The Contribution of Primary Care Systems to Health Outcomes within Organization for Economic Cooperation and Development (OECD) Countries, 1970–1998

James Macinko, Barbara Starfield, and Leiyu Shi

Objective. To assess the contribution of primary care systems to a variety of health outcomes in 18 wealthy Organization for Economic Cooperation and Development (OECD) countries over three decades.
Data Sources/Study Setting. Data were primarily derived from OECD *Health Data 2001* and from published literature. The unit of analysis is each of 18 wealthy OECD countries from 1970 to 1998 (total $n = 504$).
Study Design. Pooled, cross-sectional, time-series analysis of secondary data using fixed effects regression.
Data Collection/Extraction Methods. Secondary analysis of public-use datasets. Primary care system characteristics were assessed using a common set of indicators derived from secondary datasets, published literature, technical documents, and consultation with in-country experts.
Principal Findings. The strength of a country's primary care system was negatively associated with (a) all-cause mortality, (b) all-cause premature mortality, and (c) cause-specific premature mortality from asthma and bronchitis, emphysema and pneumonia, cardiovascular disease, and heart disease ($p<0.05$ in fixed effects, multivariate regression analyses). This relationship was significant, albeit reduced in magnitude, even while controlling for macro-level (GDP per capita, total physicians per one thousand population, percent of elderly) and micro-level (average number of ambulatory care visits, per capita income, alcohol and tobacco consumption) determinants of population health.
Conclusions. (1) Strong primary care system and practice characteristics such as geographic regulation, longitudinality, coordination, and community orientation were associated with improved population health. (2) Despite health reform efforts, few OECD countries have improved essential features of their primary care systems as assessed by the scale used here. (3) The proposed scale can also be used to monitor health reform efforts intended to improve primary care.

Key Words. Primary care, health system assessment, health reform

The World Health Organization (WHO) has recently proposed new methods for assessing the performance of national health systems (World Health Organization 2000). In spite of criticism of the methodologies employed (see Almeida et al. 2001; Blendon, Minah, and Benson 2001), the WHO report has nevertheless reinvigorated research into the international comparison of health systems worldwide.

This renewed interest is justified because although the public health literature indicates some question as to the overall contribution of medical care to the improvement of population health worldwide (McKeown 1976; McKeown, Record, and Turner 1975; McKinlay and McKinlay 1977), there is evidence that access to certain types of care may be more beneficial than others in reducing a country's overall burden of disease (Starfield 1996). Some authors have suggested that within European countries at least some of the historical differences in health status between the rich and the poor may be in part due to differential access to basic health services (Mackenbach, Stronks, and Kunst 1989).

The evidence is more striking when the unit of analysis shifts from generic systemwide assessments to more detailed analyses of specific health system components. In particular, primary care, defined as "that level of a health service system that provides entry into the system…provides person-focused care over time, provides care for all but very uncommon or unusual conditions, and coordinates or integrates care provided elsewhere or by others" (Starfield 1998), has been shown to exert a positive influence on health costs, appropriateness of care, and outcomes for some of the most common medical problems (Bindman et al. 1996; Engel et al. 1989; Kohn and White 1976; Moore 1992; Roos 1979). There is some evidence that countries characterized by a strong primary care orientation have more equitable health outcomes than those systems oriented toward specialty care, because primary care is thought to be both less costly to individuals and more cost-effective to

---

This study was partially funded by a grant (T32 HS 00029) from the Agency for Healthcare Research and Quality.

Address correspondence to James Macinko, Ph.D., 624 N. Broadway, Room 452, Baltimore, MD 21205. James Macinko is Visiting Professor, Department of Health Administration and Planning, National School of Public Health/Fundação Oswaldo Cruz, Rio de Janeiro, Brazil. Barbara Starfield, M.D., M.P.H., is University Distinguished Professor, Department of Health Policy and Management Johns Hopkins Bloomberg School of Public Health, Baltimore, MD. Leiyu Shi, M.B.A., Dr.P.H., is Associate Professor, Department of Health Policy and Management, Johns Hopkins Bloomberg School of Public Health.

society—thus freeing up resources to attend to the health needs of the most disadvantaged (Shi 1994; Starfield and Shi 2002; Starfield 1998).

Examination of the strength of primary care systems is particularly important as during the past 20 years most industrialized countries have undergone some kind of health reform. These reforms have primarily had the aim of containing the growth of health care costs, but in recent years have increasingly incorporated other objectives such as improving the quality of care and enhancing equity (Organization for Economic Cooperation and Development 1995a). Primary care should be central to achieving all three of these objectives, although few studies have examined multiple countries' primary care systems over time. This study aims to assess the contribution of national primary care systems to health outcomes (while controlling for other known determinants of health) during the period 1970–1998.

CONCEPTUAL FRAMEWORK

The conceptual framework for this study is operationalized in the following way. First, the most distal determinants of health in the model include national policies and culture. National policies refer to policies outside the health sector that affect overall macroeconomics, international relations, and income redistribution, for example. Culture refers to the broad set of beliefs and practices specific to one's national, subnational, religious, and/or ethnic identity that might contribute to different preferences for types of political and legal institutions, social participation, institutional development, lifestyle choices, and overall priorities. Both national policies and culture are conceptualized as antecedents to more proximal determinants of health, but are not explicitly analyzed in the study.

More proximal health determinants include factors thought to influence health at both the macro and the micro levels. Of these, the most macro-level health determinants include the environment and demographics. Environment refers to the overall availability and quality of natural resources affected by national policies. Demographics are considered a macro-level health determinant because a greater number of deaths can be expected in countries with a higher percentage of elderly or children. Another macro-level factor exerting an important influence on health status is the national socio-economic environment, the most common measure of which is Gross Domestic Product (GDP) per capita. This measure has consistently been found to be positively associated with better health outcomes (World Bank 1993).

In this model, the health system serves as an interface between individuals and the larger political system in which they live. Health care is produced by system inputs (physicians, medicines, facilities) that interact with the population through various processes (medical consultations, surgeries, deliveries) and result in health outcomes (Donabedian 1973).

Individual resources can be social, economic, or biological. Social resources include social networks and support considered to be important determinants of health (Berkman and Syme 1979). Economic resources include income and working conditions. Both social and economic conditions are often grouped together as socioeconomic status (SES). Socioeconomic status is defined as "a composite measure that typically incorporates economic status, measured by income; social status, measured by education; and work status, measured by occupation" (Adler 1994). Other individual resources include one's genetic makeup, although there are no international measures available to compare health risks or resiliencies conferred by heredity.

Behavioral factors include lifestyle choices that make up an increasingly significant percentage of health risks, particularly in industrialized countries (Murray and Lopez 1997). Two of the most important factors are drinking and smoking, both of which have been linked with premature mortality (Dawson 2000; Enstrom and Heath 1999; World Health Organization 1997). Other individual behaviors relevant to this framework include participation in politics or volunteering, both of which have been linked to improved health status, at least at the aggregate level (Blakely, Kennedy, and Kawachi 2001; Kawachi et al. 1997).

Although primary care is an integral part of the overall health system, it is treated separately in this framework because of evidence of its unique contribution to population health. The organization and delivery of primary care is measured through a scale adapted from Starfield (1994; 1998). Primary care is thought to mediate the effects of other health determinants. For example, we hypothesize that a strong primary care system will improve preventive care, reducing the national burden of preventable deaths. Access to good primary care is thought to reduce at least some of the ill health effects of social inequalities associated with income and resource distribution (Shi and Starfield 2000; Shi et al. 1999). Finally, good primary care is expected to be associated with improved functioning of the health system at large since strong primary care not only means more prevention, but also better referral, coordination, and continuity of care (Casanova and Starfield 1995). Thus, all else being equal, a country with a strong primary care system should have better health indicators (particularly those most sensitive to primary care) than a country with a weaker primary care system.

METHODS

The study uses a pooled, cross-sectional, time series analysis of secondary data on 18 of the Organization for Economic Cooperation and Development (OECD) member countries during the period 1970 to 1998. These countries and years have been chosen based on availability of data and similarities among countries. The unit of analysis is each country at each year (country-year).

Data

Data for this study come primarily from OECD *Health Data 2001* (Organization for Economic Cooperation and Development 2001). For reasons of missing or noncomparable data, not all of the 30 OECD member countries were included in regression analyses. The World Bank has categorized Czech Republic, Hungary, Korea, Mexico, Poland, Slovakia, and Turkey as "middle income" countries (World Bank 1993). Since there is some evidence that health determinants may vary by level of socioeconomic development (Omran 1971), these countries were excluded from the regression analyses. Austria, Iceland, Ireland, Luxembourg, and New Zealand were not included in the analyses because of extensive missing values. All countries included in the study are listed in Appendix A.

It is appropriate to use an ecological comparison across countries in this study because country-to-country differences are a well-established method of assuring sufficient variability in health policy and health system organization, financing, and delivery (Ellencweig 1992).

This study uses ecological-level data only. One potential risk in using ecological measures is the "ecological fallacy." This fallacy occurs when factors that are associated with national disease rates are assumed to be likewise associated with disease in individuals. Because this study only uses the country itself as the unit of analysis and does not make generalizations about individuals or specific population groups within each country, it does not risk an ecological fallacy (Schwartz 1994).

