The Effects of a Prepaid Group Practice on Mental Health Outcomes of a General Population

Results from a Randomized Trial

Kenneth B. Wells, Willard G. Manning, Jr., R. Burciaga Valdez
The research described in this report was supported by the National Institute of Mental Health and by the Health Care Financing Administration, U.S. Department of Health and Human Services, Grant No. 99-C-98489/9-03.

Library of Congress Cataloging in Publication Data
Wells, Kenneth B., 1948-
   The effects of a prepaid group practice on mental health outcomes of a general population.

"R-3834-NIMH/HCFA."
"December 1989."
"Supported by the National Institute of Mental Health and the Health Care Financing Administration, U.S. Department of Health and Human Services."
RA790.6.W45 1989 362.2 89-24120
ISBN 0-8330-1001-8

The RAND Publication Series: The Report is the principal publication documenting and transmitting RAND's major research findings and final research results. The RAND Note reports other outputs of sponsored research for general distribution. Publications of The RAND Corporation do not necessarily reflect the opinions or policies of the sponsors of RAND research.

Published by The RAND Corporation
1700 Main Street, P.O. Box 2138, Santa Monica, CA 90406-2138
The Effects of a Prepaid Group Practice on Mental Health Outcomes of a General Population

Results from a Randomized Trial

Kenneth B. Wells, Willard G. Manning, Jr.,
R. Burciaga Valdez

December 1989

Supported by the National Institute of Mental Health and Health Care Financing Administration,
U.S. Department of Health and Human Services

RAND
PREFACE

This report compares mental health status outcomes of persons receiving care from a prepaid group practice to outcomes achieved with fee-for-service care. This work uses data from the RAND Health Insurance Experiment. This report represents a fuller version of work presented in Wells, Manning, and Valdez (in press). Among other RAND reports and papers that deal with prepaid group practices and mental health status are:


Valdez, R. B., The Effect of a Prepaid Group Practice on Child Health Outcomes Compared to Fee-for-Service Care, N-2618, 1989.

Ware, J. E., Jr., R. H. Brook, W. H. Rogers, E. B. Keeler, A. R. Davies, C. D. Sherbourne, G. A. Goldberg, P. Camp, and J. P. Newhouse, Health Outcomes for Adults in Prepaid and Fee-for-Service Systems of Care: Results from the Health Insurance Experiment, R-3459-HHS, 1987.

This report's research was supported by funding from the National Institute of Mental Health and the Health Care Financing Administration, U.S. Department of Health and Human Services. The opinions and conclusions expressed herein are solely those of the authors and should not be construed as representing the opinions or policy of any agency of the United States government.
SUMMARY

Prepaid group practices deliver less outpatient mental health care than do comparable fee-for-service plans. Given this difference, do participants enrolled in a prepaid group practice have different mental health outcomes than participants enrolled in fee-for-service plans? To answer this question, we used data from a randomized trial, the RAND Health Insurance Experiment. The study randomly assigned Seattle families to either a prepaid group practice—the Group Health Cooperative of Puget Sound (GHC)—or to one of several fee-for-service insurance plans that varied in the amount of cost sharing required.

We observed no statistically significant or clinically meaningful differences in general mental health outcomes between persons enrolled in the GHC and fee-for-service plans. Further, no insurance plan–related differences were observed for populations of special interest, including those with initially low mental health scores and/or low incomes.

Thus, the less intensive style of mental health treatment in the prepaid group practice is not associated with noticeable adverse effects on general mental health status.
ACKNOWLEDGMENTS

We would like to thank Audrey Burnam, Beth McGlynn, and Joseph Newhouse at RAND for their suggestions; Bernadette Benjamin for her meticulous programming and data management; Paul Widem (National Institute of Mental Health), and Thomas Kickham (Health Care Financing Administration) for their support. We are indebted to the Group Health Cooperative of Puget Sound—and especially to its director of research during the experimental period, Richard Handschin—for assistance in the trial’s implementation and suggestions on this report. Neither the Group Health Cooperative, The RAND Corporation, nor any of the above-named people or institutions necessarily agree with or endorse the findings we report here.

We dedicate this report to Carl Taube, formerly of Johns Hopkins and the National Institute of Mental Health. In addition to being a good friend and valued colleague, he facilitated our continued examination of the Health Insurance Experiment data under the RAND/HCFA Health Policy Research Center.
CONTENTS

PREFACE ......................................................... iii
SUMMARY ......................................................... v
ACKNOWLEDGMENTS .............................................. vii
TABLES .......................................................... xi
Section
  I. INTRODUCTION .............................................. 1
  II. METHODS .................................................. 4
      The HMO Study Design within the Health Insurance
      Experiment ................................................. 4
      Methods of Analysis ...................................... 7
      Threats to Validity ....................................... 12
  III. RESULTS .................................................. 16
  IV. DISCUSSION ............................................... 18
Appendix: STATISTICAL METHODS ............................ 21
REFERENCES .................................................... 25
TABLES

1. Seattle sample size ........................................... 6
2. Item content of Mental Health Inventory, adult
   version .................................................................. 8
3. Mental health status at enrollment by insurance plan .... 13
4. Reasons for failure to complete study in the Seattle
   area by plan .......................................................... 14
5. Mean differences among insurance plans in change in
   mental health status ............................................. 16
I. INTRODUCTION

Public policies over the past 20 years have promoted various alternatives to fee-for-service medical care delivery, most notably the health maintenance organization (HMO) (Enthoven, 1980). Previous studies have suggested that the prepaid group practice HMO achieves cost savings in the delivery of medical care (Gaus, Cooper, and Hirschman, 1976; Perkoff, Kahn, and Haas, 1976; Luft, 1981). They have indicated that the HMO reduces hospital admissions by some 40 percent, while general medical ambulatory visit rates remain similar to fee-for-service medical care. Other studies suggest that HMOs also lower mental health care costs primarily by providing fewer outpatient visits per user (Williams et al., 1979; Diehr et al., 1984).

Because HMOs reduce both use and costs of care, interest in this form of medical care delivery remains high. Federal entitlement programs (Medicare and Medicaid) and several states are testing HMO programs for their beneficiaries, including the seriously mentally ill (Lehman, 1987; Harris and Bergman, 1988).

