

© Health Research and Educational Trust
DOI: 10.1111/1475-6773.12274
RESEARCH ARTICLE

Changes in Quality of Health Care Delivery after Vertical Integration

Caroline S. Carlin, Bryan Dowd, and Roger Feldman

Objectives. To fill an empirical gap in the literature by examining changes in quality of care measures occurring when multispecialty clinic systems were acquired by hospital-owned, vertically integrated health care delivery systems in the Twin Cities area.

Data Sources/Study Setting. Administrative data for health plan enrollees attributed to treatment and control clinic systems, merged with U.S. Census data.

Study Design. We compared changes in quality measures for health plan enrollees in the acquired clinics to enrollees in nine control groups using a differences-in-differences model. Our dataset spans 2 years prior to and 4 years after the acquisitions. We estimated probit models with errors clustered within enrollees.

Data Collection/Extraction Methods. Data were assembled by the health plan's informatics team.

Principal Findings. Vertical integration is associated with increased rates of colorectal and cervical cancer screening and more appropriate emergency department use. The probability of ambulatory care-sensitive admissions increased when the acquisition caused disruption in admitting patterns.

Conclusions. Moving a clinic system into a vertically integrated delivery system resulted in limited increases in quality of care indicators. Caution is warranted when the acquisition causes disruption in referral patterns.

Key Words. Vertical integration, quality of care, provider consolidation

The development of Accountable Care Organizations (ACOs) is a key part of provider payment reform in the United States. An ACO typically is formed by integrated hospitals and clinics, allowing the integrated delivery system (IDS) to manage the full spectrum of care and associated costs. The acquisition of physician practices by U.S. hospital systems has accelerated in recent years, in part due to payment reform. More than half of U.S. physicians are now employed by hospitals or integrated delivery systems (Kocher and Sahni 2011), facilitating development of ACOs in anticipation of higher quality, lower cost care.

Little empirical research has examined the impact of integration of hospitals and clinic systems on quality of care, despite theoretical groundwork by Burns and Muller (2008), Gal-Or (1999), and Trybou, Gemmel, and Annemans (2011). This empirical study analyzed the acquisition of three large, multispecialty clinic systems in the Minneapolis–St. Paul metropolitan area by two hospital-owned IDSs at the end of 2007. Using data that allow us to control for health plan enrollee, coverage, and clinic system characteristics, we examined the impact of acquisitions on claims-based quality of care indicators. Our research design allowed us to identify differences in quality indicators driven by the acquisitions, controlling for preacquisition characteristics of the acquired practices.

BACKGROUND

Conceptual Model

In our conceptual model, providers are motivated on the margin to improve quality due to market forces imposed by publication of quality metrics. Early efforts to measure health care quality focused on the quality of inpatient hospital procedures (Mukamel and Mushlin 1998). At the national level, publication of quality metrics for clinic systems was uncommon in 2007, when these acquisitions took place. However, physician- and clinic-level quality measures are increasingly common. UnitedHealth's Premium[®] Designation Program identifying low-cost, high-quality providers, and NCQA's Diabetes and Heart/Stroke Physician Recognition Programs are two examples of national quality measurement programs that focus on outpatient care. Quality recognition programs have become so widespread that the Robert Wood Johnson Foundation maintains a directory (Robert Wood Johnson Foundation 2013) to help patients find high-quality care in their communities, and many states have physician- or clinic-level quality measures available. The increasing prevalence of clinic-level quality measures makes this Minnesota-based study of national interest.