A limitation of many cross-country comparisons is sample size. We calculated the study's power using the following formula: $m = 2(Z_\alpha + Z_Q)^2 [1+(n-1)]\rho/(n\Delta^2)$, where: m = minimum sample size; Z_α = Z-score for a type 1 error of 0.05; Z_Q = Z-score for a type II error rate of 0.8; n = number of repeated observations (28 for this study); ρ = correlation among the repeated observations; and Δ = the smallest meaningful difference to be detected in standard deviation units (Diggle, Liang, and Zeger 1994). Given the study's sample size (18 countries $\times$ 28 time periods = a maximum of 504 data points), this design should be able to detect differences in standard deviations of less than 20 percent, even at high levels (0.8) of correlation.

Missing Data

Missing data were adjusted for in several ways. First, only independent variables that appeared to be missing at random were imputed. Second, independent variables were imputed only if they were missing less than 25 percent of their total values, otherwise, the variables were not analyzed. Based on these two criteria, less than 10 percent of all data was imputed. Because the statistical analysis software uses listwise deletion (that is, the entire unit of analysis will be dropped from the regression if any one data point is missing), failure to use even this relatively modest amount of imputed data could result in further biases due to artificial reductions in sample size.

Two techniques were used for imputing missing values. For those values that tended to increase over time, grand mean or regression imputation could introduce values that were too high if the missing data occurred early in the time series, or too low if the missing data appeared late in the time series. For this reason, missing values for number of physicians per one thousand population, for example, were imputed by using the midpoint between existing values in years immediately proceeding and following the missing value. For values that did not show any clear yearly pattern, but still increased over time, the within-country mean for each decade was used. Because of the lack of adequate predictor variables that were not included in the final model, regression imputation techniques were not used (Little and Rubin 1987). The vast majority of imputed data represents data missing for one or two years within a 28-year time series.

Prior to imputation, a dummy variable was created for each variable with missing values. Each of these dummy variables represents the pattern of missing values for imputed data and was included in preliminary regression analyses (see Little and Rubin 1987). In fixed effects multiple regression analyses, only one of the missing value dummy variables (that for tobacco) was found to be statistically significant. For this reason, it is included with the imputed tobacco variable in order to control for the pattern of missing values.

Multicollinearity results when independent variables in a multiple regression equation are highly correlated. This can result in a lack of precision in regression coefficient estimates (Tacq 1997). It is well known that in OECD countries, GDP per capita is highly correlated with national expenditures on health (Anderson et al. 2000; Barros 1998). In this study the correlation among these variables ranged from 0.84 to 0.96 and was statistically significant ($p < 0.05$) every year. Factor analysis did not satisfactorily reduce the number of economic variables. For this reason (and to avoid introducing potential endogeneity among GDP and expenditure data), these variables were not generally included in models where GDP per capita was also included. To correct for inflation and population differences, all economic data is reported on a per capita basis and in constant 1985 U.S. dollars. Income is further adjusted for purchasing power parities.

The GDP per capita was also correlated with the percentage of the population older than 65 years of age (r ranged from 0.23 to 0.62, $p < 0.05$ for every year prior to 1993). To retain this information in the model, the continuous variable was transformed into a binary variable reflecting whether a country at any given year had a proportion of elderly that exceeded the overall mean by at least one standard deviation. This variable thus reflects those countries that had a high proportion of elderly. The binary variable was not statistically significantly correlated with GDP for any year.

Health Outcomes

The effect of medical care generally and primary care specifically can be expected to be different depending on the etiology of different health outcomes. Therefore, this study uses a number of different health outcome variables to test the strength of primary care. Perhaps the most common outcome variables used in international comparisons are age and sex-standardized all-cause mortality and life expectancy at birth. The strength of primary care systems is tested in relation to each of these variables for both genders combined and for each gender separately.

Because primary care focuses on prevention and early identification of disease, we hypothesize that primary care would be an even more important predictor of premature mortality than for overall mortality rates. The OECD defines the Potential Years of Life Lost (PYLL) as a measure of premature mortality that weighs deaths occurring at younger ages more highly than those occurring at later ages. This is accomplished by adding up deaths occurring at each age and multiplying this sum by the number of years the average person would be expected to live until a certain age limit (age 70 in this case) (Organization for Economic Cooperation and Development 2001). Thus, the weight given to the death of an infant (70 years of life lost) is 14 times that of the death of a person aged 65 years (5 years of life lost) (Or 2000). To make them comparable across populations, PYLL measures are also age-standardized per 100,000 population, using the OECD population in 1980 as a reference.

The PYLL measures can also be disaggregated by gender and by specific disease categories. Starfield (1998) suggests that certain conditions such as premature deaths from asthma, diseases of the circulatory system, and infectious diseases may be particularly sensitive to primary care since their population prevalence and severity depend on prevention, early diagnosis, longitudinal (continuous) care, and coordination among different levels of care. For this reason, PYLL measures of premature deaths from bronchitis, asthma, and emphysema (ICD-9 490-496); cerebrovascular disease (ICD-9 430-438); ischemic heart disease (ICD-9 410-414); and pneumonia and influenza (ICD-9 480-487) are used as outcome variables in multivariate analyses.

Measures

Primary care is defined as that level of the health system that provides the majority of care to the population (World Organization of National Colleges, Academies and Academic Associations of General Practitioners/Family Physicians 1991). There are two main domains to be assessed when analyzing a country's primary care system: structural characteristics and practice features.

The most commonly identified structural characteristics of primary care systems include: (1) *health system finance*, or whether the health system is funded by taxes, social security, or private means; (2) *distribution of resources*, or the extent to which primary care resources are distributed according to need; (3) *physician inputs*, or the extent to which primary care providers are actually trained in primary versus specialty care; (4) *accessibility*, or the ability of patients to use services whenever needed; and (5) *longitudinality*, or the extent to which care is organized so as to provide a regular source of care over time (Elola, Daponte, and Navarro 1995; Lena and London 1993; Starfield 1996; Starfield and Shi 2002; Boerma and Fleming 1998).

Some comparative studies have used different measures for structural characteristics, such as physician salaries (Weiner 1987), or method of provider payment (capitation, salary, fee-for-service) (Or 2001). Other studies have measured access by including the proportion of the population covered by insurance (Elola, Daponte, and Navarro 1995). However, these measures may not be helpful in cross-country comparisons. First, the literature is unclear on the role of salaries on physician performance since there appears to be no "preferred" method of physician reimbursement (Delnoij et al. 2000; Scott and Hall 1995). Second, most OECD countries cover more than 90 percent of their population with health insurance, so there is little variation except in comparison with the United States. Third, they do not represent new structural characteristics, but can be considered less than desirable measures of the characteristics previously described under primary care provider and access.

There are five main features of primary care practice. *First contact* refers to the extent that primary care provides entry into the health system; *coordination* refers to the ability of primary care providers to coordinate use of other levels of health care; *comprehensive care* includes curative, preventive, and rehabilitative services; *longitudinality* refers to care that is patient-focused over time; and a *family and/or community orientation* places the patient within the wider familial and social context necessary for addressing multiple causes of illness or health (Boerma and Fleming 1998; Starfield 1998). Each of these features is addressed in the primary care scale used in this study.

Table 1 provides a description of the 10 components of the primary care scale developed for this study. This scale borrows from Starfield (1994; 1998; 2002), although it differs in three important ways. First, it includes only one measure for each of the structural characteristics and practice features. Measures of structural characteristics excluded from the scale include: "percent active physicians who are specialists," "professional earnings of primary care physicians relative to specialists," "requirements for 24-hour coverage," and "strength of academic departments of family medicine." These measures were not included because they were considered duplicative and because there is less empirical evidence on their contribution to population health than other measures used. Another variable, "patient lists," was dropped because it contains the same data present within the variable "longitudinality."

Second, instead of creating separate "structure" and "practice" subscales, this analysis sums together all 10 variables. This was done for both conceptual and empirical reasons. The division between structural characteristics and practice features is not always conceptually clear. For example, longitudinality as measured by patient lists is both a structural characteristic that determines from where primary care practices draw their clients, and a practice feature that ensures that clients see the same provider while residing

Table 1: Primary Care System Score Components

Component	*Indicator*	*Rationale*	*Scoring*
Regulation	Do specific national policies exist that regulate the distribution of primary care providers and facilities?	These policies are intended to improve equity in distribution of primary care services.	0 = no overall primary care regulation 1 = limited (only some regions or populations) 2 = entire system regulated
Financing	What is the method of financing health care for the majority of the population?	Scored by level of progressivity, tax-based systems considered most progressive.	0 = primarily private 1 = social security 2 = primarily tax-based
Primary care provider	What is the predominant type of primary care provider?	Generalists (general practitioners, family doctors) considered best providers of primary care.	0 = majority are specialists 1 = majority are pediatricians, internists 2 = majority are generalists
Access	What is the level of cost-sharing for primary care visits?	High primary care copays areaccess. considered to be a barrier to	0 = high copay 1 = moderate 2 = none or very low
Longitudinality	Are individual patient lists required for all primary care units?	Patient lists considered optimal way to track patients over time.	0 = never required 1 = limited use (or group lists only) 2 = mandatory and ubiquitous
First contact	Is there a requirement that primary care practitioners serve as gatekeepers to other levels of care?	First contact is an essential if primary care is to attend to the majority of health problems.	0 = never required 1 = required but not enforced or required for limited population only 2 = always required

Comprehensiveness	Is a full range of primary care services and procedures available for all age groups?	Specific list of services includes: prevention, mental health, minor surgery, and routine obstetric care.	0 = not comprehensive (some services offered *only* in specialty care) 1 = somewhat (all offered but not in every primary care unit) 2 = comprehensive (all offered in most locations)
Coordination	Are guidelines for the transfer of information between primary care and other levels available and required?	Data transfer (either through electronic means or through client-held records) is essential for coordinating care between levels.	0 = no guidelines present 1 = guidelines present but not widely used 2 = guidelines present and required
Family-centered	Is there a requirement that client records be organized by family as opposed to by individual?	Indicator that primary care considers patient's family environment in diagnosis and treatment.	0 = never required 1 = required for only some regions or populations 2 = generally required
Community-oriented	Is there a policy that requires use of community-based data and/or presence of community members in primary care management or priority-setting?	Primary care is more effective when it treats patients in their larger social context.	0 = never required 1 = required for limited population only 2 = generally required

Source: Adapted from Starfield and Shi 2002; Starfield 1998.

in the same geographic area. Previous studies found that measures of the two scales were highly correlated ($r = 0.98$, $p < 0.01$) lending further support to the hypothesis that they measure the same concept (Starfield and Shi 2002). In terms of statistical analyses, this also implies that the two scales should not be included together in analytical models without introducing multicollinearity.

