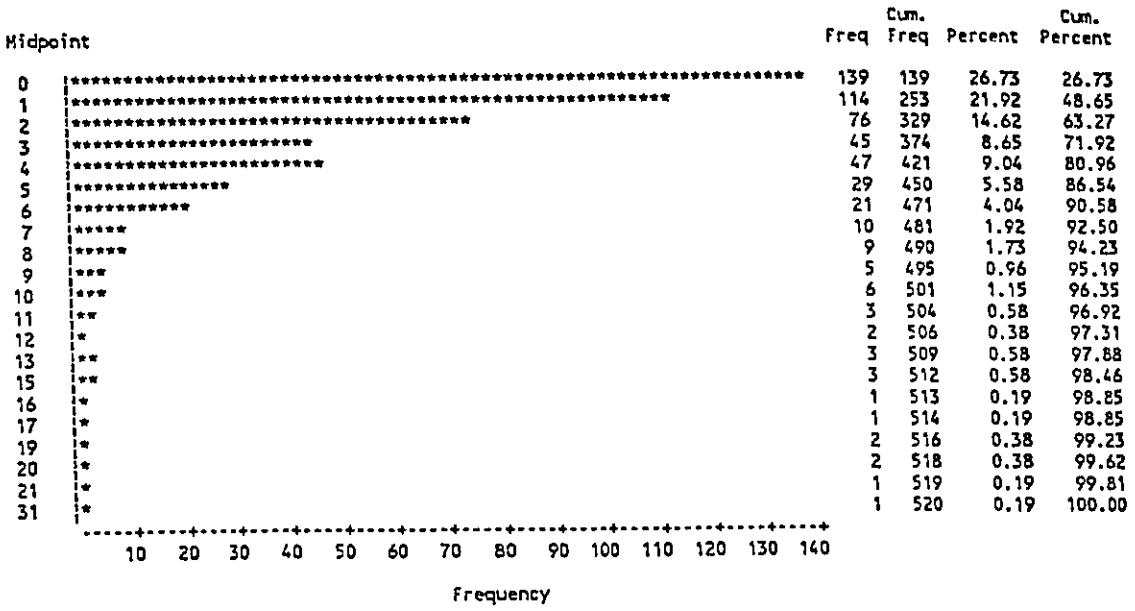
77. Hibbard JH, Pope CR. Gender roles, illness orientation and use of medical services. Soc Sci Med 1983;17(3):129-137.

