The Effect of First-Contact Care with Primary Care Clinicians on Ambulatory Health Care Expenditures

Christopher B. Forrest, MD, PhD, and Barbara Starfield, MD, MPH
Baltimore, Maryland

Background. A study was undertaken to examine the relationship between first-contact care, an essential feature of primary care, and expenditures for frequent ambulatory episodes of care in a nationally representative sample.

Methods. A nonconcurrent cohort study was conducted using data from the 1987 National Medical Expenditure Survey. Ambulatory claims data of respondents with an identified primary care source were used to develop 20,282 episodes of care for 24 preventive and acute illness conditions. The study examined the relationship of first-contact care, defined as the use of an identified primary care source for the first visit in an episode, and ambulatory episode-of-care expenditures.

Results. Episodes that began with visits to an individual's primary care clinician, as opposed to other sources of care, were associated with reductions in expenditures of 53% overall ($63 vs $134, \( P < .001 \)), 62% for acute illnesses ($62 vs $164, \( P < .001 \)), and 20% for preventive care ($64 vs $80, \( P < .001 \)). For 23 of the 24 health problems studied, first-contact care was associated with reductions in expenditures. Multivariate regression analyses that controlled for sociodemographic characteristics, health status, case-mix, length of the episode, and number of visits to the emergency room did not substantively alter these results.

Conclusions. First-contact care was associated with reductions in ambulatory episode-of-care expenditures of over 50% in a nationally representative sample. These findings suggest that systems of care that promote the first-contact aspect of primary care may reduce ambulatory expenditures.

Key words. Health care costs; delivery of health care; episode of care; primary health care; health personnel.

(J Fam Pract 1996; 43:40-48)

Driven by the necessity to contain costs, health care in the United States is undergoing dramatic changes in its organization and financing. The basis of ambulatory care appears to be shifting from a specialty focus to a primary care orientation, largely because of the proliferation of managed care organizations. These new systems of care, funded by both public and private payers, rely on primary care clinicians to deliver the majority of a population's health care services while simultaneously acting as fiduciary gatekeepers to contain costs. From 1988 to 1993, the proportion of privately insured individuals enrolled in managed care plans increased from 29% to 51%. A growing number of states are enrolling Medicaid beneficiaries in health maintenance organizations (HMOs), a trend that has been accelerated by regulatory changes. Similarly, mounting political pressure to restrain growth in Medicare expenditures is apt to substantially increase the proportion of Medicare enrollees in HMOs over the next few years.

While a good deal is known about the health care financing techniques that lead to cost savings in HMOs, much less is known about the specific methods by which primary care as an approach to the delivery of the clinical aspects of health care influences costs. There is scant empirical evidence for a cost-savings effect. National health systems with a strong primary care orientation spend less on health care than those with a specialty care...
focus. Within the United States, health care expenditures are inversely related to the ratio of primary care physicians to specialists within small areas. Few studies, however, examine the relationship between specific attributes of primary care—such as first contact, continuity, comprehensiveness, or coordination—and health care expenditures. This type of information is needed to promote the development of high-quality, cost-efficient primary care systems.

Prior studies have found that gatekeeping influences patterns of ambulatory care for privately and publicly insured populations by reducing use of specialists, emergency rooms, and, as a consequence, health care expenditures. Unfortunately, no study has disaggregated the effect of gatekeeping into its financial (eg, limiting use of technology and specialty care) and clinical (eg, serving as the entry point into the medical care system and matching patients' health care needs with appropriate health care resources) components.

In their capacity as gatekeepers, primary care clinicians perform two broadly defined clinical roles. They serve as the entry point for patients presenting to the medical system with new health problems (ie, first-contact care). As one of the essential attributes of primary care, first-contact care has been a traditional role for primary care clinicians. Once health problems are identified, gatekeepers manage patients' care by matching health needs with appropriate health care resources. The relative contribution of each of these functions to cost reductions is unknown.

In this study, we used data from the 1987 National Medical Expenditure Survey, which was conducted at an optimal time for assessing the impact of gatekeeping, since unrestricted access to all aspects of the medical system was still the dominant organizational mode. We hypothesized that individuals beginning an episode of care for a new health problem with the clinician they identify as their primary care source would receive less resource-intensive medical care, resulting in lower expenditures, than if they began the episode with an alternative clinician.