A composite scale was used in order to measure the total impact of primary care on population health. This was done because the intent of this paper is to measure the impact of the primary care system as a whole on population health. Further justification for creating a single primary care measure includes the fact that individual elements of primary care structure and practice may not be independent. For example, most primary care systems that use patient lists will also employ primary care providers as gatekeepers. Thus, it would be difficult to parse out the separate effects of these two elements within the same model. Further research is necessary to determine mechanisms for weighing the relative contribution of the individual components of the scale.

Finally, the scale used in this study differs from that of Starfield in that scores for each of the countries included in this study are calculated not for one but for 28 years.

Each primary care component was assigned a score based on an objectively verifiable indicator. These indicators and their rationale are described in Table 1. Data used for determining the value of each indicator for each country and year were obtained through published literature and public use datasets (European Observatory on Health Care Systems 1996a; 1996b; 1996c; 1999a; 1999b; 1999c; 2000a; 2000b; 2000c; 2000d; 2000e; 2001a; 2001b; Organization for Economic Cooperation and Development 2001). Boerma and Fleming's (1998) survey of European primary care practitioners provided additional data on primary care physicians and practice characteristics.

Each of the 10 components for each country at each time period was scored using the criteria presented in Table 1. For each country and year, each of the 10 components was scored from 0 to 2, with 0 representing the absence of the characteristic and 2 representing the presence of the characteristic. A score between 0 and 2 represented either poor implementation or that only a percentage of the population was affected by that primary care component. In the case of health care financing and primary care provider, a score of 1 referred to social security and pediatrician/internist, respectively—both considered improvements over scores of 0, but less conducive to good primary care than responses earning a score of 2.

Changes in scores over time represent changes in national policy. In general, the change in score corresponds with the change in the policy, not necessarily in its implementation. The primary care score for each country during each year was calculated by using the sum of all 10 components for that year. The overall primary care score has a range of 0 (worst) to 20 (best).

Several procedures were followed to ascertain the validity and accuracy of the scoring system and the scale. First, country scores were shared with a selected group of international primary care experts in order to clarify scores and confirm accuracy of country rankings. However, no formal group consensus procedure was used. Second, Cronbach's alpha was calculated. The scale's overall alpha score was 0.91. Unstandardized scores (0.91) did not significantly vary from standardized ones (0.88) and all alpha scores were well within the acceptable range for the scale to be considered valid (DeVellis 1991). Third, factor (principal factor) analyses were performed on the 10 items included in the primary care score. The 10 items loaded on one major factor (Eigenvalue 5.32), which explained 75 percent of item variance. This also supports the authors' decision to use a composite scale, as the individual elements did not naturally break down into statistically distinct primary care subcategories. Finally, national primary care rankings were compared with published studies that ranked national primary care systems using a slightly different combination of variables (Starfield and Shi 2002; Starfield 1998). Country ranks calculated in this study are broadly consistent with those obtained in these prior studies. Due to the different components of the two scoring systems, however, these two scales cannot be directly compared.

As a final sensitivity test, all regression analyses using the composite score were also run using a primary care dummy variable that represented whether a country's primary care score was above or below the yearly mean primary care score for that year. The results of this analysis were generally consistent with those obtained in multivariate statistical models using the composite (0–20) primary care score.

Score reliability was not calculated. Further applications of the primary care scale presented here will be necessary before this aspect of the scale's reliability can be tested empirically.

Analyses

This study uses a pooled, cross-sectional, time series design to assess the relationships between the dependent and independent variables over a 28-year period. A pooled, cross-sectional, time series design is one in which variables for a number of different cross-sections are observed over a time span (Sayrs 1989).

This analysis uses fixed effects (FE) regression because ordinary least squares (OLS) regression will not yield proper estimates on data containing repeat measures, and the alternative (the random effects model) was found to be inappropriate for use with this data due to results obtained from performing a Hausman test (Hsiao 1986).

The fixed effects model uses a differencing estimator in order to remove the systematic variance in the error term. The OLS regression can then be used on the transformed model (Hsiao 1986). This estimation technique yields results identical to the least squared dummy variable (LSDV) approach described by Sayrs (1989).

Another advantage of the fixed effects model is that it controls for time-invariant heterogeneity among countries. Examples of factors potentially captured in the fixed effects include: underlying aspects of national culture, historical patterns that shape social institutions and policy systems, value systems that influence citizens' outlook on life and perceptions of illness, and factors shaping health-seeking behaviors.

The F-test was used throughout this study to test nested models. As per Tacq (1997), R^2 values used to calculate F-tests are adjusted for the number of covariates in the model in order to calculate the most conservative estimate possible. Although some models may have different effective sample sizes, list-wise deletion was used to assure that all F-tests were performed on models containing exactly the same data. All data were analyzed using the *Stata* software package, version 7 (Statacorp 2001).

RESULTS

Table 2 presents the total primary care scores for 18 members of the OECD for three decades: the 1970s, the 1980s, and the 1990s. Countries were scored based on each of the 10 criteria listed in Table 1. Scores reported reflect those at the midpoint of each decade. Statistical analyses were performed using scores calculated every year from 1970 to 1998. Appendix A contains the score for each component used to calculate the total primary care score for each country.

Several trends are apparent from an examination of Table 2. First, the average primary care score has improved by nearly one point over the three-decade period, although this difference was not statistically significant.

Table 2: Primary Care System and Practice Scores for OECD Countries

1975		*1985*		*1995*	
Country	*Score*	*Country*	*Score*	*Country*	*Score*
Countries Scoring above Mean					
Denmark	18	Denmark	18	U.K.	19
U.K.	17	U.K.	17	Denmark	18
Netherlands	14	Netherlands	15	Spain	16.5
Norway	13	Italy	13.5	Netherlands	15
Australia	12	Australia	13	Italy	14
Spain	11	Norway	13	Finland	14
Italy	10.5	Spain	11	Norway	13
Finland	10	Finland	10.5	Australia	13
Sweden	9.5	Canada	10.5	Canada	11.5
Canada	8	Sweden	9.5	Sweden	11
Countries Scoring below Mean					
Japan	7.5	Japan	7.5	Japan	7.5
Portugal	6	Portugal	7	Portugal	7
Greece	4	Germany	4	Belgium	4
Belgium	4	Belgium	4	Greece	4
Germany	4	Greece	4	U.S.A.	3
Switzerland	2.5	Switzerland	2.5	Germany	3
France	2	France	2	Switzerland	2.5
U.S.A.	1	U.S.A.	1	France	2
Summary Statistics by Decade					
Observations	18		18		18
Mean	8.85		9.27		9.65
Std. Dev.	5.01		5.26		5.51

Data Sources: Starfield 1998; Starfield and Shi 2002; European Health Observatory "Health Systems in Transition" publication series 1996–2001; OECD 2000, 2001; personal communications.

Second, countries that were high performers in the 1970s remained high performers in each subsequent decade. If countries are divided into high and low performers (above or below the mean for each decade), then no country crossed the threshold from low to high or from high to low.

Certain system characteristics also appear to be associated with high primary care scores. For example, those countries with tax-based health financing also tend to score higher on other primary care components. One exception to this is the Netherlands. Although it does not have a purely tax-based health financing system, the Netherlands does share other features of primary care (such as few barriers to access, geographic regulation of primary care, use of family practitioners as gatekeepers, and a family-orientation) with its Scandinavian neighbors, all of which have high primary care scores.

Within the two groups of high and low performers there were significant movements over time. In general, these changes reflected improvements in primary care. For example, in the late 1980s and early 1990s Spain experienced reorganization and strengthening of primary care by improving health system features (moving to a tax-based financing system, improving geographic allocation of funds, and increasing the supply of family physicians) as well as practice features (improved integration, family orientation, coordination, and health promotion) (Larizgoitia and Starfield 1997). Although the reform has not achieved all of its aims in every region (Larizgoitia and Starfield 1997), there is some evidence that health outcomes have improved in regions where reform has been fully implemented (Villalbi et al. 1999). The United States also showed a slight improvement over time. This improvement comes almost entirely from increased participation of Americans in health maintenance organizations (HMOs), which have tended, on average, to use a higher percentage of primary care providers who act as gatekeepers to higher levels of care, and which had (at least among the not-for-profit HMOs) a tradition of community involvement. This trend may reverse itself with the decrease in not-for-profit HMOs in the United States that was seen in the late 1990s.