APPENDIX I Data distributions for year 2 of follow-up

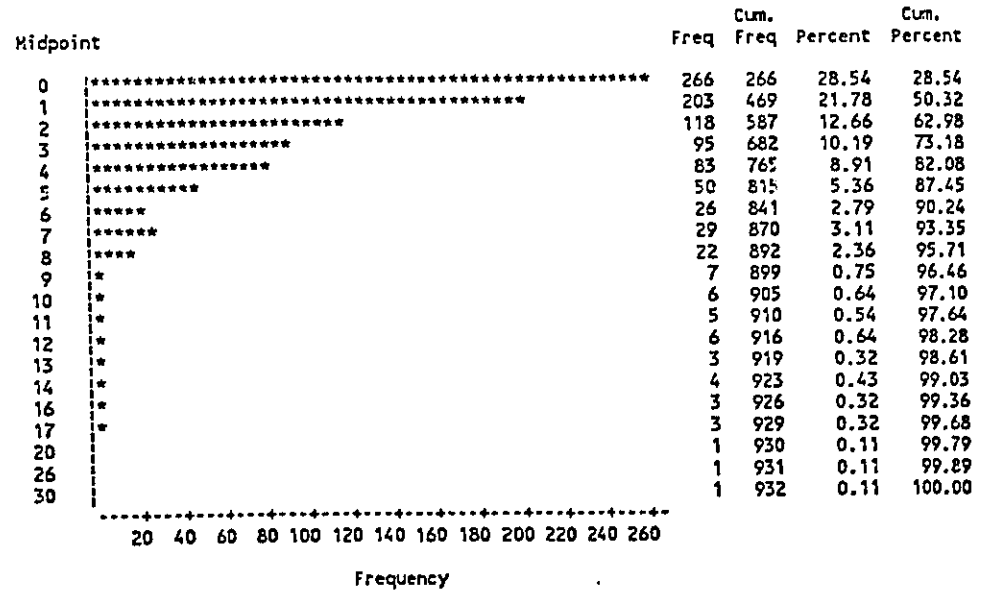


Midpoint	Freq	Cum. Freq	Percent	Cum. Percent
0	143	143	15.34	15.34
1	133	276	14.27	29.61
2	104	380	11.16	40.77
3	103	483	11.05	51.82
4	86	569	9.23	61.05
5	70	639	7.51	68.56
6	57	696	6.12	74.68
7	38	734	4.08	78.76
8	40	774	4.29	83.05
9	39	813	4.18	87.23
10	19	832	2.04	89.27
11	14	846	1.50	90.77
12	14	860	1.50	92.27
13	9	869	0.97	93.24
14	4	873	0.43	93.67
15	16	889	1.72	95.39
16	6	895	0.64	96.03
17	5	900	0.54	96.57
18	5	905	0.54	97.10
19	4	909	0.43	97.53
20	5	914	0.54	98.07
21	2	916	0.21	98.28
22	1	917	0.11	98.39
23	3	920	0.32	98.71
24	1	921	0.11	98.82
27	2	923	0.21	99.03
28	2	925	0.21	99.25
29	1	926	0.11	99.36
30	1	927	0.11	99.46
33	1	928	0.11	99.57
36	1	929	0.11	99.68
37	1	930	0.11	99.79
54	1	931	0.11	99.89
136	1	932	0.11	100.00

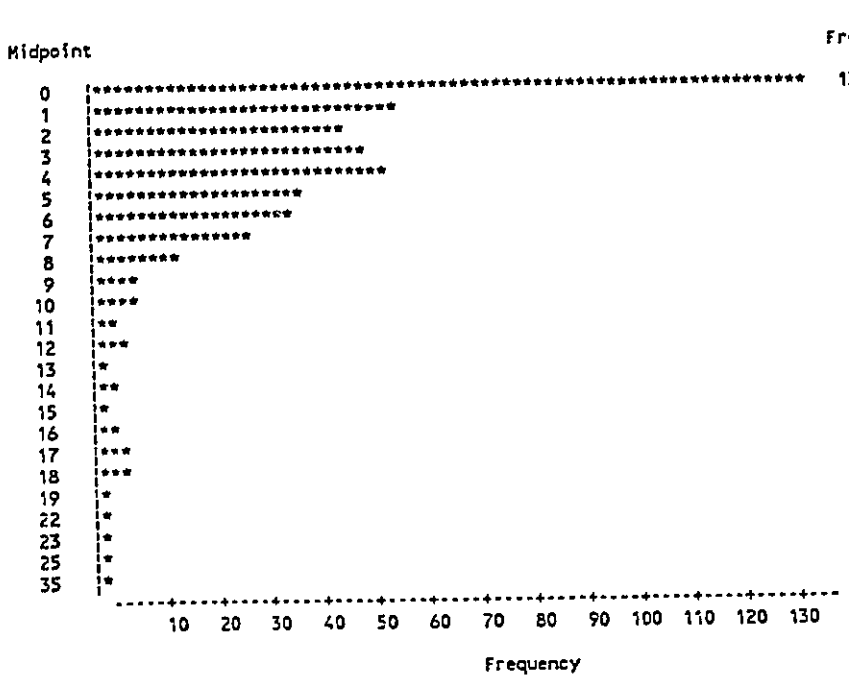
Visits for Ambulatory Services in Year 2: Controls