Methods

Data Source and Study Population

The National Medical Expenditure Survey (NMES) provided information on utilization, financing, and expenditures for a representative sample of noninstitutionalized persons in the United States from January 1 through December 31, 1987. Respondents supplied the majority of data items during a series of household interviews separated by 3- to 4-month intervals. To enhance recall of use and expenditures, survey staff encouraged respondents to keep all medical bills and claims information. These bills were used during the household survey to confirm self-reports. Data on use and expenditures were assembled into an ambulatory claims data file.

A medical provider survey supplemented information obtained from the household panel survey. This survey was targeted to presumed high-cost visits, persons with Medicaid, and a 25% sample of visits of respondents who completed the first wave of interviews. The sample of providers was obtained from respondent reports. Data were collected from providers using personal, telephone, and self-administered surveys. The response rate for office-based practices was 75.9%.

We limited the study population to persons who met the following criteria: (1) presence of an identified site of primary care; (2) presence of a specific clinician within the reported site; and (3) a community-based site of primary care, ie, neither a hospital clinic nor an emergency room. There were 19,835 persons (68.1% of the total sample) who met these criteria. Use of the primary care clinician rather than the primary care site was studied because the NMES did not assign provider identification numbers to sites of care for each medical provider visit.

Construction of Episodes of Care

The NMES ambulatory encounter data file, which included ICD-9-CM codes, was used to develop ambulatory episodes of care (AECs). ICD codes were aggregated into more broadly defined diagnostic categories using the method developed by Schneeweiss et al. This method groups ICD codes into broadly defined clinical conditions based on the clinical homogeneity of the codes and the assumption that they would evoke similar physician responses with respect to the cognitive process and resources used. Twenty-four of the 92 Schneeweiss diagnostic clusters were selected based on the following criteria: (1) they are frequently occurring health problems; (2) they would reasonably require the use of services for a period of no longer than 3 months; and (3) they would not be expected to require hospital care. Episodes of care for chronic conditions were excluded because of the difficulty in identifying the start of a new chronic-care episode in a 1-year data set.

The study population of 19,835 persons made 129,255 ambulatory visits. Of these visits, 6,863 (5.3%) were missing ICD information and thus were excluded from further analyses. Visits for pregnancy-related care were also excluded from the data set, because of the focus on conditions that would resolve within 90 days. Preventive care, some laboratory services, and physical and
speech therapy were not assigned an ICD code in the NMES. Respondents, however, were queried about their reason for the encounter for each visit. Using this information, the following categories of "reasons for encounter" were assigned appropriate ICD codes: eye examination; well-child care; immunizations; general medical examination; physical therapy; speech therapy; radiography; computed tomography (CAT) scan; throat, blood, and urine cultures; preadmission testing; and hearing test. Disability conditions were given "X-codes" in the NMES. These unique codes were developed by the National Center for Health Statistics specifically to classify impairments. These codes were reassigned to appropriate ICD codes.

First-contact care was measured in the context of an episode of care. An episode was defined as a set of health care services in the same diagnostic category that occurred within a specified interval. Decision rule logic was developed for grouping visits into ambulatory episodes of care within a diagnostic cluster type. A team of clinicians assigned each diagnostic cluster a "window" period (range, 30 to 90 days), which was defined as the maximum amount of time for which follow-up of the condition would be reasonable. Sensitivity analyses revealed that the conclusions of the study were not altered by changing the duration of the window periods to a maximum of 6 months. Visits that occurred within the interval of a window period for a single diagnostic cluster type were grouped into an AEC. The logic stipulated that each AEC had an empty window period that preceded it ("run-in period") during which no visits for the diagnostic cluster occurred. The last visit in the AEC then became a new starting point for a subsequent run-in period, and the cycle repeated itself. To ensure that a visit could not have been part of an AEC that began in the year preceding the study period, the first run-in period for all diagnostic clusters began on the first day of the study period. Thirty-three percent of all visits with ICD-9-CM information were for one of the 24 selected diagnostic clusters. The decision logic grouped 75% of the 40,884 visits given diagnostic cluster codes into discrete AECs, for a total of 20,282 unweighted AECs.