Although researchers have studied HMOs' ability to save money, whether these reductions adversely affect health remains unclear. If reductions eliminate unnecessary services or increase the efficiency of care, then the health of beneficiaries could be affected positively by reducing the risk of iatrogenic conditions (conditions induced inadvertently by a physician or a physician's treatment). Alternatively, if effective services are reduced, health could be affected negatively. The large reduction in hospitalizations and differences in ambulatory mental health practice by HMOs could thus produce beneficial or adverse effects on health. Additionally, persons who are poor or sick may fare differently within HMOs than do others.

Despite the importance of these issues, few comparison studies on the effects of HMO and fee-for-service care on health outcomes—especially mental health outcomes—exist. Because such information on outcomes was unavailable and prior studies were limited to using self-selected samples, the federal government sponsored the RAND Health Insurance Experiment (HIE).

The HIE included a randomized controlled trial of one well-established HMO, the Group Health Cooperative of Puget Sound (GHC). This trial eliminated or reduced design problems of previous HMO studies. First, the HIE randomly assigned families to either the HMO plan or to a fee-for-service insurance plan, avoiding self-selection problems among the
comparison groups. Second, the HIE plans provided an identical benefits package for the HMO and fee-for-service participants, eliminating concerns related to noncomparability of coverage. Third, the HIE sampled participants from the nonelderly general population rather than relying on treated or other nonrepresentative populations. Previous studies on HMOs’ effects were limited by one or more of these issues.

Major HIE findings indicate that the HMO reduced costs of care by providing a different mix of ambulatory services and by reducing hospital admissions. First, total medical care expenditures were some 25 percent lower in the HMO than under comparable (free) fee-for-service coverage, almost entirely because of the HMO’s 40 percent probability of having an inpatient admission in a year (Manning, Leibowitz, Goldberg, et al., 1984). Second, total annual mental health outpatient expenditures in the HMO were one-third those of a comparable (free) fee-for-service plan, entirely because of the HMO’s fewer visits per user—not because of a smaller proportion of users per year (Wells, Manning, and Benjamin, 1986a). Third, although the annual probability of having any outpatient mental health visit was the same in each type of plan, a relatively higher proportion of HMO enrollees used some mental health care over several years. This difference can be accounted for by a higher turnover of new patients in the HMO than in fee-for-service plans (Manning, Wells, and Benjamin, 1986b).

Given these differences in use, we might expect differences in health outcomes. Ware et al. (1986) and Sloss et al. (1987) compared baseline and exit health status (outcome) measures for adults in the HMO and fee-for-service plans. Valdez et al. (1989) conducted a similar analysis for children. These three studies found few adverse effects on the typical HMO participants’ physical and mental health compared to fee-for-service participants’ health. Ware et al. (1986) reported two adverse effects of HMO participation (out of 26 comparisons): more bed days and a higher probability of reported serious symptoms. No mental health outcome differences were observed using a general measure of mental health status. Valdez et al. (1989) found that children assigned to the HMO exhibited worse general health and more behavior problems after three or five years than did children under fee-for-service care.

In this report, we reexamine the HIE data on the HMO’s effects on mental health for adults and children. A reassessment of these data is warranted for several reasons. First, because the HIE is the only major randomized trial of an HMO and continued interest in promoting HMOs exists, the HIE findings will inform both public and private policy decisions.

Second, the earlier analyses had less statistical power (that is, precision) for plan comparisons than are achievable with these data. Ware et
al. (1986) and Valdez et al. (1989) reported findings on outcomes assessed at the experiment's end (after three or five years). In the HIE, however, health status was assessed yearly. Because many psychiatric problems (for example, major depression) are episodic, we thought that the additional information on mental health outcomes from the middle years of participation might reveal differences between the HMO and fee-for-service plans that were not previously detectable. Further, the earlier results separated children and adults. By using all available data on adults and children, we increased the estimates' precision.

Third, the previous HIE analyses estimated the HMO's effects using the mental health index (MHI) (Veit and Ware, 1983), a summary indicator that aggregates information about two distinct dimensions, psychological distress (PSDS; symptoms of anxiety and depression) and psychological well-being (PWB; level of positive affect and quality of interpersonal ties). Based on earlier work, we know that psychological distress is much more predictive of mental health service use than is psychological well-being (Ware et al., 1984), and that effects of mental health status on use are similar in the HMO and fee-for-service plans (Wells, Manning, and Benjamin, 1986b). We reasoned that the HMO, by lowering use of mental health services relative to comparable fee-for-service coverage, would have stronger effects on psychological distress (for which people seem to be seeking care) than on psychological well-being (which has little independent relationship to use). The distress measure, similar to a measure of psychopathology severity, consists of items that assess the frequency and intensity of anxiety and depression symptoms. It is similar in content to the National Institute of Mental Health Center for Epidemiology Depression Scale (Radloff, 1977) and the General Health Questionnaire (Goldberg, 1972). It is increasingly in use as an outcome measure in medical and psychiatric research (Lehman, 1983; Smith, Monson, and Ray, 1986; Cassileth et al., 1984). In contrast, the interpretation of psychological well-being as a clinically meaningful component of mental health status remains uncertain. Therefore, we examined the HMO's effect on each dimension separately.
II. METHODS

THE HMO STUDY DESIGN WITHIN THE HEALTH INSURANCE EXPERIMENT

Seattle is the only site of six in the HIE with participants assigned to an HMO. The HMO, the Group Health Cooperative of Puget Sound, serves the greater Seattle, Washington, area.

The GHC, a nonprofit prepaid group practice, operated two hospitals, three mental health clinics, and 13 medical clinics in the Seattle area during the period under study. It had been operating for 30 years when the HIE began in Seattle in 1976. At that time, the GHC had enrolled approximately 15 percent of consumers in the Seattle area. The GHC encourages its primary care providers (for example, family practitioners) to provide mental health care or to coordinate care with formally trained mental health specialists. Patients at the GHC can obtain care from mental health specialists by referral from a general medical provider or on their own initiative.