Visits for medical consultations in year 2: participants

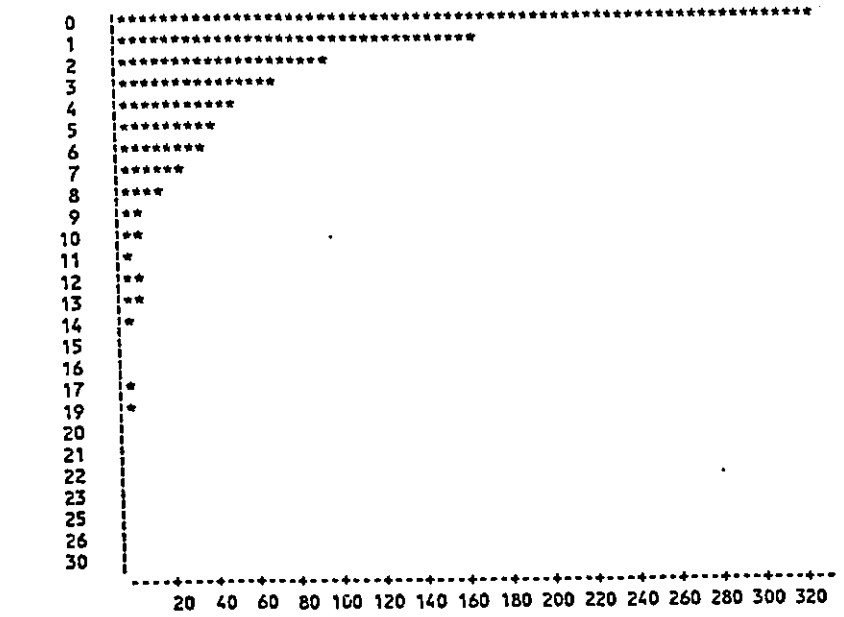


Visits for medical consultations in year 2: controls



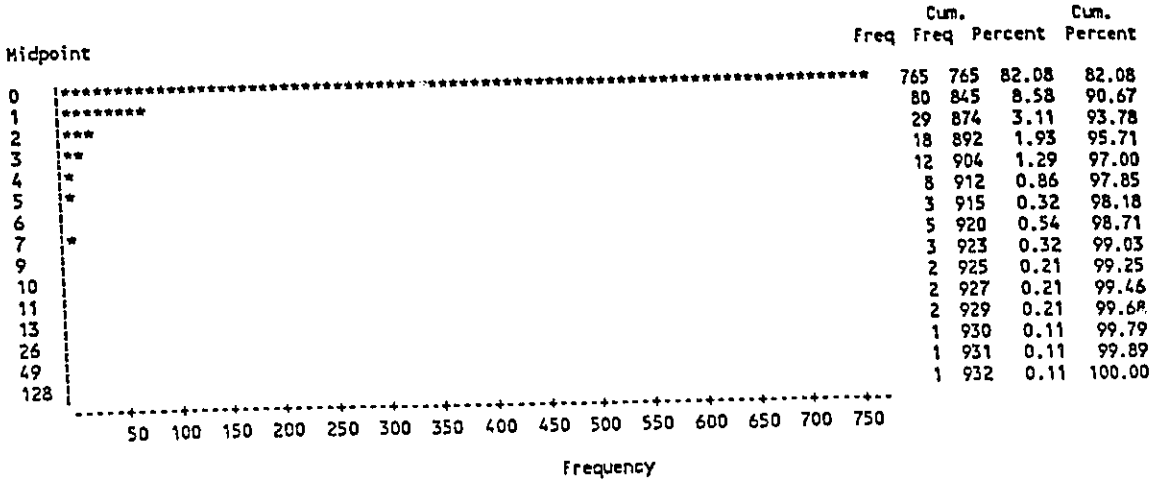
Midpoint	Freq	Cum. Freq	Percent	Cum. Percent
0	136	136	26.15	26.15
1	58	194	11.15	37.31
2	47	241	9.04	46.35
3	51	292	9.81	56.15
4	56	348	10.77	66.92
5	39	387	7.50	74.42
6	37	424	7.12	81.54
7	29	453	5.58	87.12
8	15	468	2.88	90.00
9	7	475	1.35	91.35
10	7	482	1.35	92.69
11	4	486	0.77	93.46
12	6	492	1.15	94.62
13	2	494	0.38	95.00
14	4	498	0.77	95.77
15	1	499	0.19	95.96
16	3	502	0.58	96.54
17	6	508	1.15	97.69
18	5	513	0.96	98.65
19	2	515	0.38	99.04
22	1	516	0.19	99.23
23	2	518	0.38	99.62
25	1	519	0.19	99.81
35	1	520	0.19	100.00