To examine the accuracy of the computer algorithm for assigning visits to AECs, the raw claims data for a random sample of 5% of the study population were assigned to AECs manually. The results of the manual assignment corresponded perfectly with those of computer assignment.

**Definition of Variables**

The NMES asked individuals whether there was a particular clinic, health center, doctor's office, or other place that he or she usually visits when sick or needing advice about his or her health. This was considered the site of primary care. A primary care site is where persons would turn for the majority of their health needs, whether they were for curative ("sick") care or preventive ("advise about . . . health") care. This is one of the fundamental premises behind the notion of first contact: persons use their site of primary care as an entry point for all new health problems over time, regardless of the nature of the condition. The primary care clinician was defined as the "particular doctor you usually see at [the primary care site]."

The provider identification number of an individual's reported primary care clinician was linked with the identification numbers given to clinicians seen for the first visit of an episode. First-contact care occurred when the first visit was made with the primary care clinician.

Expenditures reflected charges for each visit in an episode and were summed to obtain the expenditures per AEC (dependent variable). When there was duplication of information for a specific visit, expenditure data from the medical provider survey superseded information obtained from the household survey. Expenditures consisted of physician, facility, laboratory, and radiography fees, but did not include any hospital inpatient or pharmaceutical expenses. Expenses reflected payments when third-party payers paid a discounted amount to the provider. Expenditures were assigned to visits that occurred in settings that do not attach a charge for each visit (e.g., HMOs, government-financed clinics, and charities) by imputing data from visits for similar types of services provided in fee-for-service settings. A recent study showed that the medical expenditure estimates from the NMES compare favorably with the Consumer Expenditure Survey, which is an alternative method for estimating costs of health care.

All analyses were risk adjusted using sociodemographics, a measure of illness burden, ie, case-mix, and a rating of health status. Sociodemographic data consisted of age, sex, race/ethnicity, percentage of poverty level, geographic region in the United States, and urban vs rural residence (urban defined as within a standard metropolitan statistical area). Case-mix was measured using the Ambulatory Care Group method for categorizing diagnoses in ambulatory care. The Ambulatory Care Group system of case-mix measurement was developed to predict resource utilization. Using computer software, all ICD codes from all visits, irrespective of assignment to one of the 24 diagnostic clusters, were matched to a unique ambulatory diagnostic group. The number of ambulatory diagnostic groups reflecting the presence of one or more chronic conditions was summed to obtain a measure of the chronic illness burden for each individual.
Respondents’ global rating of their health as excellent, good, fair, or poor was used as the measure of health status.

Data Analysis

To obtain nationally representative parameter estimates, the NMES employed a probability sample of the US population using multistage, stratified cluster sampling. This type of sample survey design complicates variance estimation techniques. Ignoring it during data analysis would lead to spuriously low standard errors because of the effect of clustering.25 Furthermore, each person had an associated sampling weight, which is the number of individuals in the US population that the respondent represented. The sampling weights used in this study adjusted for non-response to both the main questionnaires and to the health status/access to care supplement.26

The statistical package SUDAAN was used in all data analyses to account for both the design effect and the sampling weights.27 SUDAAN computes variance estimates for descriptive statistics, cross-tabulations, and linear regression parameter estimates using a Taylor series method of variance estimation.28

The unit of analysis was ambulatory episodes of care. Analysis of variance with K-1 design variables was used to assess differences between the means of K levels using the REGRESS statistical procedure.27 The F-statistic for the overall model was used to test the null hypothesis that all the means were equal. Multivariate linear regression was performed to control for sociodemographics, health status, case-mix, and type of diagnostic cluster. All controlling variables entered regression models as design variables. Variables related to expenditures at $P<.25$ in univariate analyses were entered into the regression models.

Results

Forty-nine percent of all episodes began with a visit to the primary care clinician. The proportion of episodes associated with first-contact care was highest for children aged 1 to 4 years (61.6%), lowest for young adults aged 18 to 44 (40.9%), and midlevel for those over age 74 years (54.4%). First-contact care was more common for individuals without any chronic condition (50.1%) than for those with two or more (45.1%).