Only one country's score decreased over time; Germany experienced decreased access due to increased out-of-pocket payments, thus lowering its overall primary care score (Organization for Economic Cooperation and Development 2001).

Table 3 presents descriptive data on independent and dependent variables used in multivariate analyses. Physicians per one thousand population, doctor visits per capita, and the percentage of population older than age 65 all showed increases in mean values at each decade. Tobacco consumption showed a statistically significant decrease each decade, from an average of 2,661 grams per capita in the 1970s to 1,970 grams per capita in 1990s. Even while controlling for inflation, both GDP and income per capita showed statistically significant increases over time. From the 1970s to the 1990s, income per capita nearly tripled, and average GDP per capita nearly quadrupled.

Table 3 shows overall improvement in dependent (health outcome) variables over three decades. Age and sex-standardized all-cause mortality (as well as male- and female-specific all-cause mortality and cause-specific mortality) showed statistically significant declines each decade. All-cause

Table 3: Descriptive Statistics of Variables Used in the Study

		1970–1979		1980–1989		1990–1998	
Variable	*Definition*	*Mean*	*Std. Dev.*	*Mean*	*Std. Dev.*	*Mean*	*Std. Dev.*
Independent Variables							
Primary care	Total score	8.85	5.01	9.27	5.26	9.65	5.51
Physicians	Per 1,000 pop.	1.55	0.40	2.29*	0.65	2.82*	0.88
Alcohol	Liters/capita	11.82	4.19	11.16	3.19	9.91*	2.57
Tobacco	Grams/capita	2661.38	609.70	2417.79*	551.64	1970.31*	413.35
Income	$US/capita	589.55	533.42	923.41	858.23	1405.14**	1376.80
Doctor visits	Visits/capita	5.28	3.00	6.13	2.89	6.54**	3.11
GDP	$US/ capita	5926.78	3058.56	12480.40*	5390.30	23928.16*	6955.29
Elderly	% pop. over 65	11.72	2.26	13.17*	2.05	14.38*	1.64
Dependent Variables (All Standardized and Expressed as per 100,000)							
Mortality	All-cause (both genders)	943.29	112.28	791.98*	81.62	695.07*	79.57
Female mortality	All-cause (women)	743.02	99.22	611.03*	69.35	534.91*	68.67
Male mortality	All-cause (male)	1211.76	149.85	1038.48*	113.61	913.44*	102.26
Premature mortality	All-cause (both genders)	6933.48	1515.39	5355.63*	892.49	4494.81*	818.79
Premature mortality (cerebrovascular)	PYLL (both genders)	282.65	112.25	198.88*	77.23	146.75*	58.45
Premature mortality (heart)	PYLL (both genders)	662.91	317.44	498.26*	231.64	340.67*	128.21
Premature mortality (pneumonia)	PYLL (both genders)	225.00	264.16	81.14*	55.40	56.36*	29.40
Premature mortality (asthma)	PYLL (both genders)	88.98	43.05	59.52*	26.54	39.72*	15.60

* = Significantly different from previous decade average ($p < 0.05$)

** = Significantly different from 1970–1979 average only ($p < 0.05$)

mortality for both genders declined from an average of 943 per 100,000 in the 1970s to 695 per 100,000 in the 1990s, a statistically significant decrease of 26 percent. All-cause and cause specific premature mortality (Potential Years of Life Lost—PYLL) declined significantly over the period. Average all-cause PYLL declined from 6,933 years in 1970s to 4,495 years in the 1990s, a decline of approximately 35 percent.

Regression results are found in Tables 4–6. Fixed effects regression analyses are presented in separate tables for each health outcome measure. In each of the tables presented, three nested models are compared. The first model (model 1) contains only primary care. The second model (model 2) contains primary care in addition to macro-level factors (physician supply, GDP per capita, and whether or not a country had a high proportion of elderly). The third model (model 3) contains primary care, macro-level factors, and aggregated individual measures expressed on a per capita basis (number of doctor visits, liters of alcohol consumed, grams of tobacco smoked, and the log of income earned adjusted for purchasing power parities). An F-test was performed to test the hypothesis that additional variables improved each model.

In Table 4, fixed effects regression analyses are presented for all-cause standardized mortality for both genders, for women only, and for men only. Primary care is negatively associated with all-cause mortality rates independently (model 1), within a model of macro-level health determinants (model 2), and also in the full model (model 3) that includes aggregated individual determinants of health. In all models the primary care score is statistically significant ($p < 0.05$), although the effect of primary care is partially reduced in the presence of environmental factors, and further reduced by the presence of aggregate individual health determinants. As expected, the numbers of physicians and GDP per capita are also negatively associated with all-cause mortality. In the full model, doctor visits and alcohol are not statistically significantly related with mortality, although income per capita ($p < 0.001$) and tobacco ($p < 0.05$) are. Model 2 and model 3 have unadjusted R^2 values of 0.80 and 0.84, respectively.

The results are somewhat different for gender-specific all-cause mortality rates. For women, primary care is negatively associated with mortality in models 1 and 2, but it is not statistically significant in model 3. Macro-level factors have the expected effect, but alcohol, tobacco, and doctor visits per capita are not significant. All-cause mortality for men shows yet another pattern. Primary care is negatively associated with mortality in all three models, and both alcohol and tobacco are each positively associated with male

Table 4: Fixed Effects Regression Estimates for Mortality Outcomes

Variables	*Model 1*	*P*	*Model 2*	*P*	*Model 3*	*P*
All-Cause Standardized Mortality (Both Genders) per 100,000 Population (n = 467)						
Primary care score	– 66.950	0.000	– 14.496	0.000	– 7.520	0.013
Doctors/1,000 pop.			– 83.258	0.000	– 63.867	0.000
GDP/capita			– 0.007	0.000	– 0.004	0.000
Elderly			31.340	0.001	36.484	0.000
Doctor visits/capita					– 3.349	0.137
Alcohol (l/capita)					3.350	0.122
Tobacco (g/capita)					0.014	0.026
(Tobacco-missing)					35.664	0.000
Log income (ppp)					– 121.747	0.000
_Cons	1,439.970	0.000	1,221.818	0.000	1,758.118	0.000
F (df)	177 (1,448)	0.000	457(4,445)	0.000	253 (9,440)	0.000
R-2 (within)	0.283		0.804		0.838	
F-test (vs. model 1)	—		391 (3,445)	0.000	184 (8,440)	0.000
F-test (vs. model 2)	—		—	—	17 (5,440)	0.001
All-Cause Standardized Mortality (Women) per 100,000 (n = 467)						
Primary care score	– 54.076	0.000	– 7.680	0.005	– 2.680	0.325
Doctors/1,000 pop.			– 81.743	0.000	– 68.130	0.000
GDP/capita			– 0.006	0.000	– 0.003	0.000
Elderly			29.449	0.000	37.834	0.000
Doctor visits/capita					– 3.616	0.075
Alcohol (l/capita)					2.020	0.302
Tobacco (g/capita)					0.000	0.934
(Tobacco-missing)					30.090	0.000
Log income (ppp)					– 99.905	0.000
_Cons	1,138.449	0.000	953.610	0.000	1447.690	0.000
F (df)	149 (1,448)	0.000	423(4,445)	0.000	225 (9,440)	0.000
R-2 (within)	0.250		0.792		0.822	
F-test (vs. model 1)	—		383(3,445)	0.000	172(8,440)	0.000
F-test (vs. model 2)	—		—	—	13(5,440)	0.001
All-Cause Standardized Mortality (Men) per 100,000 (n = 467)						
Primary care score	– 83.642	0.000	– 24.065	0.000	– 15.014	0.000
Doctors/1,000 pop.			– 85.707	0.000	– 56.302	0.000
GDP/capita			– 0.009	0.000	– 0.004	0.000
Elderly			31.735	0.007	35.117	0.002
Doctor visits/capita					– 1.566	0.582
Alcohol (l/capita)					7.280	0.008
Tobacco (g/capita)					0.027	0.001
(Tobacco-missing)					45.848	0.000
Log income (ppp)					– 156.026	0.000

continued

Table 4: (Continued)

Variables	*Model 1*	*P*	*Model 2*	*P*	*Model 3*	*P*
_Cons	1,841.332	0.000	1,584.604	0.000	2,190.536	0.000
F (df)	197 (1,448)	0.000	390(4,445)	0.000	224 (9,440)	0.000
R-2 (within)	0.305		0.778		0.821	
F-test (vs. model 1)	—		313 (3,445)	0.000	154 (8,440)	0.000
F-test (vs. model 2)	—		—	—	19 (5,440)	0.001

mortality rates. Full models for female and male mortality both have unadjusted R^2 values of 0.82.

Table 5 presents results for all-cause Potential Years of Life Lost (PYLL)—a measure of premature mortality. Primary care is negatively associated with all-cause PYLL for both genders combined as well as for male-only and female-only PYLL, even in the presence of other known determinants of health (models 2 and 3). This is in keeping with the hypothesis that primary care would be most closely associated with health outcomes that represent preventable deaths. The number of physicians per one thousand population and income per capita were also found to be strongly and significantly negatively associated with PYLL in all models, and for both combined and separate gender-specific rates. In a pattern consistent with results for all-cause mortality, alcohol and tobacco use are positively associated with all-cause PYLL for men, but not for women. Finally, as expected, GDP per capita is negatively associated with PYLL, although in full models for both genders and for women only, it is not statistically significant. Unadjusted coefficients of determination for models 2 and 3 were between 0.69 and 0.79.