Visits for diagnostic services in year 2: participants



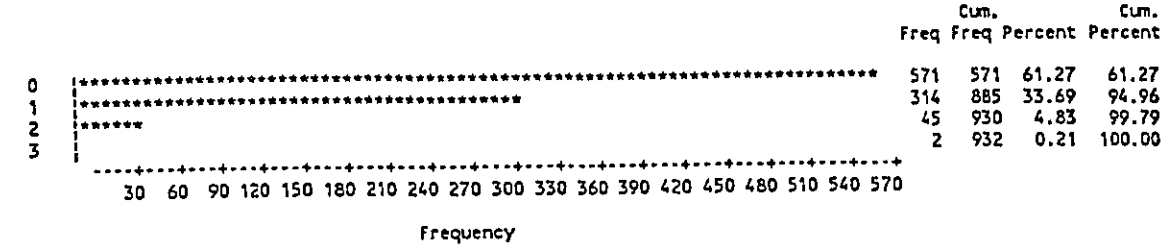
Midpoint	Freq	Cum. Freq	Percent	Cum. Percent
0	331	331	35.52	35.52
1	169	500	18.13	53.65
2	101	601	10.84	64.48
3	74	675	7.94	72.42
4	54	729	5.79	78.22
5	43	772	4.61	82.83
6	40	812	4.29	87.12
7	29	841	3.11	90.24
8	21	862	2.25	92.49
9	12	874	1.29	93.78
10	11	885	1.18	94.96
11	6	891	0.64	95.60
12	9	900	0.97	96.57
13	8	908	0.86	97.42
14	6	914	0.64	98.07
15	2	916	0.21	98.28
16	1	917	0.11	98.39
17	3	920	0.32	98.71
19	3	923	0.32	99.03
20	1	924	0.11	99.14
21	2	926	0.21	99.36
22	2	928	0.21	99.57
23	1	929	0.11	99.68
25	1	930	0.11	99.79
26	1	931	0.11	99.89
30	1	932	0.11	100.00

Visits for diagnostic services in year 2: controls



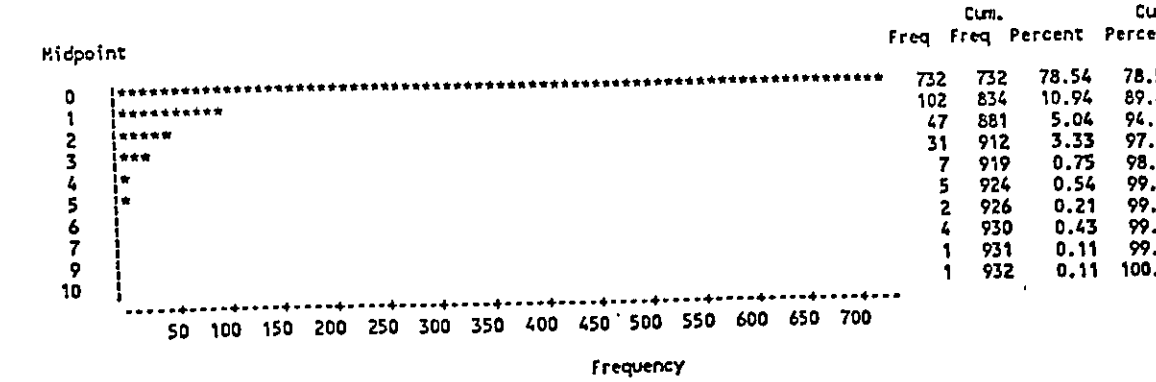
Freq	Cum. Freq	Percent	Cum. Percent
765	765	82.08	82.08
80	845	8.58	90.67
29	874	3.11	93.78
18	892	1.93	95.71
12	904	1.29	97.00
8	912	0.86	97.85
3	915	0.32	98.18
5	920	0.54	98.71
3	923	0.32	99.03
2	925	0.21	99.25
2	927	0.21	99.46
2	929	0.21	99.68
1	930	0.11	99.79
1	931	0.11	99.89
1	932	0.11	100.00