In this study, we randomly assigned participants to either the GHC or to one of several fee-for-service plans that differed with regard to the level of required cost sharing. The participants were sampled from the Seattle area population that was not enrolled in the GHC in 1976 but that was otherwise eligible for the trial. Ineligible for the study were people over 62 at the time of enrollment, Seattle-area families whose incomes exceeded $61,000 (in 1985 dollars), the institutionalized, the military and their dependents, veterans with service-connected disabilities, and people eligible for disability Medicare and end-stage renal dialysis programs.

We assigned families to their plans using the Finite Selection Model (Morris, 1979). This model is designed to achieve as much balance across plans as possible while retaining randomization. It uses a

---

1Newhouse (1974) and Brook et al. (1979) fully describe the HIE design. Newhouse et al. (1979) discuss the measurement issues for the second generation of social experiments, to which the HIE belongs. Ware, Brook, Davies-Avery, et al. (1980) and Valdez (1986) discuss many aspects of data collection and measurement of health status.

2The study also included a random sample of individuals already enrolled at the GHC. That group served as a control group for estimating the magnitude of any adverse selection on observed differences in GHC use. We found no significant evidence of either favorable or adverse selection into the GHC (see Manning, Leibowitz, Goldberg, et al. [1984]; Manning, Wells, and Benjamin [1984a]; and Wells, Manning, and Benjamin [1986a] for a fuller discussion).

3This excluded about 1 percent of the families we contacted.
statistical method analogous to stratification to obtain greater comparability among the groups than would be expected if simple random assignment were used.

The covered services were identical for families assigned to the HMO and for those in the fee-for-service plans. These services included acute, chronic, and preventive ambulatory care; all hospital care; mental health services; visual and auditory services; prescription drugs; supplies; and all dental services except cosmetic orthodontia. Services of nonphysician providers such as audiologists, chiropractors, clinical psychologists, and speech therapists were also covered.4 The only noteworthy exclusion for this investigation was outpatient psychotherapy in excess of 52 visits per year per person.5

Families assigned to the Group Health Cooperative were not subject to out-of-pocket charges for services received within the HMO. Services unavailable from the HMO but available in the community (for example, chiropractic or dental) were fully covered. Services available from the HMO but acquired elsewhere were reimbursed at a rate of 5 percent of charges unless the services were for out-of-area emergency care or were approved referrals.6

Families participating in the fee-for-service plans were assigned to health insurance plans that had different levels of cost sharing. One group of participants received services with no out-of-pocket expenses (0 percent coinsurance); we refer to this plan as the “free” plan. Another group had plans that required a 25 or 95 percent coinsurance for all health services.7 We refer to these plans as the “family pay” plans. All plans with cost sharing limited out-of-pocket expenses (the maximum dollar expenditure [MDE]) to 5, 10, or 15 percent of family income, up to a maximum of $1000. Beyond the MDE, the insurance plan reimbursed all expenses in full. Finally, on one plan, families faced a 95 percent coinsurance rate for outpatient services, subject to a $150 annual limit on out-of-pocket expenses per person (or $450 per family). In this plan, all inpatient services were free, so that, in effect, this plan had an outpatient individual deductible. Hence, we refer to this plan as the “individual deductible” plan. Because the samples on

---

4See Claquin (1973) for a discussion of the reasons for the HIE structure of mental health benefits.

5Nonpreventive orthodontia and cosmetic surgery (not related to preexisting conditions) were also excluded.

6To encourage the filing of claims for noncovered out-of-GHC use, these participants were reimbursed a nominal 5 percent.

7One of these plans had different coinsurance rates for inpatient and ambulatory medical services (25 percent) than for dental and ambulatory mental health services (50 percent).
some fee-for-service plans are small, we aggregated the free and individual deductible (F/ID) plans\(^6\) for this analysis.

Table 1 shows the enrollment sample size and the number of observations (unique years of enrollment for each participant) by plan. Because children under five years old at the time of enrollment do not have a baseline measure of mental health status, we have excluded them from these analyses. Thus, the numbers in Table 1 are lower than have been reported in some other publications based on the full HIE sample.

Families were enrolled in the insurance plans as a unit, with only eligible members participating. To reduce plan-related refusals, the HIE paid families on the fee-for-service plans a lump sum equal to their worst-case out-of-pocket expense; the size of the payment was independent of their use of health services during the experiment.\(^9\) Therefore, no family was financially worse off for participating in the study. A family’s previous insurance coverage was maintained by the HIE during the study, with the benefits of the policy assigned to the

Table 1

<table>
<thead>
<tr>
<th>Plan</th>
<th>Initial Enrollment</th>
<th>Estimation Sample (Person-Years)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prepaid group practice</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group Health Cooperative</td>
<td>1026</td>
<td>3040</td>
</tr>
<tr>
<td>Fee-for-service</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Free and individual deductible</td>
<td>625</td>
<td>1542</td>
</tr>
<tr>
<td>Family pay</td>
<td>455</td>
<td>1145</td>
</tr>
<tr>
<td>Total</td>
<td>2106</td>
<td>5727</td>
</tr>
</tbody>
</table>

\(^6\)To avoid precision problems resulting from small sample sizes, we have grouped these two plans together. Such a grouping is plausible for two reasons. First, our prior work indicates the use of outpatient mental health care is not significantly different for these two plans (Wells, Manning, Duan, et al., 1982; Manning et al., 1984; Keeler et al., 1986). Second, the two plans exhibit similar mental health status responses using fee-for-service data from all six sites (Wells, Manning, and Valdez, 1989).

\(^9\)Because the payment did not vary with use, participants should have treated the lump sum as a transient change in income. Preliminary evidence suggests that a very small amount of this money is actually used for health care (Newhouse, Manning, Morris, et al., 1981).
HIE. If the family had no coverage, the HIE purchased a policy on its behalf.\textsuperscript{10}

Half the GHC participants were enrolled for five years; the remainder, for three years. One-quarter of the fee-for-service participants were enrolled for five years; the remainder, for three years. Assignment to three- or five-year participation was made at random. Because the participation period differs, on average, for persons in the GHC and fee-for-service plans, we adjusted all comparisons of mental health status by the amount of time the individual participated.