Table 1 presents univariate analyses of episode expenditures for several personal characteristics of the respondents. The relationship between age and expenditures mirrored the U-shaped relationship between age and first-contact care. Children and the elderly incurred the lowest expenditures, and episodes for individuals aged 45 to 64 years were twice as expensive as those for children aged 1 to 4 years. Episodes were more costly for men than women, and more expensive for urban dwellers compared with those in rural locations. Health status was significantly related to ambulatory expenditures. For those in poor health, expenditures were 46% higher than for respondents in excellent health. Episodes for persons with two or more chronic conditions were 18% more expensive than for those with no chronic conditions.

Table 2 presents the mean number of visits and mean
Table 2. Mean Expenditures and Mean Number of Visits per Ambulatory Episode of Care by First-Contact Care, Overall, and for 24 Preventive and Acute Illness Conditions

<table>
<thead>
<tr>
<th>Reason for Episode</th>
<th>No. of Episodes*</th>
<th>Visits per Episode</th>
<th>Expenditures per Episode†</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>With and Without</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>First-Contact Care</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>With</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Without</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>With, $</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Without, $</td>
<td></td>
</tr>
<tr>
<td>Overall</td>
<td>20,282</td>
<td>1.36</td>
<td>63</td>
</tr>
<tr>
<td>Preventive care only</td>
<td>8,080</td>
<td>1.17</td>
<td>64</td>
</tr>
<tr>
<td>Acute illness conditions</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acute upper respiratory infection</td>
<td>3,742</td>
<td>1.24</td>
<td>39</td>
</tr>
<tr>
<td>Lacerations, contusions, abrasions</td>
<td>1,842</td>
<td>1.82</td>
<td>74</td>
</tr>
<tr>
<td>Acute sprains and strains</td>
<td>900</td>
<td>1.95</td>
<td>74</td>
</tr>
<tr>
<td>Acute lower respiratory infection</td>
<td>748</td>
<td>1.71</td>
<td>76</td>
</tr>
<tr>
<td>Otitis media</td>
<td>861</td>
<td>1.55</td>
<td>58</td>
</tr>
<tr>
<td>Dermatitis</td>
<td>439</td>
<td>1.45</td>
<td>53</td>
</tr>
<tr>
<td>Fractures and dislocations</td>
<td>480</td>
<td>2.72</td>
<td>211</td>
</tr>
<tr>
<td>Urinary tract infection</td>
<td>733</td>
<td>1.74</td>
<td>65</td>
</tr>
<tr>
<td>Vaginitis, vulvitis, cervicitis</td>
<td>241</td>
<td>1.26</td>
<td>62</td>
</tr>
<tr>
<td>Bursitis, synovitis, tenosynovitis</td>
<td>219</td>
<td>1.82</td>
<td>198</td>
</tr>
<tr>
<td>Gastroenteritis</td>
<td>201</td>
<td>1.32</td>
<td>54</td>
</tr>
<tr>
<td>Nonfungal skin infection</td>
<td>238</td>
<td>1.59</td>
<td>65</td>
</tr>
<tr>
<td>Headache</td>
<td>438</td>
<td>1.62</td>
<td>106</td>
</tr>
<tr>
<td>Conjunctivitis</td>
<td>186</td>
<td>1.17</td>
<td>25</td>
</tr>
<tr>
<td>Viral warts</td>
<td>146</td>
<td>1.93</td>
<td>72</td>
</tr>
<tr>
<td>Viral exanthems</td>
<td>161</td>
<td>1.24</td>
<td>58</td>
</tr>
<tr>
<td>Burns</td>
<td>84</td>
<td>1.31</td>
<td>42</td>
</tr>
<tr>
<td>Abdominal pain</td>
<td>92</td>
<td>1.93</td>
<td>173</td>
</tr>
<tr>
<td>Mononucleosis and hepatitis</td>
<td>42</td>
<td>2.84</td>
<td>154</td>
</tr>
<tr>
<td>Chest pain</td>
<td>140</td>
<td>1.74</td>
<td>165</td>
</tr>
<tr>
<td>Fatigue</td>
<td>48</td>
<td>1.92</td>
<td>112</td>
</tr>
<tr>
<td>Scabies, pediculosis, helminthiasis</td>
<td>64</td>
<td>1.22</td>
<td>13</td>
</tr>
<tr>
<td>Diseases of nail</td>
<td>157</td>
<td>1.60</td>
<td>61</td>
</tr>
</tbody>
</table>

*The number of unweighted episodes of care is reported.
†First-contact care occurred when individuals began episodes with their primary care clinicians.
‡95% confidence intervals of the two means do not overlap.