Table 6 presents results for four PYLL measures thought to be particularly sensitive to primary care. The first panel shows that primary care is negatively associated with premature mortality due to asthma, bronchitis, and emphysema, even in the presence of all other covariates, although, as in other models, the magnitude of the primary care coefficient was reduced with the introduction of other covariates. Total physician supply per one thousand population and income per capita were also negatively associated with premature deaths from these conditions. Somewhat surprisingly, GDP per capita was associated with a slightly increased prevalence of premature deaths. Coefficients of determination were only 0.43 and 0.49 for models 2 and 3, respectively. Because GDP per capita is positively associated with air

Table 5: Fixed Effects Regression Estimates for Premature Mortality (PYLL)

Variables	*Model 1*	*P*	*Model 2*	*P*	*Model 3*	*P*
All-Cause Potential Years of Life Lost (Both Genders) per 100,000 Population (n = 467)						
Primary care score	– 701.027	0.000	– 179.462	0.000	– 101.473	0.007
Doctors/1,000 pop			– 1,131.273	0.000	– 880.732	0.000
GDP/capita			– 0.045	0.000	– 0.010	0.221
Elderly			353.080	0.002	384.179	0.001
Doctor visits/capita					– 22.348	0.425
Alcohol (l/capita)					33.894	0.210
Tobacco (g/capita)					0.181	0.023
(Tobacco-missing)					157.961	0.081
Log income (ppp)					– 1,248.410	0.000
_Cons	12,177.950	0.000	10,311.230	0.000	15,573.000	0.000
F (df)	180 (1,448)	0.000	317(4,445)	0.000	162 (9,440)	0.000
R-2 (within)	0.286		0.740		0.768	
F-test (vs. model 1)	—		256 (3,445)	0.000	111 (8,440)	0.000
F-test (vs. model 2)	—		—	—	10 (5,440)	0.001
All-Cause Potential Years of Life Lost (Women Only) per 100,000 (n = 467)						
Primary care score	– 539.976	0.000	– 123.274	0.000	– 69.200	0.041
Doctors/1,000 pop.			– 1006.525	0.000	– 797.298	0.000
GDP/capita			– 0.027	0.000	– 0.004	0.569
Elderly			314.508	0.001	365.393	0.000
Doctor visits/capita					– 16.228	0.519
Alcohol (l/capita)					43.964	0.070
Tobacco (g/capita)					0.036	0.612
(Tobacco-missing)					72.262	0.374
Log income (ppp)					– 932.090	0.000
_Cons	9,008.761	0.000	7,620.663	0.000	11,619.500	0.000
F (df)	155 (1,448)	0.000	254(4,445)	0.000	124 (9,440)	0.000
R-2 (within)	0.258		0.695		0.717	
F-test (vs. model 1)	—		211 (3,445)	0.000	87 (8,440)	0.000
F-test (vs. model 2)	—		—	—	6 (5,440)	0.001
All-Cause Potential Years of Life Lost (Men Only) per 100,000 (n = 467)						
Primary care score	– 879.720	0.000	– 240.022	0.000	– 140.402	0.001
Doctors/1,000 pop.			– 1280.665	0.000	– 986.521	0.000
GDP/capita			– 0.064	0.000	– 0.016	0.080
Elderly			419.926	0.001	441.604	0.001
Doctor visits/capita					– 15.506	0.630
Alcohol (l/capita)					17.899	0.564
Tobacco (g/capita)					0.334	0.000
(Tobacco-missing)					256.455	0.014
Log income (ppp)					– 1,624.971	0.000

continued

Table 5: (Continued)

Variables	*Model 1*	*P*	*Model 2*	*P*	*Model 3*	*P*
_Cons	15,548.950	0.000	13,153.060	0.000	20,014.180	0.000
F (df)	193 (1,448)	0.000	354 (4,445)	0.000	190 (9,440)	0.000
R-2 (within)	0.301		0.761		0.795	
F-test (vs. model 1)	—		282 (3,445)	0.000	130 (8,440)	0.001
F-test (vs. model 2)	—		—	—	14 (5,440)	>0.05

pollution, it may be that the positive association between GDP and these respiratory conditions is actually due to increased air pollution.

Premature deaths from pneumonia and influenza are negatively associated with primary care in all three models. Physician supply and income per capita were also associated with reduced premature deaths from pneumonia. As expected, there were higher numbers of pneumonia deaths in countries with a higher proportion of elderly. Two results are unexpected. The GDP per capita was associated with slightly higher premature pneumonia deaths, and alcohol use was associated with lower pneumonia deaths. Coefficients of determination were low for these models (0.35 and 0.36), indicating that there are likely to be other important determinants of pneumonia and influenza deaths that were not included in the models.

Primary care is significantly associated with reduced premature deaths from cerebrovascular diseases. In the presence of other covariates the relationship is reduced in magnitude, but remains statistically significant. Physicians, GDP, doctor visits, and income per capita were all negatively associated with cerebrovascular PYLL measures. Those countries with a high proportion of elderly also experienced greater potential years of life lost due to cerebrovascular disease. Alcohol and tobacco were not found to have a significant effect. Unadjusted R^2 measures for the multivariate models were 0.68 and 0.72.

Ischemic heart disease is one of the most prevalent causes of death in OECD countries. It follows a pattern similar to that of cerebrovascular disease. Primary care is significantly and negatively associated with premature mortality from heart disease in all three models. Income and GDP per capita are also negatively associated with heart disease. Alcohol and tobacco use are positively associated with premature heart disease deaths. The unadjusted coefficient of determination was 0.64 for model 2 and 0.78 for model 3.

Table 6: Fixed Effects Regression Estimates for PYLL (Cause-Specific)

Variables	*Model 1*	*P*	*Model 2*	*P*	*Model 3*	*P*
PYLL–Asthma and Bronchitis (Both Genders) per 100,000 Population (n = 467)						
Primary care score	– 16.081	0.000	– 8.009	0.000	– 6.087	0.000
Doctors/1,000 pop.			– 19.707	0.000	– 11.857	0.000
GDP/capita			0.000	0.048	0.001	0.015
Elderly			2.502	0.535	8.547	0.037
Doctor visits/capita					0.174	0.865
Alcohol (l/capita)					– 0.497	0.613
Tobacco (g/capita)					– 0.003	0.337
(Tobacco-missing)					3.182	0.335
Log income (ppp)					– 51.878	0.000
_Cons	213.605	0.000	186.721	0.000	466.839	0.000
F (df)	149 (1,448)	0.000	84 (4,445)	0.000	47 (9,440)	0.000
R-2 (within)	0.249		0.429		0.489	
F-test (vs. model 1)	—		45 (3,445)	0.000	24 (8,440)	0.000
F-test (vs. model 2)	—		—	—	9 (5,440)	0.001
PYLL–Pneumonia and Influenza (Both Genders) per 100,000 Population (n = 467)						
Primary care score	– 62.338	0.000	– 23.403	0.002	– 18.709	0.017
Doctors/1,000 pop.			– 166.180	0.000	– 149.002	0.000
GDP/capita			0.004	0.002	0.006	0.000
Elderly			36.053	0.103	56.793	0.015
Doctor visits/capita					– 6.065	0.299
Alcohol (l/capita)					– 1.727	0.758
Tobacco (g/capita)					– 0.035	0.033
(Tobacco-missing)					– 10.384	0.581
Log income (ppp)					– 114.132	0.007
_Cons	702.491	0.000	644.488	0.000	1,383.982	0.000
F (df)	75 (1,448)	0.000	59(4,445)	0.000	28 (9,440)	0.000
R-2 (within)	0.144		0.346		0.364	
F-test (vs. model 1)	—		45 (3,445)	0.000	17 (8,440)	0.001
F-test (vs. model 2)	—		—	—	1 (5,440)	> 0.05
PYLL–Cerebrovascular Disease (Both Genders) per 100,000 Population (n = 467)						
Primary care score	– 39.304	0.000	– 11.833	0.000	– 8.340	0.001
Doctors/1,000 pop.			– 24.124	0.000	– 22.342	0.000
GDP/capita			– 0.006	0.000	– 0.003	0.000
Elderly			23.285	0.002	28.107	0.000
Doctor visits/capita					– 5.076	0.005
Alcohol (l/capita)					– 1.620	0.355
Tobacco (g/capita)					0.005	0.352
(Tobacco-missing)					37.779	0.000
Log income (ppp)					-64.779	0.000

continued

Table 6: (Continued)

Variables	*Model 1*	*P*	*Model 2*	*P*	*Model 3*	*P*
_Cons	578.836	0.000	445.944	0.000	806.266	0.000
F (df)	152 (1,448)	0.000	234(4,445)	0.000	129 (9,440)	0.000
R-2 (within)	0.254		0.678		0.725	
F-test (vs. model 1)	—		193 (3,445)	0.000	92 (8,440)	0.000
F-test (vs. model 2)	—		—	—	14 (5,440)	0.001
PYLL–Heart Disease (Both Genders) per 100,000 Population (n = 467)						
Primary care score	– 115.646	0.000	– 71.014	0.000	– 53.048	0.000
Doctors/1,000 pop.			8.393	0.532	22.830	0.074
GDP/capita			– 0.013	0.000	– 0.004	0.000
Elderly			5.562	0.761	– 33.285	0.032
Doctor visits/capita					– 5.668	0.144
Alcohol (l/capita)					– 16.136	0.000
Tobacco (g/capita)					0.158	0.000
(Tobacco-missing)					48.153	0.000
Log income (ppp)					– 153.697	0.000
_Cons	1,585.478	0.000	1,320.365	0.000	1,788.172	0.000
F (df)	287 (1,448)	0.000	201 (4,445)	0.000	171 (9,440)	0.000
R-2 (within)	0.391		0.643		0.778	
F-test (vs. model 1)	—		103 (3,445)	0.000	93 (8,440)	0.000
F-test (vs. model 2)	—		—	—	52 (5,440)	0.000

DISCUSSION

The results of this analysis suggest that primary care systems are important for population health. Strong national primary care systems were found to be negatively associated with aggregate and gender-specific mortality rates, overall levels of premature deaths, and premature deaths from a variety of important preventable or treatable conditions including deaths from asthma, heart and cerebrovascular diseases, and pneumonia. The effect of primary care was generally found to be significant, albeit reduced, in the presence of such important determinants of population health as demographics, income, and GDP per capita, and behavioral factors such as smoking and drinking.