\section*{METHODS OF ANALYSIS}

\subsection*{Mental Health Status Measures}

In this report, we analyze three outcome measures developed for the HIE that assess mental health outcomes: the mental health index, the psychological distress index, and the psychological well-being index (Veit and Ware, 1983). We collected data for these measures from a medical history questionnaire that was self-administered before the experiment began (at enrollment), annually, and three or five years later (at exit).

The Mental Health Inventory was designed specifically to measure mental health status in the HIE (Veit and Ware, 1983). The adult version consists of 38 self-administered items (see Table 2) adapted from the General Well-Being Schedule (Dupuy, 1974) and other measures (Costello and Comrey, 1967; Dohrenwend et al., 1980; Beck, 1967). A similar measure was developed for children aged 5 to 13, based on 12 questionnaire items completed by a parent (Eisen et al., 1980). The items assess the frequency and intensity of symptoms of both psychological distress (for example, anxiety and depression) and psychological well-being. These subdimensions are each represented by a multi-item scale; the scales combine (for adults, with two additional items) to form the summary MHI score.

Each index is scored on a scale ranging from 0 to 100. For both the MHI and the PWB index, a higher score indicates better mental health. For the PDS index, a higher score indicates poorer mental health. The correlation between psychological distress and psychological well-being is $-0.75$, which is substantially lower than the internal consistency reliability estimates of 0.92 to 0.96 for the HIE population, indicating that although these measures overlap, they also assess

\textsuperscript{10}Thus, no family could become uninsurable as a result of its participation in the study.
Table 2

ITEM CONTENT OF MENTAL HEALTH INVENTORY, ADULT VERSION

<table>
<thead>
<tr>
<th>Anxiety</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very nervous person</td>
</tr>
<tr>
<td>Bothered by nervousness</td>
</tr>
<tr>
<td>Felt tense or high-strung</td>
</tr>
<tr>
<td>Anxious, worried</td>
</tr>
<tr>
<td>Difficulty trying to calm down</td>
</tr>
<tr>
<td>Nervous or jumpy</td>
</tr>
<tr>
<td>Restless, fidgety, impatient</td>
</tr>
<tr>
<td>Rattled, upset, flustered</td>
</tr>
<tr>
<td>Hands shake when doing things</td>
</tr>
<tr>
<td>Relax with difficulty</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Depression</th>
</tr>
</thead>
<tbody>
<tr>
<td>Moody, brooded about things</td>
</tr>
<tr>
<td>Low or very low spirits</td>
</tr>
<tr>
<td>Felt downhearted and blue</td>
</tr>
<tr>
<td>Felt depressed</td>
</tr>
<tr>
<td>Strain, stress, pressure</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Loss of Behavioral/Emotional Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control behavior, thoughts, feelings</td>
</tr>
<tr>
<td>Concern about losing control of mind</td>
</tr>
<tr>
<td>Felt emotionally stable</td>
</tr>
<tr>
<td>Nothing turns out as wanted</td>
</tr>
<tr>
<td>Felt like crying</td>
</tr>
<tr>
<td>Better off if dead</td>
</tr>
<tr>
<td>Down in the dumps</td>
</tr>
<tr>
<td>Think about taking own life</td>
</tr>
<tr>
<td>Nothing to look forward to</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>General Positive Affect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Happy person</td>
</tr>
<tr>
<td>Happy, satisfied or pleased</td>
</tr>
<tr>
<td>Daily life interesting</td>
</tr>
<tr>
<td>Felt calm and peaceful</td>
</tr>
<tr>
<td>Felt cheerful, lighthearted</td>
</tr>
<tr>
<td>Generally enjoyed things</td>
</tr>
<tr>
<td>Relaxed and free of tension</td>
</tr>
<tr>
<td>Living a wonderful adventure</td>
</tr>
<tr>
<td>Expect an interesting day</td>
</tr>
<tr>
<td>Wake up fresh, rested</td>
</tr>
<tr>
<td>Future hopeful, promising</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Emotional Ties</th>
</tr>
</thead>
<tbody>
<tr>
<td>Felt loved and wanted</td>
</tr>
<tr>
<td>Love relations full, complete</td>
</tr>
<tr>
<td>Time felt lonely</td>
</tr>
</tbody>
</table>

NOTE: All items in this table are positively scored.
unique information. A substantial proportion of the variance in these indices is maintained from year to year. One-year stability coefficients range from 0.61 to 0.69. These coefficients can be interpreted as the proportion of true score variance that remained unchanged during the year. Comparisons of internal consistency and stability estimates demonstrate that these indices are sensitive to changes rather than that they reflect a stable unchanging trait over time.

The MHI has excellent construct validity (Veit and Ware, 1983; Ware, Johnston, et al., 1979; Williams, Ware, and Donald, 1981). The total MHI score strongly predicts the use of outpatient mental health services. For example, the HIE participants in the bottom third of the MHI at baseline used nearly three times as much outpatient mental health services as did those in the upper third (Ware et al., 1984).

To facilitate interpretation of an annual change in mental health scores, we estimated the impact of common negative life events in a given year on change in mental health status over the subsequent year, controlling for prior mental health status. In these analyses, we used the longitudinal data on all the fee-for-service participants in all six sites. We used multiple regression to estimate one-year change in mental health status. The main independent variable was a life event in the previous year—for example, being fired. We included the following variables as covariates: age, sex, and mental health status in the prior year. We found that being fired or laid off in the previous year lowered overall MHI by 2.34 scale points (standard error = 0.35); lowered the PWB index by 2.85 (standard error = 0.61) scale points; and increased the PSDS index by 1.66 (standard error = 0.41) scale points. We obtained similar results using other life events. We used these scale differences as a metric for a clinically meaningful change in mental health status (see Wells, Manning, and Valdez, 1989).

For each mental health status measure, we defined initially “sick” and “well” groups as those in the the lowest and highest thirds of the index’s distribution at enrollment. The groups are the same as those used in the analysis of fee-for-service participants from all six sites (Wells, Manning, and Valdez, 1989).

**Independent Variables**

We used three groups of independent variables: insurance plan, initial mental and physical health status measures, and sociodemographic measures.