Type of diagnostic cluster was an important determinant of the duration and expense of an episode. The longest and most expensive AECs were for injuries and musculoskeletal conditions, such as fractures/dislocations, sprains/strains, and bursitis/tenosynovitis. Minor infections, including vaginitis/cervicitis, otitis media, gastroenteritis, and respiratory infections, were among the shortest and least expensive types of episodes.

First-contact use of an individual's self-reported primary care clinician was significantly associated with reductions in episode expenditures of 53% for all types of episodes ($63 vs $134, P<.001), 62% for acute illness episodes ($62 vs $164, P<.001), and 20% for preventive care episodes ($64 vs $80, P<.001). The cost-savings effect of first-contact care was found for preventive care and all types of acute illness episodes except fatigue. For the majority (15/23) of acute illness AECs, the size of the cost-savings effect of first-contact care ranged between 30% and 60%.

Expenditures were directly related to the number of visits in the episode. The addition of one visit increased episode expenditures by an average of $77. Episodes with three or more visits were six times as expensive as those with a single visit ($378 vs $58, P<.001). The effect of first-contact care on number of visits in an episode was weaker than the cost-savings effect. Overall, first-contact care was associated with 0.32 fewer visits per episode (P<.001), a 19% reduction in the length of an AEC. For 10 types of episodes, the percent reduction was less than 10%, and for 8 others, first-contact care was associated with a statistically insignificant increase in the number of visits.
Individuals who began an episode with their primary care clinician rather than using more costly types of care, such as the emergency room (ER), at the beginning of an episode of care incurred lower expenditures. AECs that began with a visit to the ER were approximately four times more expensive than those that began with visits to individuals’ self-identified primary care clinicians ($255 vs $63, \( P < .001 \)). While use of the ER had a substantial effect on expenditures for an AEC, only 9% of AECs were associated with any use of the ER. Furthermore, first-contact care decreased the chances of using the ER during an ambulatory episode. The odds that episodes beginning with a visit to the primary care clinician involved a subsequent ER visit were reduced by 37% relative to episodes that began with a visit to any other non-ER provider (odds ratio: 0.63, 95% confidence interval [CI]: 0.41 to 0.96).

Table 3 presents a series of multivariate linear regression models that test the alternative hypotheses that the reduction in expenditures associated with first-contact use of the primary care clinician could be a result of differences in the personal characteristics of survey respondents, duration of the episode, or number of emergency room visits. Adjusting for sociodemographics, health status, and case-mix did not substantively alter the magnitude of the cost savings effect of first-contact care on expenditures for any of the conditions in the table (model 1). Adjustment for duration of the episode as well as personal characteristics of the respondent (model 2) decreased the cost-savings effect of first-contact care for preventive care episodes by 31% and for acute illness episodes by 37%. Thus, the majority of the effect of first-contact care was primarily a result of using fewer resources, which resulted in decreased expenditures per visit rather than fewer visits per episode.

Model 3 in Table 3 demonstrates that use of the ER accounted for part of the reductions in expenditures attributable to first-contact care. Controlling for both personal characteristics of respondents and number of ER visits per episode was associated with a 39% overall decrease of the effect of first-contact care on episode expenditures. For acute lower respiratory infections and burns, however, the decrease in the cost-savings effect of first-contact care associated with adjustment for the number of ER visits in the episode was substantially greater (73% and 65%, respectively).

When AECs in which the first visit was to the ER were removed from the analysis, first-contact care was significantly associated with an overall reduction of $59 in episode expenditures, which was just 17% lower than the unadjusted effect.