Minnesota Community Measurement (www.mncm.org), a collaboration between Minnesota health plans and the Minnesota Medical Association, was

Address correspondence to Caroline S. Carlin, Ph.D., Medica Research Institute, MR-CW105, PO Box 9310, Minneapolis, MN 55440-9310; e-mail: caroline.carlin@medica.com. Bryan Dowd, Ph.D., and Roger Feldman, Ph.D., are with the Division of Health Policy and Management, University of Minnesota, Minneapolis, MN.

launched as a nonprofit 501(c)(3) organization in 2005. Their earliest comprehensive report, based on 2003 visits and published in 2004, summarized quality of care metrics for diabetes, asthma, depression, and high blood pressure. In addition, quality of preventive care was summarized by child and adolescent immunizations, well-child visits, and women's health (screening for breast and cervical cancer, and chlamydia). All three of the acquired clinic systems and eight of nine control groups participated in this initial report. By 2005, preventive care measures were expanded to include colorectal cancer screening and all of our acquired and control clinic systems were participating in Minnesota Community Measurement. To respond to publication of quality measures, health care delivery systems need to invest in clinic infrastructure, including electronic medical records, and attract new talent to replace retiring physicians. It might be difficult for smaller clinic systems to raise the capital necessary to finance these improvements in infrastructure.

As a part of this project, we conducted a series of key informant interviews to understand the context in which these acquisitions took place, summarized in Christianson, Carlin, and Warrick (2014). In these interviews, it became clear that access to capital for improved electronic medical records and other infrastructure was perceived as a key factor in quality improvement. The authors state, "Virtually all [acquired] respondents identified the perceived need on the part of practices to invest in electronic medical records (EMRs) as the most important single motivating factor in their desire to be acquired by an IDS. . . [EMRs] also were seen as important in moving beyond claims data for measuring and publicly reporting patient outcomes."¹ These practices also expressed a need for greater leverage with health plans to secure reimbursement rates that support EMR investment, quality improvement activities, and health care home certification. They stated that major infrastructure improvements were necessary to retain existing and attract new physicians.

Acquiring IDSs are motivated to increase their physician base, particularly in primary care. From Christianson, Carlin, and Warrick: "IDSs see primary care physicians, linked to other IDS components through organizational structures, financial incentives, and a common EMR, as essential for developing a continuum of care." Acquiring IDSs also believe EMRs have potential to improve quality through better care coordination and use of embedded clinical guidelines.

Thus, by entering into an acquisition agreement, the multispecialty practices expected to improve their infrastructure (including EMR) and increase their ability to attract top clinical talent, both of which could plausibly

improve their ability to respond to the publication of quality metrics. The quality metrics published by Minnesota Community Measurement include the three types of cancer screening we evaluated, where we expect to see a direct effect of improved EMR systems on improved population health. With better infrastructure and talent, the acquired entities should be positioned to improve care for diabetes, asthma, and high blood pressure, motivated by publication of these quality metrics. The improvements in care for these chronic conditions may result in reductions in inappropriate emergency department (ED) use and ambulatory-care sensitive (ACS) admissions. Better continuity of care—as postulated by the acquiring IDs above—may result in lower readmission rates. Reductions in inappropriate use of inpatient and ED services may eventually result in overall reductions in incidence of inpatient admissions and ED visits, but this may be difficult to detect in our relatively short 4-year follow-up period.

Previous Literature

In their recent review of the empirical literature examining the impact of integrated delivery systems on cost and quality of care, Hwang et al. (2013) found only four peer-reviewed studies (Shortell et al. 2005; Mehrotra, Epstein, and Rosenthal 2006; Rittenhouse et al. 2010; Weeks et al. 2010) that examined the link between clinic ownership, or clinic size and structure, and quality of care. Mehrotra, Epstein, and Rosenthal (2006) compared horizontally integrated medical groups (IMGs)² with the more decentralized independent physician associations, finding that IMGs provided higher levels of preventive care screening. Rittenhouse et al. (2010) found that physician practices owned by a hospital or health maintenance organization were more likely to use evidence-based care management processes. Shortell et al. (2005) found that medical groups affiliated with a hospital or health plan were significantly more likely to be in the top quartile of care management and health promotion indices. Finally, Weeks et al. (2010) compared the care received by Medicare patients in large multispecialty groups affiliated with the Council of Accountable Physician Practices (CAPP), a consortium³ of 27 large group practices, against care delivered by other practices in the same markets. They found that patients assigned to the CAPP practices received higher levels of evidence-based care.