During the past 20 years, most countries in the OECD have undergone health reform efforts that have included strengthening their primary care systems. Much of this reform has been initiated in an effort to control costs, by introducing primary care gatekeeping, for example. There is evidence that these efforts have been partially successful in at least slowing the rate of growth in health care costs (Delnoij et al. 2000; Organization for Economic Cooperation and Development 1995b). With a few notable exceptions, less attention has been paid to the impact of these reform efforts on improving population health.

Nevertheless, based on this analysis, there are still several features of primary care systems that could be better addressed within most OECD countries. Average scores for each primary care component for all countries show that, by and large, most countries are performing well on system characteristics such as system financing, type of primary care provider, and access (average score for each is greater than one). Nearly universal areas of deficiency include practice components such as patient lists (longitudinality), coordination, and community orientation (average score for each is less than 0.5 out of a total possible score of 2).

Several anomalies are present within the analysis and deserve further exploration. First is the apparent difference between gender-disaggregated rates. Primary care was not significantly related to women's all-cause mortality after adjusting for all covariates, even though it was related to male mortality. This difference may be due to different causes of death and disability (e.g., higher rates of mortality from heart disease in men), the differential impact of environmental and behavioral factors on women's and men's health (e.g., tobacco and alcohol use were significantly associated with poorer health outcomes only in the case of men and not women), differences in life expectancies, or differential patterns of primary health care use by men and women. This study does not provide the means to test these hypotheses directly, but does suggest that investigation of gender differences is an important area for further inquiry.

The impact of alcohol and tobacco would be thought to be stronger than found in this study. International studies have consistently found tobacco to be an important determinant of premature mortality (World Health Organization 1997). One explanation for the lack of significant findings in this study may have to do with the definition of tobacco and alcohol consumption used. Because of data availability, this study used a measure of grams of tobacco consumed per capita. A more appropriate measure of health risk would be "percent of population that smokes every day," since grams per capita does not reveal how much of the population is actually at risk of ill health due to tobacco smoke.

The relationship of alcohol to premature mortality is more complex. Studies have shown that the relationship between alcohol consumption and premature mortality is actually U-shaped (Liao et al. 2000). Moderate drinking has been associated with lower premature mortality, while higher drinking and alcohol dependency are associated with higher mortality rates (Dawson 2000). This relationship is not easily modeled here because the measure of alcohol consumption used is liters per capita, whereas a more appropriate measure of health risk might be "percent of population that consumes more than one alcoholic drink per day."

Second, expecting even aggregate individual-level measures to explain macro-level variation in health outcomes among countries can be viewed as a variation of the ecological fallacy. It is clear that within any individual country, smokers have a higher risk of premature mortality than nonsmokers. But countries with high rates of smoking (e.g., France and Japan) do not necessarily have higher rates of premature deaths than countries with lower rates of smoking (Macintyre and Ellaway 2000). This consideration limits the extent to which this study can propose mechanisms for the effect of primary care on health outcomes.

Sensitivity tests included removing five highly influential outlying points: Denmark 1992, 1994, 1995; Japan 1970; and Portugal 1971. Excluding these points did not significantly alter regression results, but it did improve overall model fit for several outcomes. Additional sensitivity tests employed statistical models that adjust for heteroskedastic errors and panel level AR (1) auto-regression. The results of these analyses were not significantly different from those obtained using the fixed effects models reported here.

Limitations

Although the primary care scoring system used here improves on previous studies by incorporating a time dimension, there are aspects of the scale that could be enhanced. First, the study is ecological in nature. This limits the extent to which causal relations between primary care and health outcomes can be drawn. Moreover, because of the statistical models employed (fixed effects regression) the results of the study cannot be generalized to other countries not already included in the sample. Second, although the primary care score is helpful in determining the overall contribution of primary care to population health, it does not provide specific policy advice as to which part of the primary care system a policymaker might want to improve to most effectively benefit population health. This is partly because the objective of the study was to ascertain the overall impact of primary care systems on population health and not to discern the relative contribution of specific primary care features. Future studies should test the effects of weighting different components in terms of their contribution to different health goals.

For example, if evaluating primary care systems in terms of their impact on equity, then perhaps geographic regulation, financing, and access (copayments) would weigh more heavily than provider type and family orientation. Further work needs to be done to refine the scaling system to reflect policy-relevant domains.

Third, two important considerations are not included in this study because there are no comparable international data available for multiple years. The first of these is the quality of care delivered. Countries could have a well-organized and well-funded primary care system, but the actual care provided could be inappropriate or of very low quality. Although some attempt was made to adjust scores for implementation, overall health system quality is not explicitly captured in the scale.

Equity is another important consideration. Health policy objectives are usually concerned not only with overall population health, but also with the distribution of health and health care resources within countries and across population groups. Although several measures of health equity have been proposed (Gakidou, Murray, and Frenk 2000; van Doorslaer et al. 2000), there are few data and little consensus on which measures are most appropriate for international comparisons (Macinko and Starfield 2002). Further use of the primary care score developed here could be used to assess the contribution of primary care (or its components) on the distribution of health outcomes within and among countries.

Fourth, the model of health production is crude. More comprehensive models should include policy, political, social, and cultural factors that modify both the health system and individual propensities for illness. For example, there has been much discussion of factors such as social capital as determinants of population health. Although the measures of social capital employed in health studies are far from ideal (Macinko and Starfield 2001), there is some evidence that they do play a role in population health (Berkman 1986; Bobak et al. 2000). If and when comparable cross-national time series data become available, they should be included in more complete models. Ideally, a multilevel design could be used to estimate the separate influences of individual, health system, and country-level factors.

Finally, dynamic and reciprocal causal effects are not considered in this study. There are several reasons for this. First, previous studies on OECD countries showed that such dynamic effects were not significant (Or 2000). Second, it is not clear what sort of time lag would best model the relationship between primary care system and health outcomes. Third, the possibility of endogeneous independent variables is not addressed directly in this study.

This is partly due to the lack of data on acceptable instruments that could be used to model this endogeneity properly.

CONCLUSIONS

This article presents one of the first cross-sectional, time series analyses of the association between national primary care systems and health outcomes. Keeping in mind the ecological nature of the analysis, and the limitations presented by the data and measures employed, several tentative conclusions can be drawn. First, the financing, organization, and delivery of primary care appear to have a significant impact on health outcomes at the national level. This effect is particularly influential on all-cause and several categories of cause-specific premature mortality thought to be sensitive to primary care. Even though the magnitude of the association is reduced in the presence of other health determinants, primary care still exerted a health-enhancing role on most of the outcomes examined.

Second, it appears that health reform in OECD countries has not uniformly targeted primary care. Those countries that began to reform their primary care systems in the 1970s and 1980s—most notably the Nordic countries, the United Kingdom, Spain, and Italy—have made progress in improving both structural features and practice characteristics of these systems. However, in spite of the potential benefits of improved primary care on population health, countries with the weakest primary care systems—and therefore those with the most potential to benefit from improvements—have, in general, not made much progress in improving either primary care structure or practice.

The argument could be made that those aspects of the primary care system that most need change, such as the system of health care financing, are those that are the most politically and logistically difficult to realize. Such an argument is often made as one explanation for the failure of the United States to make progress in health reform overall. However, an analysis of the primary care systems reviewed here reveals that even if more difficult structural features such as geographic regulation and patient lists cannot be easily implemented, other practice features such as improving coordination and community orientation are lacking in nearly every OECD country. It is hoped that evidence for the potential impact of primary care will give primary care improvement a more prominent place in health reform efforts currently underway within the OECD and elsewhere.