### Discussion

This investigation provides evidence for the cost-efficiency of the first-contact aspect of primary care for individuals in the United States. The study’s main finding is
that significant reductions in ambulatory expenditures were realized when individuals used their primary care clinician rather than other sources of care at the beginning of an episode for a new health problem. This finding was reproduced for 23 of the 24 different types of preventive and acute illness episodes of care, which accounted for one third of the study population's ambulatory utilization. A recent study of patients with back pain similarly found lower resource intensity for patients who had first contact with a primary care clinician rather than alternative sources of care, such as chiropractors and orthopedic surgeons.

The unadjusted effect of age on episode expenditures suggested that young and middle-aged adults were more costly than both children and the elderly. This finding can be explained by a greater tendency of the young and elderly to use primary care clinicians as their entry point into the medical system and by differences in patterns of morbidity across the age groups. Young and middle-aged adults reported a greater proportion of their care for the more expensive injury and musculoskeletal episodes. Regression models that controlled for type of condition showed that episodes for children up to age 17 were less expensive than for adults and elderly individuals, but there was no difference in expenses for individuals over age 18.

Although regression models in this study were risk adjusted with patient sociodemographic characteristics, case-mix, and health status, we did not directly control for the severity of the primary diagnosis. It is possible that individuals who had health problems with greater severity of illness chose to use a clinician other than their primary care source. Currently, there are no adequate methods to measure severity of illness using claims data for undifferentiated and common acute conditions of the sort investigated in this study (eg, upper respiratory infections, urinary tract infections, preventive care, and abdominal pain). Moreover, it is unclear whether patients make decisions about where to enter the medical system based on perceived severity of their condition. One study found that baseline functional status for patients with back pain was no different between those who used primary care clinicians and those who used orthopedic surgeons as their entry point into the medical system. Lastly, the consistency of the effect of first-contact care on expenditures across a broad range of acute illness conditions, some of which have minimal variation in severity (eg, upper respiratory infections, otitis media, urinary tract infection, gastroenteritis, viral warts, and viral exanthems) argues against severity as a principal explanation for the study's findings.

Differences in sociodemographics, health status, case-mix, number of visits, and use of the emergency room between episodes with and without first contact explained a small part of the reductions in expenditures associated with first-contact care. The bulk of the cost-saving effect of first-contact care is probably a result of lower payments for physician services to primary care clinicians, less resource-intensive practice styles of generalists as compared with specialists, differences in severity of illness, and lower reimbursement for established rather than new patients. Prior research has found less diagnostic testing for visits of established rather than new patients. Disaggregating these components will be an important challenge for further inquiry into the efficiency of first-contact care. Moreover, because this study did not examine how first contact affects the quality or outcomes of care, more research is needed to determine how the cost-efficiency of first-contact care influences these critical aspects of health care delivery.

This investigation does not provide estimates of the effect of first-contact care for individuals whose regular source of care is reported to be a hospital or emergency room, even if they had a specific physician within that site; for those without an identified source of care; or for respondents who consider advanced practice nurses or physician assistants to be their primary care source. Furthermore, the effect of first-contact care for a primary care site, rather than a clinician, was not studied because limitations in data collection precluded linking an identified primary care facility with sites used for ambulatory services. Future research should focus on estimating the cost-savings effect of first contact with primary care sites (such as staff/group model HMOs and primary care group practices), nonphysician practitioners, and hospital-based clinicians.

This study's findings have important policy implications for the organization and financing of health care delivery systems. The first step in promoting first-contact care is to ensure that all members of an eligible population identify a specific clinician for their primary care gatekeeper. In 1987, nearly one in six individuals in the United States did not have a specific primary care clinician (based on data used for this study). Traditional fee-for-service delivery systems do not require individuals to register with a specific primary care clinician. One possibility to increase the proportion of the US population that identifies a specific primary care clinician is through enrollment into delivery systems in which individuals select primary care gatekeepers. The growth of managed care health systems may enhance patient identification with a specific primary care source, which also serves to improve primary care clinician accountability for the health care of their population of patients. Merely having a primary care source does not, however, assure first-contact care. Facilitating access to the primary care source so that care can be obtained for a new health problem when it is needed will...
Acknowledgment

Supported by NRSA training grants F32-HS00070 and T32-HS00029 from the Agency for Health Care Policy and Research, Public Health Service, DHHS, Rockville, Maryland.

References