**Insurance Plan.** We used dummy variables to represent each of the following insurance plan groups: the GHC plan; the free or indi-
individual deductible plans; and the family pay plans. For most of our analyses, the F/ID plans served as the omitted or contrast group.

**Health Status Measures.** We used three measures of baseline health status to increase the precision of our plan estimates: general health perceptions, physical or role limitations, and mental health status. Each of these measures was based on information gathered from the self-administered medical history questionnaire (parents served as proxy respondents for children under 14 years of age). All health status data were collected at the beginning of the study; a summary description of each is presented below.

The general health index (GHINDEX), based on 22 questionnaire items for individuals aged 14 and over and 7 items for children, provides a favorably scored summary of health perceptions. We scored the GHINDEX using a simple summated ratings method with scores ranging from 0 to 100. These ratings are considered measures of general health because they do not focus on a specific health status attribute and are associated with a wide range of physical and mental health concepts. Items assess perceptions of health in general at the present, in the past, and in the future, as well as beliefs about resistance to illness and health worry (Ware, 1976; Davies and Ware, 1981; and Eisen et al., 1980).

Among the more widely used measures of physical health are those that assess limitations in functioning caused by poor health. We used a measure of role or physical limitation that is scored dichotomously (1 = limited; 0 = otherwise) to indicate the presence of one or more limitations. Thus, we assigned persons completely free of limitation in their usual activities a score of 0 and those reporting one or more limitations a score of 1. The measure is based on 12 questionnaire items for adults (Stewart et al., 1977, 1978, 1981a, 1981b) and 5 items for children (Eisen et al., 1980).

**Sociodemographic and Other Factors.** To enhance the precision of our estimates and to adjust for any residual plan imbalance, we also included covariates for age, sex, race, and family income. In addition, we included a dummy variable for type of questionnaire—adult or child—because children under age 14 may have used both forms during the course of the study. Adolescents 14 years and older provided self-reports on our annual medical history questionnaire, whereas parents provided proxy responses for younger children.

Except for family income, we collected the data prior to or at enrollment. Our measure of income, collected during the first year of the study, is the ratio of family income to the square root of family size. It represents disposable income per person.
Statistical Methods

We used both analysis of variance (ANOVA) and least-squares regression methods to estimate the effects of insurance coverage on changes in mental health status—the summary mental health index, psychological distress index, and psychological well-being index. Annual assessments of these mental health measures provided multiple years of observations for each person in the study. We calculated changes as the difference between mental health status measured at enrollment and the end of each study year.

Differences were stated in terms of annual changes, whether they occurred in the first or last year of the study. We adjusted differences for the amount of time elapsed between that year's measurements and the enrollment measurements: Each difference was divided by the amount of time from the beginning of the study to the time of assessment. This adjustment was important for this application because of the time-at-risk difference among plans. Because more GHC participants were enrolled for five years, failure to correct for time at risk might have given the appearance of a difference among the plans when no such true difference existed, or it might have masked a true difference.

By using differences, we allowed each individual to act as his own control. To the extent that some individuals may have had a stable tendency to have higher or lower mental health status outcomes, differencing removed this effect. Our estimates were more precise than if we had simply regressed the assessed outcome on insurance plan, income, and other covariates. Differencing also tends to purge the results of any residual imbalance among the plans that exists after randomization, such as might occur from refusal of the initial offer or subsequent attrition from the study.

We used two approaches to estimate the insurance plan effects. First, we compared mean difference scores for participants on each plan using ANOVA and least-squares regression. Second, we predicted the average difference based on the estimated parameters of the least-squares regression model. The predictions were standardized to the enrollment sample. The second procedure was analogous to an age-sex adjustment, but corrected for any imbalance by the plan in all measured characteristics. We presented plan contrasts for both ANOVA (unadjusted means) and predicted differences (adjusted means) for each of the three mental health status measures.
No statistically significant interactions between mental health status, income, and insurance plan were observed. Our final prediction models do not include two- and three-way interactions between initial mental health status, family income, and insurance plan (for example, GHC × initial MHI, or GHC × initial MHI × family income). We tested these interactions in preliminary analyses and they were insignificant. In preliminary analyses, we also tested interactions among age, insurance plan, and initial MHI; these interactions were not significant. Therefore, we dropped them from our final models.

Our data exhibit positive correlations among individuals in the same family and among observations for the same individual over time. Such correlations affect the inference statistics and standard errors. To obtain inference statistics adjusted for these correlations, we modeled them using a nested variance component or intraclass cluster model (Maddala, 1971; Searle, 1971).

The appendix provides additional details on our estimation approach.

THREATS TO VALIDITY

Two major threats could potentially bias our estimates of mental health outcomes or lead to incorrect inferences about differences in mental health between groups. First, offers to participate could have been accepted by different kinds of people, whose characteristics could have affected the outcome. Second, individuals could have dropped out of the different plans at different rates as a function of their health. In addition, missing data from participants could also have biased results. We used several strategies to detect and—when present—counter any bias these problems may have introduced.

Refusals

Most people who refused to participate in the trial did so before we made assignments to experimental treatments. In all, 29 percent of those persons originally contacted refused to participate in preliminary interviews.

Of those people who were assigned to a plan and to whom an offer was made to participate, 21 percent refused to participate in the GHC sample, 5 percent refused to participate in the free plan, and 19 percent refused to participate in the family pay and individual deductible plans.

Characteristics of families who refused participation were similar to those who accepted the offer. Health status, prior use patterns, demographic variables, family size and income, race, and AFDC (Aid to
Families with Dependent Children status did not differ significantly among those who enrolled in the fee-for-service groups and the Group Health Cooperative (Manning, Leibowitz, Goldberg, et al., 1984).

The enrollees in the various groups did not differ significantly in their mental health status at the time of enrollment (see Table 3). This result held true for measures of overall mental health status, as well as for psychological distress and psychological well-being. Our regression models included both pre-HIE health status and scores for other variables known to predict health outcomes so that we could control statistically for measured differences between persons assigned to different plans.