Appendix A: Components of Primary Care Score by Decade[1]

Country	*Year*	*Geographic Distribution*	*Financing*	*Primary Care Provider*	*Primary Care Copay*	*Longitudinality (Patient Lists)*	*First Contact*	*Comprehen-siveness*	*Coorination*	*Family-Centered*	*Community Orientation*	*Primary Care Score*
		0 = none, 1 = limited, 2 = highly regulated	0 = private, 1 = social security, 2 = taxes	0 = specialist 1 = pediatrician / internist, 2 = GP	0 = high 1 = moderate, 2 = none or low	0 = none, 1 = limited 2 = required	0 = no gatekeeper, 1 = limited, 2 = required	0 = not, 1 = somewhat, yes	0 = no guidelines, 1 = limited, 2 = guidelines	0 = none, 1 = somewhat, 2 = entirely	0 = none, 1 = somewhat, 2 = yes	Sum
Australia	1975	0	2	2	2	0	2	2	0	2	0	12
Australia	1985	1	2	2	2	0	2	2	0	2	0	13
Australia	1995	1	2	2	2	0	2	2	0	2	0	13
Belgium	1975	0	1	2	1	0	0	0	0	0	0	4
Belgium	1985	0	1	2	1	0	0	0	0	0	0	4
Belgium	1995	0	1	2	1	0	0	0	0	0	0	4
Canada	1975	0	2	1.5	1	0	1	1	0.5	1	0	8
Canada	1985	0	2	1.5	2	0	1	2	0.5	1	0.5	10.5
Canada	1995	1	2	1.5	2	0	1	2	0.5	1	0.5	11.5
Denmark	1975	2	2	2	2	2	2	2	1	2	1	18
Denmark	1985	2	2	2	2	2	2	2	1	2	1	18
Denmark	1995	2	2	2	2	2	2	2	1	2	1	18
Finland	1975	1	2	2	2	0	1	2	0	0	0	10
Finland	1985	2	2	2	1.5	0	1	2	0	0	0	10.5
Finland	1995	2	2	2	1	1	1	2	1	0	2	14
France	1975	0	1	1	0	0	0	0	0	0	0	2
France	1985	0	1	1	0	0	0	0	0	0	0	2
France	1995	0	1	1	0	0	0	0	0	0	0	2
Germany	1975	0	1	1	2	0	0	0	0	0	0	4
Germany	1985	0	1	1	2	0	0	0	0	0	0	4
Germany	1995	0	1	1	1	0	0	0	0	0	0	3
Greece	1975	0	2	0	2	0	0	0	0	0	0	4
Greece	1985	0	2	0	2	0	0	0	0	0	0	4
Greece	1995	0	2	0	2	0	0	0	0	0	0	4
Italy	1975	0	1	1.5	2	1.5	1.5	1	1	0	1	10.5
Italy	1985	0	2	1.5	2	2	2	1	1	0	2	13.5
Italy	1995	0	2	2	2	2	2	1	1	0	2	14
Japan	1975	0	1	0.5	1	0	1	1	1	1	1	7.5
Japan	1985	0	1	0.5	1	0	1	1	1	1	1	7.5
Japan	1995	0	1	0.5	1	0	1	1	1	1	1	7.5

Appendix A: Continued[1]

Country	Year	Geographic Distribution	Financing	Primary Care Provider	Primary Care Copay	Longitudinality (Patient Lists)	First Contact	Comprehen-siveness	Coorination	Family-Centered	Community Orientation	Primary Care Score
		0 = none, 1 = limited, 2 = highly regulated	0 = private, 1 = social security, 2 = taxes	0 = specialist 1 = pediatrician / internist, 2 = GP	0 = high 1 = moderate, 2 = none or low	0 = none, 1 = limited 2 = required	0 = no gatekeeper, 1 = limited, 2 = required	0 = not, 1 = somewhat, yes	0 = no guidelines, 1 = limited, 2 = guidelines	0 = none, 1 = somewhat, 2 = entirely	0 = none, 1 = somewhat, 2 = yes	Sum
Netherlands	1975	1	1	2	0	2	2	2	1	2	1	14
Netherlands	1985	2	1	2	0	2	2	2	1	2	1	15
Netherlands	1995	2	1	2	0	2	2	2	1	2	1	15
Norway	1975	1	2	2	1	0	2	2	1	2	0	13
Norway	1985	1	2	2	1	0	2	2	1	2	0	13
Norway	1995	1	2	2	1	0	2	2	1	2	0	13
Portugal	1975	1	1	1	0	1	1	1	0	0	0	6
Portugal	1985	1	2	1	0	1	1	1	0	0	0	7
Portugal	1995	1	2	1	0	1	1	1	0	0	0	7
Spain	1975	1	1	1	2	2	1	1	1	0	1	11
Spain	1985	1	1	1	2	2	1	1	1	0	1	11
Spain	1995	2	2	1.5	2	2	2	1	1	2	1	16.5
Sweden	1975	1.5	2	2	1	0	0	1	0	2	0	9.5
Sweden	1985	1.5	2	2	1	0	0	1	0	2	0	9.5
Sweden	1995	2	2	2	1	0	0	1	0	2	1	11
Switzerland	1975	0	0.5	1	0	0	0	1	0	0	0	2.5
Switzerland	1985	0	0.5	1	0	0	0	1	0	0	0	2.5
Switzerland	1995	0	0.5	1	0	0	0	1	0	0	0	2.5
U.K.	1975	1.5	2	2	2	2	2	1.5	0.5	2	1.5	17
U.K.	1985	1.5	2	2	2	2	2	1.5	0.5	2	1.5	17
U.K.	1995	2	2	2	2	2	2	2	1	2	2	19
U.S.A.	1975	0	0	1	0	0	0	0	0	0	0	1
U.S.A.	1985	0	0	1	0	0	0	0	0	0	0	1
U.S.A.	1995	0	0	1	0	0.5	1	0	0	0	0.5	3
Average Score		0.57	1.42	1.17	1.08	0.50	0.77	0.88	0.33	0.64	0.38	7.73

[1]Sources: Starfield 1998; Starfield and Shi 2001; European Observatory 1995–2001; OECD 2000, OECD 2001; Boerma and Fleming 1998; personal communications.

Acknowledgments

The authors would like to thank Juan Gervas, Wienke Boerma, Zeynep Or, Lingxin Hao, and Naoki Ikegami for data and advice. Responsibility for all country ratings, analyses, and conclusions is the authors' alone.

References

Adler, N. 1994. "Socioeconomic Status and Health: The Challenge of the Gradient." *American Psychologist* 49 (1): 15–24.

Almeida, C., P. Braveman, M. Gold, C. Szwarcwald, S. Ribeiro, A. Miglionico, J. Millar, S. Porto, N. Costa, V. Rubio, M. Segall, B. Starfield, C. Travessos, A. Uga, J. Valente, and F. Viacava. 2001. "Methodological Concerns and Recommendations on Policy Consequences of the World Health Report 2000." *Lancet* 357 (9268): 1692–7.

Anderson, G. F., J. Hurst, P. S. Hussey, and M. Jee-Hughes. 2000. "Health Spending and Outcomes: Trends in OECD Countries, 1960–1998." *Health Affairs* 19 (3): 150–7.

Barros, P. 1998. "The Black Box of Health Care Expenditure Growth Determinants." *Health Economics* 7 (6): 533–44.

Berkman, L. 1986. "Social, Networks Support and Health: Taking the Next Step Forward." *American Journal of Epidemiology* 123 (2): 559–62.

Berkman, L., and S. Syme. 1979. "Social Networks, Host Resistance, and Mortality: A Nine-Year Follow-up Study of Alameda Country Residents." *American Journal of Epidemiology* 109 (2): 186–204.

Bindman, A., K. Grumbach, D. Osmond, K. Vranizan, and A. Stewart. 1996. "Primary Care and Receipt of Preventive Services." *Journal of General Internal Medicine* 11 (5): 269–76.

Blakely, T., B. Kennedy, and I. Kawachi. 2001. "Socioeconomic Inequality in Voting Participation and Self-rated Health." *American Journal of Public Health* 91 (1): 99–104.

Blendon, R., K. Minah, and J. Benson. 2001. "The Public versus the World Health Organization on Health System Performance." *Health Affairs* 20 (3): 10–20.

Bobak, M., H. Pikhart, R. Rose, C. Hertzman, and M. Marmot. 2000. "Socioeconomic Factors, Material Inequalities, and Perceived Control in Self-rated Health: Cross-sectional Data from Seven Post-Communist Countries." *Social Science and Medicine* 51 (9): 1343–50.

Boerma, W., and D. Fleming. 1998. *The Role of General Practice in Primary Health Care.* Norwich: The Stationery Office/World Health Organization Regional Office for Europe.

Casanova, C., and B. Starfield. 1995. "Hospitalizations of Children and Access to Primary Care: A Cross-national Comparison." *International Journal of Health Services* 25 (2): 283–94.

Dawson, D. A. 2000. "Alcohol Consumption, Alcohol Dependence, and All-Cause Mortality." *Alcoholism: Clinical & Experimental Research* [erratum in: *Alcoholism: Clinical & Experimental Research* 24 (3): 395] 24 (1): 72–81.

Delnoij, D., G. Van Merode, A. Paulus, and P. Groenewegen. 2000. "Does General Practitioner Gatekeeping Curb Health Care Expenditure?" *Journal of Health Services Research and Policy* 5 (1): 22–6.

DeVellis, R. 1991. *Scale Development: Theory and Applications.* Newbury Park, CA: Sage.

Diggle, P., K. Y. Liang, and S. Zeger. 1994. *Analysis of Longitudinal Data.* New York: Oxford University Press.

Donabedian, A. 1973. *Aspects of Medical Care Administration: Specifying Requirements for Health Care.* Cambridge, MA: Harvard University Press.

Ellencweig, A. Y. 1992. *Analysing Health Systems: A Modular Approach.* Oxford: Oxford University Press.

Elola, J., A. Daponte, and V. Navarro. 1995. "Health Indicators and the Organization of Health Care Systems in Western Europe." *American Journal of Public Health* 85 (10): 1397–1401.