Visits for counselling and psychotherapy in year 2: controls



Freq	Cum. Freq	Percent	Cum. Percent
571	571	61.27	61.27
314	885	33.69	94.96
45	930	4.83	99.79
2	932	0.21	100.00

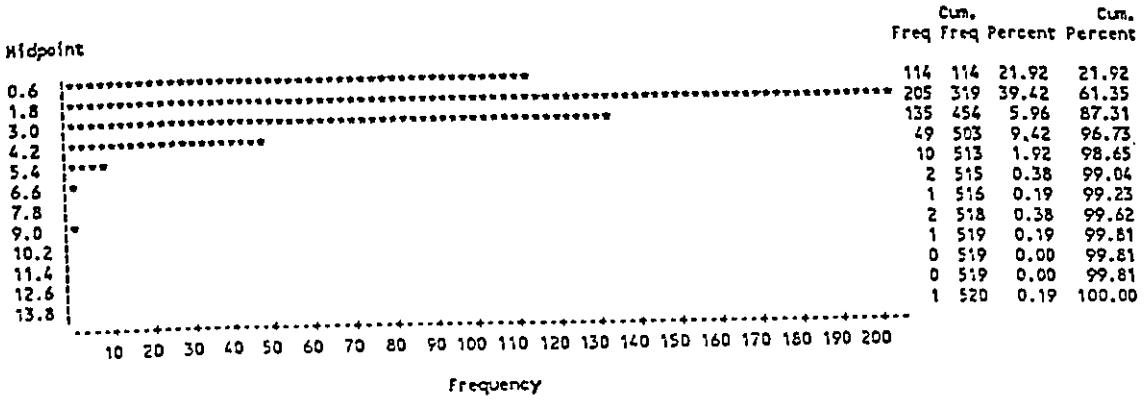
Visits for preventive services in year 2: controls



Freq	Cum. Freq	Percent	Cum. Percent
732	732	78.54	78.54
102	834	10.94	89.48
47	881	5.04	94.53
31	912	3.33	97.85
7	919	0.75	98.61
5	924	0.54	99.14
2	926	0.21	99.36
4	930	0.43	99.79
1	931	0.11	99.89
1	932	0.11	100.00

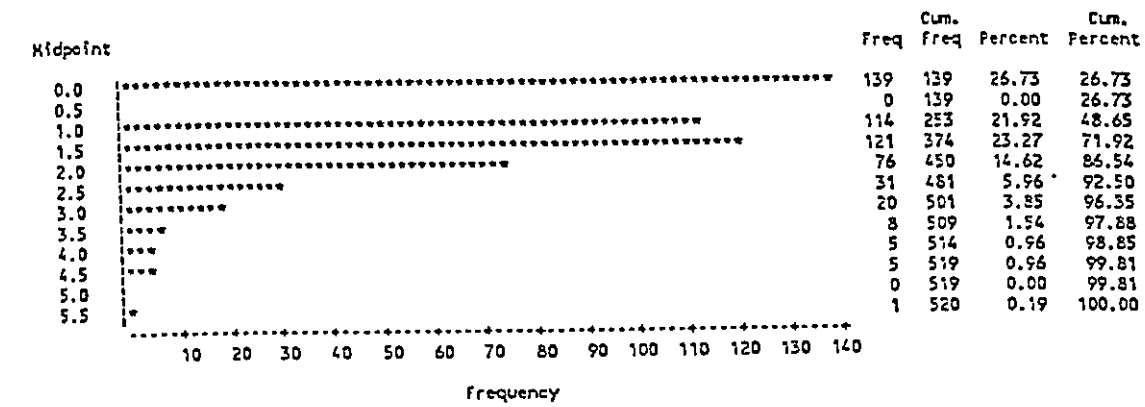
Visits for surgical services in year 2: controls