Table 3
MENTAL HEALTH STATUS AT ENROLLMENT
BY INSURANCE PLAN

<table>
<thead>
<tr>
<th>Plan</th>
<th>Mental Health Index</th>
<th>Psychological Well-being</th>
<th>Psychological Distress</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prepaid group practice</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group Health Cooperative</td>
<td>75.35</td>
<td>65.42</td>
<td>19.13</td>
</tr>
<tr>
<td></td>
<td>(0.492)</td>
<td>(0.639)</td>
<td>(0.464)</td>
</tr>
<tr>
<td>Fee-for-service plans</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Free and individual deductible</td>
<td>74.91</td>
<td>64.45</td>
<td>19.14</td>
</tr>
<tr>
<td></td>
<td>(0.686)</td>
<td>(0.823)</td>
<td>(0.666)</td>
</tr>
<tr>
<td>Family pay</td>
<td>75.47</td>
<td>65.02</td>
<td>18.88</td>
</tr>
<tr>
<td></td>
<td>(0.848)</td>
<td>(1.073)</td>
<td>(0.768)</td>
</tr>
<tr>
<td>Chi-square(2)</td>
<td>0.35</td>
<td>0.88</td>
<td>0.09</td>
</tr>
</tbody>
</table>

NOTE: Mean standard error. A higher value indicates better mental health on the mental health index and for psychological well-being; a higher value indicates poorer mental health status for psychological distress. Standard errors are corrected for intrafamily correlation.

Sample Loss

The sample we used in our analyses consisted of enrolled participants while they remained in the experiment and in the Seattle area. People who moved from the Seattle area could obviously no longer receive services at the GHC. To maintain comparability with the GHC participants, those in the fee-for-service plans who moved from the Seattle area were also omitted from our analyses starting at the time of their move. Moves from the Seattle area account for about half of the partial years of participation (see Table 4).
Table 4
REASONS FOR FAILURE TO COMPLETE STUDY
IN THE SEATTLE AREA BY PLAN
(In percent)

<table>
<thead>
<tr>
<th>Reason</th>
<th>Prepaid Group Practice</th>
<th>Fee-for-Service</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Group Health Cooperative</td>
<td>Free</td>
</tr>
<tr>
<td>Voluntarily withdrew</td>
<td>4.4</td>
<td>0.2</td>
</tr>
<tr>
<td>Terminated because of failure</td>
<td>2.5</td>
<td>3.0</td>
</tr>
<tr>
<td>to meet study obligations</td>
<td>0.3</td>
<td>0.2</td>
</tr>
<tr>
<td>Died</td>
<td>22.3&lt;sup&gt;a&lt;/sup&gt;</td>
<td>20.7&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Moved from Seattle area</td>
<td>0.5</td>
<td>0.2</td>
</tr>
<tr>
<td>Other</td>
<td>Total not completing study</td>
<td>29.9</td>
</tr>
</tbody>
</table>

SOURCE: Manning, Wells, and Benjamin (1986a).

NOTE: The column numbers do not necessarily add to the total because of rounding error. The sample for this table consists of the 3095 individuals initially enrolled in the study (young children are included in the table but excluded from the present analysis).

<sup>a</sup>These individuals were kept in the experiment; the Group Health Cooperative experimentals, however, were switched to the free fee-for-service plan once they moved from the Seattle area.

Sample loss should not appreciably affect our estimates of the differences in mental health status outcomes among plan groups. First, rates of sample loss were not significantly different by plan. Although the numbers in Table 4 indicate greater sample loss for the GHC participants than for the fee-for-service plan participants, this greater loss is entirely an artifact of the greater proportion of GHC experimental participants enrolled for five years (that is, they had more time to drop out). Using a Weibull survival model to adjust for time at risk, we found no significant differences in loss rates among the GHC and fee-for-service plans ($\chi^2(3) = 0.30$); the free fee-for-service versus GHC comparison was also insignificant ($t = 0.50$) (see Manning, Wells, and Benjamin, 1986a for further details). In their analysis of medical service use, Manning, Leibowitz, Goldberg, et al. (1984) found no evidence of significant bias from sample loss.

Second, no significant differences in initial mental health status existed among the dropouts from the various plans. However, those individuals who left the study or the Seattle metropolitan area early had lower mental health status than those who remained. The difference is 2.3 units of MHI, $p < 0.01$, which corresponds roughly to the effect of being laid off or fired. But when we considered the
enrollment mental health status of those who stayed to normal completion, no significant differences by plan existed; chi-square(2) = 0.69, 0.74, and 0.56 for overall mental health status, psychological well-being, and psychological distress, respectively, which are all insignificant (p > 0.50).

Third, we observed no differential response to plan by the dropouts during the period they remained in the study (p > 0.50).
III. RESULTS

The markedly less inpatient-care intensive and outpatient mental health–care intensive—and consequently, less expensive—style of medicine practiced at the Group Health Cooperative, a staff model HMO, produced no significant differences in mental health outcomes relative to fee-for-service plans. Table 5 contrasts mean annual change—actual and predicted—for participants in the GHC or the family pay plans.

Table 5
MEAN DIFFERENCES AMONG INSURANCE PLANS IN CHANGE IN MENTAL HEALTH STATUS

<table>
<thead>
<tr>
<th>Plan Comparison</th>
<th>MHI Difference</th>
<th>MHI t</th>
<th>PWB Difference</th>
<th>PWB t</th>
<th>PSDS Difference</th>
<th>PSDS t</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unadjusted means</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pay – free/IDP</td>
<td>-0.063</td>
<td>-0.17</td>
<td>0.183</td>
<td>0.38</td>
<td>0.137</td>
<td>0.40</td>
</tr>
<tr>
<td>Pay – GHC</td>
<td>0.037</td>
<td>0.11</td>
<td>-0.076</td>
<td>-0.18</td>
<td>0.027</td>
<td>0.09</td>
</tr>
<tr>
<td>Free/IDP – GHC</td>
<td>0.100</td>
<td>0.33</td>
<td>-0.258</td>
<td>-0.66</td>
<td>-0.111</td>
<td>-0.39</td>
</tr>
<tr>
<td>Predicted meansa</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pay – free/IDP</td>
<td>-0.065</td>
<td>-0.19</td>
<td>-0.004</td>
<td>-0.01</td>
<td>0.066</td>
<td>0.20</td>
</tr>
<tr>
<td>Pay – GHC</td>
<td>0.048</td>
<td>0.15</td>
<td>0.104</td>
<td>0.26</td>
<td>-0.088</td>
<td>-0.29</td>
</tr>
<tr>
<td>Free/IDP – GHC</td>
<td>0.113</td>
<td>0.40</td>
<td>0.099</td>
<td>0.27</td>
<td>-0.153</td>
<td>-0.57</td>
</tr>
</tbody>
</table>