Engel, W., D. Freund, J. Stein, and R. Fletcher. 1989. "The Treatment of Patients with Asthma by Specialists and Generalists." *Medical Care* 27 (3): 306–14.

Enstrom, J. E., and C. W. Heath. 1999. "Smoking Cessation and Mortality Trends among 118,000 Californians, 1960–1997." *Epidemiology* 10 (5): 500–12.

European Observatory on Health Care Systems. 1996a. *Health Care Systems in Transition: Finland 1996.* Copenhagen: European Observatory on Health Care Systems/WHO.

———. 1996b. *Health Care Systems in Transition: Greece 1996.* Copenhagen: European Observatory on Health Care Systems/WHO.

———. 1996c. *Health Care Systems in Transition: Sweden 1996 (Preliminary).* Copenhagen: European Observatory on Health Care Systems/WHO.

———. 1999b. *Health Care Systems in Transition: Luxembourg 1999.* Copenhagen: European Observatory on Health Care Systems/WHO.

———. 1999d. *Health Care Systems in Transition: Portugal 1999.* Copenhagen: European Observatory on Health Care Systems/WHO.

———. 1999e. *Health Care Systems in Transition: United Kingdom 1999.* Copenhagen: European Observatory on Health Care Systems/WHO.

———. 2000b. *Health Care Systems in Transition: Germany 2000.* Copenhagen: European Observatory on Health Care Systems/WHO.

———. 2000c. *Health Care Systems in Transition: Norway 2000.* Copenhagen: European Observatory on Health Care Systems/WHO.

———. 2000d. *Health Care Systems in Transition: Spain 2000.* Copenhagen: European Observatory on Health Care Systems/WHO.

———. 2000e. *Health Care Systems in Transition: Switzerland 2000.* Copenhagen: European Observatory on Health Care Systems/WHO.

———. 2001a. *Health Care Systems in Transition: Denmark 2001.* Copenhagen: European Observatory on Health Care Systems/WHO.

———. 2001b. *Health Care Systems in Transition: Italy 2001.* Copenhagen: European Observatory on Health Care Systems/WHO.

Gakidou, E., C. Murray, and J. Frenk. 2000. "Defining and Measuring Health Inequality: An Approach Based on the Distribution of Health Expectancy." *Bulletin of the World Health Organization* 78 (1): 42–54.

Hsiao, C. 1986. *Analysis of Panel Data.* Cambridge: Cambridge University Press.

Kawachi, I., B. Kennedy, K. Lochner, and D. Stith. 1997. "Social Capital, Income Inequality, and Mortality." *American Journal of Public Health* 87 (9): 1491–8.

Kohn, R., and K. L. White. 1976. *Health Care: An International Study.* London: Oxford University Press.

Larizgoitia, I., and B. Starfield. 1997. "Reform of Primary Health Care: The Case of Spain." *Health Policy* 41 (2): 121–37.

Lena, H., and B. London. 1993. "The Political and Economic Determinants of Health Outcomes: A Cross-Sectional Analysis." *International Journal of Health Services* 23 (3): 585–602.

Liao, Y., D. L. McGee, G. Cao, and R. S. Cooper. 2000. "Alcohol Intake and Mortality: Findings from the National Health Interview Surveys (1988 and 1990)." *American Journal of Epidemiology* 151 (7): 651–9.

Little, R. J., and D. B. Rubin. 1987. *Statistical Analysis with Missing Data.* New York: Wiley.

Macinko, J., and B. Starfield. 2001. "The Utility of Social Capital in Studies on Health Determinants." *Milbank Quarterly* 79 (3): 387–428.

———. 2002. "Annotated Bibliography on Equity in Health." *International Journal of Equity in Health* 1 (1): 1–20.

Macintyre, S., and A. Ellaway. 2000. "Ecological Approaches: Rediscovering the Role of the Physical and Social Environment." In *Social Epidemiology*, edited by L. Berkman, and I. Kawachi, pp. 332–48. Oxford: Oxford University Press.

Mackenbach, J., K. Stronks, and A. E. Kunst. 1989. "The Contribution of Medical Care to Inequalities in Health." *Social Science and Medicine* 29 (3): 369–76.

McKeown, T. 1976. *The Role of Medicine: Dream, Mirage or Nemesis.* London: Nuffield Provincial Hospitals Trust.

McKeown, T., R. G. Record, and R. D. Turner. 1975. "An Interpretation of the Decline in Mortality in England and Wales during the Twentieth Century." *Population Studies* 29 (3): 391–422.

McKinlay, J., and S. McKinlay. 1977. "The Questionable Contribution of Medical Measures to the Decline of Mortality in the United States in the Twentieth Century." *MMFQ Health and Society* 55 (3): 405–28.

Moore 1992. "The Case of the Disappearing Generalist: Does It Need to Be Solved?" *Milbank Quarterly* 70 (2): 361–79.

Murray, C. J., and A. D. Lopez. 1997. "Global Mortality, Disability, and the Contribution of Risk Factors: Global Burden of Disease Study." *Lancet* 349 (9063): 1436–42.

Or, Z. 2000. *Determinants of Health in Industrialized Countries: A Pooled, Cross-Country, Time Series Analyses.* Paris: Organization for Economic Cooperation and Development.

———. 2001. *Exploring the Effects of Health Care on Mortality across OECD Countries.* Paris: Organization for Economic Cooperation and Development (OECD).

Organization for Economic Cooperation and Development. 1995a. *New Directions in Health Care Policy.* Paris: Organization for Economic Cooperation and Development (OECD).

———. 1995b. *Health Care Reform: Controlling Spending and Increasing Efficiency.* Paris: Organization for Economic Cooperation and Development (OECD).

———. 2001. *Health Data 2001: A Comparative Analysis of Thirty Countries.* Paris: Organization for Economic Cooperation and Development (OECD).

Omran, A. 1971. "The Epidemiological Transition. A Theory of the Epidemiology of Population Change." *Milbank Memorial Fund Quarterly* 49: 509–38.

Roos, N. 1979. "Who Should Do the Surgery? Tonsillectomy and Adenoidectomy in One Canadian Province." *Inquiry* 16 (1): 73–83.

Sayrs, L. 1989. *Pooled Time Series Analysis.* Thousand Oaks, CA: Sage.

Schwartz, S. 1994. "The Fallacy of the Ecological Fallacy: The Potential Misuse of a Concept and the Consequences." *American Journal of Public Health* 84 (5): 819–24.

Scott, A., and J. Hall. 1995. "Evaluating the Effects of GP Remuneration: Problems and Prospects." *Health Policy* 31 (3): 183–95.

Shi, L. 1994. "Primary Care, Specialty Care, and Life Chances." *International Journal of Health Services* 24 (3): 431–58.

Shi, L., and B. Starfield. 2000. "Primary Care, Income Inequalities, and Self-rated Health in the United States: A Mixed-Level Analysis." *International Journal of Health Services* 30 (3): 541–55.

Shi, L., B. Starfield, B. Kennedy, and I. Kawachi. 1999. "Income Inequality, Primary Care, and Health Indicators." *Journal of Family Practice* 48 (4): 275–84.

Starfield, B. 1994. "Is Primary Care Essential?" *Lancet* 344 (8930): 1129–33.

———. 1996. "Is Strong Primary Care Good for Health Outcomes?" In *The Future of Primary Care: Papers for a Symposium held on 13th September 1995*, edited by J. Griffin, pp. 18–29. London: Office of Health Economics.

———. 1998. *Primary Care: Balancing Health Needs, Services and Technology.* New York: Oxford University Press.

Starfield, B., and L. Shi. 2002. "Policy Relevant Determinants of Health: An International Perspective." *Health Policy* 60 (3): 201–16.

Statacorp. 2001. *Stata* (release 7.0) statistical software. College Station, TX: Stata Corporation.

Tacq, J. 1997. *Multivariate Analysis Techniques in Social Science Research.* Thousand Oaks: Sage.

van Doorslaer, E., A. Wagstaff, H. van der Burg, T. Christiansen, D. De Graeve, I. Duchesne, U.G. Gerdtham, M. Gerfin, J. Geurts, L. Gross, U. Hakkinen, J. John, J. Klavus, R.E. Leu, B. Nolan, C. O'Donnell, C. Propper, F. Puffer, M. Schellhorn, G. Sundberg, and O. Winkelhake. 2000. "Equity in the Delivery of Health Care in Europe and the U.S." *Journal of Health Economics* 19 (5): 553–83.

Villalbi, J. R., A. Guarga, M. Pasarin, M. Gil, C. Borrell, M. Ferran, and E. Cirera. 1999. "Evaluacion del impacto de la reforma de la atencion primaria sobre la salud." *Atención primaria* 24 (1): 468–74.

Weiner, J. 1987. "Primary Care Delivery in the United States and Four Northwest European Countries: Comparing the 'Corporatized' with the 'Socialized'." *Milbank Quarterly* 65 (3): 426–61.

World Bank. 1993. *World Development Report 1993: Investing in Health.* New York: Oxford University Press.

World Health Organization. 1997. *Tobacco or Health: A Global Status Report.* Geneva: World Health Organization.

———. 2000. *World Health Report 2000.* Geneva: World Health Organization.

World Organization of National Colleges, Academies and Academic Associations of General Practitioners/Family Physicians. 1991. *The Role of the General Practitioner/ Family Physician in Health Care Systems.* Victoria: World Organization of National Colleges, Academies and Academic Associations of General Practitioners/ Family Physicians (WONCA).