APPENDIX II Distribution of data after square root transformations



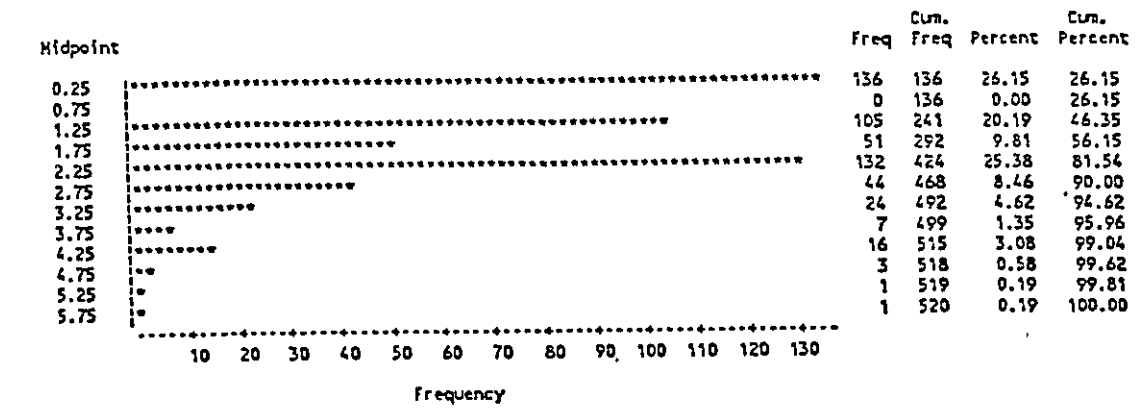
Midpoint	Freq	Cum. Freq	Percent	Cum. Percent
0.6	114	114	21.92	21.92
1.8	205	319	39.42	61.35
3.0	135	454	5.96	87.31
4.2	49	503	9.42	96.73
5.4	10	513	1.92	98.65
6.6	2	515	0.38	99.04
7.8	1	516	0.19	99.23
9.0	2	518	0.38	99.62
10.2	1	519	0.19	99.81
11.4	0	519	0.00	99.81
12.6	0	519	0.00	99.81
13.8	1	520	0.19	100.00

Square root number of visits for ambulatory services in year 2: participants



Midpoint	Freq	Cum. Freq	Percent	Cum. Percent
0.0	139	139	26.73	26.73
0.5	0	139	0.00	26.73
1.0	114	253	21.92	48.65
1.5	121	374	23.27	71.92
2.0	76	450	14.62	86.54
2.5	31	481	5.96	92.50
3.0	20	501	3.85	96.35
3.5	8	509	1.54	97.88
4.0	5	514	0.96	98.85
4.5	5	519	0.96	99.81
5.0	0	519	0.00	99.81
5.5	1	520	0.19	100.00

Square root number of visits for medical consultations in year 2: participants



Midpoint	Freq	Cum. Freq	Percent	Cum. Percent
0.25	136	136	26.15	26.15
0.75	0	136	0.00	26.15
1.25	105	241	20.19	46.35
1.75	51	292	9.81	56.15
2.25	132	424	25.38	81.54
2.75	44	468	8.46	90.00
3.25	24	492	4.62	94.62
3.75	7	499	1.35	95.96
4.25	16	515	3.08	99.04
4.75	3	518	0.58	99.62
5.25	1	519	0.19	99.81
5.75	1	520	0.19	100.00