NOTE: MHI = mental health index; PWB = psychological well-being; PSDS = psychological distress; FREE = 0 percent cost sharing, fee-for-service plan; GHC = prepaid group practice (Group Health Cooperative); PAY = family coinsurance of 25, 50, or 95 percent; IDP = individual deductible plan with 95 percent outpatient cost sharing. A positive value indicates that the free/IDP plan is worse for the MHI and PWB; a negative value indicates that the free/IDP plan is worse for the PSDS when free/IDP serves as the contrast group (first and fourth rows). A positive value indicates that the GHC plan is worse for the MHI and PWB; a negative value indicates that the GHC plan is worse for the PSDS when the GHC is the contrast group (second, third, fifth, and sixth rows). Means were estimated using generalized least squares.

aThe predictions are from a model that controls for initial health status and other covariates.
relative to participants in the free/individual deductible plans. We found no statistically significant differences for these insurance plan contrasts.

The predicted plan differences were quite small. For example, the annual difference in outcome for participants in the F/ID plan group and the GHC group was only 0.113 units on the MHI.

With these data and our approach, we had the precision to detect (at the 5 percent level) a difference in MHI of from −0.68 to +0.45 units per year between the combined F/ID plan group and the GHC group. Thus, we had the precision to detect an effect equivalent to one-quarter the impact of being fired or laid off on overall mental health status. The predicted differences we observed, however, were on the order of 5 percent of the effect on mental health status of being laid off or fired.

Our conclusions are the same using alternative methods of analysis. We obtained the same pattern of results using ANOVA to contrast actual differences and multiple regression methods to adjust for individual characteristics (see Table 5). Further, no significant interactions among insurance plan, initial mental health status, and income in predicting mental health outcomes existed.
IV. DISCUSSION

Interest in the HMO concept remains high after almost 18 years of federal government promotion. With the recent demonstration of its cost-saving ability, it serves as the cornerstone for federal and state cost-control initiatives (Wells, Manning, and Benjamin, 1986a). Given the importance of this concept in cost-control efforts, understanding the full implications of financing and organizing health care in this manner is important. Our results from a randomized controlled trial in one well-established staff model HMO permit us to answer two questions. First, does the main conclusion of Ware et al. (1986) and Valdez et al. (1989)—that no significant difference in mental health outcomes exists between HMO plan participants and fee-for-service plan participants—still hold true using more years of information and more efficient statistical methods? Second, do differences in outcome exist along more specific mental health dimensions—psychological well-being and psychological distress—between HMO plan participants and fee-for-service plan participants?

We found no statistically significant or appreciable differences in mental health outcomes between participants enrolled in HMO plans and those enrolled in fee-for-service plans. We reached the same conclusion using each of our mental health status measures. We had enough precision to detect a clinically meaningful effect—an annual difference in outcome equivalent to one-quarter the impact of being fired or laid off. We found no evidence of the HMO’s having any detrimental or positive effects, relative to comparable fee-for-service plans, on mental health outcomes for special subgroups of interest, such as those with lower initial mental health ratings and/or lower incomes. Our precision for testing these interactions, however, was small.

We examined the direct effects of HMO versus fee-for-service participation on the mental health status of the total population, rather than first estimating the effects of participation on use of services and then the effects of variations in use on outcomes. We did this for two reasons. First, the total impact of the plan differences is the sum of the effects on mental health status for those who do and do not receive care. Both the untreated and the treated may experience changes in mental health status over time. In part, these changes may reflect effects of the different insurance plans because they affect the proportion and case mix of the covered population that remains untreated. Second, in the HIE,
participants were randomly assigned to the HMO or fee-for-service plans, not to treatment. Thus, for purposes of estimating effects of differences in use of services on outcomes, the HIE has a weak (observational) design, while for purposes of estimating the direct effects of the HMO versus fee-for-service participation, it has a strong (experimental) design.

Our main conclusion that care in an HMO leads to no differences in mental health compared to care under fee-for-service coverage is particularly noteworthy because of large differences in service use. The probability of inpatient medical use is 40 percent lower for HMO participants than for fee-for-service participants (Manning, Leibowitz, Goldberg, et al., 1984). Although HMO participants received a much less intensive form of psychotherapy than did comparable fee-for-service participants, over a period of several years a larger proportion of HMO participants received some outpatient mental health treatment (Manning and Wells, 1986; Wells, Manning, and Benjamin, 1986a). These two stylistic differences in mental health treatment could have had a counterbalancing effect on mental health outcomes.

Two other reasons why the HMO and fee-for-service plans produced similar mental health status outcomes should be considered. First, few participants ever received formal mental health care specifically designed to improve their mental health. Fewer than 20 percent of free plan and 30 percent of GHC participants over three years used such services. Thus, we might not expect large effects, on average, in plan contrasts.

Second, the HIE's fee-for-service and HMO plans were more generous than many prevailing insurance plans. The HIE plans covered all inpatient medical and psychiatric care, outpatient medical care and outpatient psychotherapy up to 52 visits per person per year. The effects on mental health outcomes of plans currently available in the community may differ (that is, may result in more adverse affects) in the context of less generous coverage, particularly of mental health services. Some HMOs, for example, do not cover psychiatric inpatient services and cover only 10 outpatient visits per year.

Thus, an insurer offering a new HMO option could observe a different effect on mental health outcomes, especially if the HMO and fee-for-service plans differed in covered benefits. Individuals who were sicker or had a greater propensity to use services could select the option with more generous benefits, making outcomes on these plans appear relatively worse (if unadjusted for case mix). In previous work, however, we found no evidence for adverse selection effects at the Group Health Cooperative (Manning, Leibowitz, Goldberg, et al., 1984; Manning and Wells, 1986).

Our results do not necessarily apply to several populations of considerable public policy interest. The HIE excluded the Medicare disabled,
the elderly, and those institutionalized in long-term hospitals and jails. Further, our study examined the experience of a single, well-established staff model HMO. Thus, our results may not apply to populations receiving care from newer types of HMOs, such as independent practice associations.

These data provide relatively limited information on the severely mentally ill. The HIE did not include measures of specific psychiatric disorders, such as the Schedule for Affective Disorders and Schizophrenia or the Diagnostic Interview Schedule (Endicott and Spitzer, 1978; Robins et al., 1981). Thus, we cannot comment on the effects of insurance coverage either on the course of a psychiatric disorder or on the general mental health status of those with specific psychiatric disorders.

In sum, we found no significant or appreciable differences in mental health outcomes for a general population assigned to receive care from an HMO compared to those covered by fee-for-service insurance plans. Nor did we observe differences for people with initially low mental health and/or low income. The absence of differences in mental health were achieved despite the fact that the HMO plan provided less-intensive outpatient therapies to a larger share of its enrolled population than did fee-for-service plans.
Appendix

STATISTICAL METHODS

In this analysis, we have used both analysis of variance (ANOVA) and least-squares regression methods to estimate the effects of insurance coverage on changes in mental health status. Specifically, we regress the difference between that year's value and the entry value of mental health status on insurance coverage (that is, our comparisons are of the before-and-after variety). All differences are stated in terms of annual changes, whether they occurred in the first or last year of the period we examined. We adjusted for the amount of time elapsed between that year's measurements and the entry measurements; each difference is decided by the amount of time from the study's beginning to the time of assessment.

HETEROSCEDASTICITY

The data exhibit a heteroscedastic response (nonconstant variance) as a result of stating all differences in annual terms. Observations averaged over a five-year period are less variable than those averaged over a one-year period. In the parallel study of fee-for-service plans (Wells, Manning, and Valdez, 1969), we could not reject the hypothesis that the variance was inversely proportional to time at risk. We have weighted all observations to reflect time at risk.

CORRELATED RESPONSES

Our data exhibit positive correlations between individuals in the same family and in the same individual over time. Using differences (response at time t minus response at time 0) removes some, but not all, of the correlations. In fact, differencing adds some correlation. Each year's response includes a permanent and transitory component. Differencing removes the permanent component but adds the (negative of the) entry transitory term to each subsequent error for differences. As a result, we observe a positive correlation among differences.
We model this correlation using a nested variance component or intraclass cluster model (Maddala, 1971; Searle, 1971). Specifically, the model is

\[ \Delta MH \text{ Status}_{it} = x'_{it} \beta + \mu_f + \nu_i + \epsilon_{it}, \]  

(1)

where \( \Delta MH \) Status and \( x \) are the dependent and independent variables (transformed to remove the annualizing heteroscedasticity) for person \( i \) in family \( f \) at time \( t \), \( \beta \) is a coefficient vector to be estimated, \( \mu \) is an unobserved family effect (distributed across families with mean 0 and variance \( \sigma_\mu^2 \)), \( \nu \) is an unobserved individual effect variance (distributed across individuals with mean 0 and variance \( \sigma_\nu^2 \)), and \( \epsilon \) is an error term (distributed across observations with mean 0 and variance \( \sigma_\epsilon^2 \)).

We use a two-step random effects estimator. We estimate the ordinary least-squares (OLS) version of Eq. (1). The variances of \( \mu \), \( \nu \), and \( \epsilon \) are estimated from the OLS residuals, correcting for the unequal number of observations per person and per family (Searle, 1971). With these estimated values, we then obtain the generalized least-squares estimate of \( \beta \)

\[ \beta_{GLS} = (X'\Omega^{-1}X)^{-1} X'\Omega Y, \]

where \( \Omega \) is block diagonal with a block \( D_F \) for each family \( f \); \( D_I \) is block diagonal with blocks \( P_i \) for each person:

\[
\Omega = \begin{pmatrix}
D_1 & 0 & & \\
0 & D_2 & & \\
& & \ddots & \\
& & & D_M
\end{pmatrix}
\]

\[
D_F = \begin{pmatrix}
P_1 & & \\
& \rho_F & \\
& & \rho_F & \\
& & & \rho_F
\end{pmatrix}
\]

\[
P_1 & & \\
\rho_F & \\
\rho_F & \\
\rho_F & \\
\rho_F & \\
P_T
\]
with intrafamily correlation:

$$\rho_F = \frac{\sigma^2_F}{\sigma^2_F + \sigma^2 + \sigma^2}$$

and intraperson correlation:

$$\rho_I = \frac{\sigma^2_I}{\sigma^2_F + \sigma^2 + \sigma^2}$$

and $DF$ has the same dimension as the number of person years in family $F$, and $P_I$ has the same dimension as the number of years for person $I$.

We use a random effects estimator for two reasons. First, by the study's design, the plan is orthogonal to $\mu_F$ and $\nu_I$. Second, we could not estimate a fixed effects model because all members of the same family have the same insurance. That is, a fixed effects model would not be full rank.

We should note that the error term does not exhibit constant correlation over time. Some evidence of mild autoregression exists. Given the short time frame and the weak autoregression, our approach is a close enough approximation that avoids the large cost of mixed variance components/autoregressive moving average models, with unequal number of observations cross-sectionally and over time.

Here, we use a more efficient estimation technique than in earlier RAND work. The earlier studies reported OLS parameter estimates and a corrected set of inference statistics. Their correction for heteroscedasticity and correlation was nonparametric, based on Huber's (1967) approach. In contrast, we used generalized least-squares methods, which yield more efficient estimates if the parametric specification of the error above (Eq. 1) is correct (Maddala, 1971; Maddala, 1977; Judge et al., 1980, 1982).
REFERENCES


Davies, A. R., and J. E. Ware, Jr., _Measuring Health Perceptions in the Health Insurance Experiment_, The RAND Corporation, R-2711-HHS, October 1981.