Square root number of visits for diagnostic services in year 2: participants

Standard deviation, skewness and kurtosis:
before and after data transformations (Year 2 of follow-up)

Type of service	Variable	SD		Skewness		Kurtosis	
		P	C	P	C	P	C
Total	cost	925	392	11.6	18.5	158.1	435.4
	sqrt(cost)	11.24	7.49	4.0	2.8	29.6	25.9
	# visits	12.2	6.9	9.0	8.6	115.7	143.0
	sqrt(# visits)	1.5	1.2	1.8	0.9	10.0	4.4
Medical consultations	cost	63	56	3.2	2.6	16.5	10.6
	sqrt(cost)	4.28	4.17	0.6	0.5	0.5	-0.1
	# visits	3.6	3.2	2.9	2.7	12.9	12.3
	sqrt(# visits)	1.0	0.9	0.6	0.4	0.4	-0.1
Diagnostic services	cost	114	76	2.8	2.7	10.6	10.2
	sqrt(cost)	5.84	5.05	0.8	0.8	0.5	0.1
	# visits	4.5	4.0	2.2	2.5	7.3	8.7
	sqrt(# visits)	1.2	1.1	0.3	0.7	-0.3	-0.1

APPENDIX III Power Calculations

1. Preventive Services and Counselling

The power curves for detection of the association between participation in Healthstyles and use of services in the preventive and counselling categories were generated based on equation 1[76]. The proportion of controls using preventive services(p_0) was estimated to be 0.2 based on the study by Brown[61]. The proportion of controls using counselling services was also estimated to be 0.2. Thus the same power curve was used for both categories.

$$\text{Equation 1: } \sqrt{\frac{nd^2r}{(r+1)\bar{p}(1-\bar{p})}} - Z_\alpha$$

n=520 (number of exposed individuals)
r=2 (ratio of controls to participants)
z=1.96
p_0=0.20
$p_1 = p_0 \times RR$
$p = \frac{p_1 + rp_0}{r+1}$
d=difference in proportions ($p_0 - p_1$)

RR	Z	Power (%)
1.2	-0.13	0
1.3	0.74	54.1
1.5	2.46	98.6
2.0	3.97	99.9

2. Medical Consultations and Total Ambulatory Visits

Power to detect differences in mean number of visits made by participants and controls for minor illness care and total ambulatory care were calculated based on equation 2 [76]. From the literature the mean number of visits for minor illnesses per person per year ranges from 1.4 to 2.8[53] [25]. The mean value of 2.1 was used for the power calculations with an estimated variance of 1. Reports of the mean number of total ambulatory visits range from 2.37 to 6.20 per person per year [53] [22] [77]. A mid-range value of 4.5 visits among the unexposed was used in the generation of the power curves. While standard errors of 0.60 to 0.82 were reported by [22] a preliminary look at the data indicated a much higher standard deviation (>8). For the power calculations the standard deviation was estimated to be 4.

d = difference in mean number of visits
n = number of exposed individuals
r = ratio of controls to participants
$Z_\alpha = 1.96$

Total Ambulatory $\sigma = 4$, mean # visits in control group = 4.5
n=520 r=2

% difference	d	Z	Power (%)
10	0.45	0.13	10.0
15	0.68	1.18	76.2
20	0.90	2.23	97.4
30	1.35	4.32	99.9

Medical Consultations $\sigma = 1$, mean # visits in controls = 2.1

a) Power for unstratified analysis n=520, r=2

% difference	d	Z	Power (%)
8	0.17	1.17	75.4
10	0.20	1.76	92.2
15	0.32	3.90	99.9
25	0.53	7.81	99.9

b) Power for stratified analysis n=173, r=2

% difference	d	Z	Power (%)
10	0.20	0.30	23.6
15	0.32	1.35	82.3
20	0.42	2.35	98.1
25	0.53	3.62	99.9

^A