These studies are all cross-sectional in design, and thus subject to selection concerns, for example, are improvements in quality driven by these organizational affiliations, or are high-quality medical groups more attractive

targets for the affiliations described in the literature? In addition, none of these papers directly compared vertically integrated systems with independent physician-owned practices.

Our study addresses these methodological issues through a differences-in-differences model using a longitudinal health care claims database with observations from clinic systems acquired by vertically integrated IDSs, and from control clinic systems. We developed variables to measure processes of care and health outcomes for preventive cancer screening, inpatient experience, and emergency department experience, and used our model to identify the impact of vertical integration on these measures. This study is an important contribution to the sparse empirical literature studying the impact of vertical integration between hospitals and clinic systems.

METHODS

Study Setting

Near the end of 2007, three large, multispecialty clinic systems in the Minneapolis–St. Paul region were acquired by two hospital-owned IDSs. Specifically, the two clinic systems we label Acq1A and Acq1B were acquired by IDS1, and Acq2 was acquired by IDS2. A regional health plan provided data from enrollees attributed by the plan to one of 12 clinic systems: the three acquired clinic systems, the legacy clinic systems in the two acquiring IDSs, or one of seven other clinic systems not affected by acquisitions.⁴ The health plan attributed enrollees to clinic systems retrospectively based on primary care spending. Primary care was defined by place of service (office visits and, for Medicaid enrollees, emergency department visits) and physician specialty code (general practice, internal medicine, family medicine, and OB-GYN). Attribution was based on allowed charges after provider discounts, including amounts paid by plan and enrollee.

Observations from enrollees attributed to the legacy clinics in IDS1 and IDS2, and seven additional clinic systems served as controls. Among the control systems, IDS3, IDS4, and IDS5 are IDSs with a variety of ownership statuses: one owned by an insurer, one by the physicians, and one by the hospital system. Two additional control systems (MS1, MS2) are physician-only multispecialty clinic systems and two (PC1, PC2) are physician-only clinic systems with limited specialty services.

Data

We used commercial insurance and managed Medicaid data provided by a health plan in the upper Midwest for enrollees whose care was attributed to one of the 12 clinic systems in our study. The attributed enrollees represent more than a quarter of the population in the plan's commercial and Medicaid populations. Our observation period includes 2 years prior to the acquisitions (2006–2007) and 4 years after the acquisitions (2008–2011), allowing us to control for time-invariant, unobserved characteristics of the clinic systems. We used the Johns Hopkins ACG system (Weiner, Starfield, and Lieberman 1992) to capture the enrollee's prior-year health status.⁵ The system-generated categorical variable, "resource utilization band," indicates no diagnosis on record, indicates the enrollee is a healthy user of care, or specifies a low, moderate, high, or very high health risk category for the enrollee. We also developed neighborhood effects using the 2011 5-year American Community Survey (U.S. Census Bureau 2012), matched at the census-tract level to the enrollee's address. These neighborhood effects captured race and ethnicity, language, education, and income.

Variables Measuring Quality of Care

The purpose of this study was to estimate the impact of vertical integration of clinic systems into IDSs on indicators of quality of care, measured by processes of care and health outcomes. (We study impacts on the price of care elsewhere [Carlin, Feldman, and Dowd unpublished data]). Specifically, we measured:

- Preventive screening for breast, colorectal, and cervical cancer;
- Inpatient admissions, ACS admissions, and admissions followed by a readmission within 30 days;
- Emergency department (ED) visits and appropriateness of the ED visits.⁶

Populations eligible for cancer screening were identified by age and gender based on the regionally accepted Institute for Clinical Systems Improvement (ICSI) guidelines (ICSI 2013), excluding those with a prior history of the disease. Male and female enrollees aged 50–75 were eligible for colorectal cancer screening; women aged 21–65 were eligible for cervical cancer screening; women aged 50–75 were eligible for breast cancer screening. For each calendar year, we identified routine mammograms, colonoscopies, or pap tests from CPT-4 codes for outpatient claims. ICSI's Adult Preventive Services

guidelines in 2006–2011 recommended screening mammograms every 1–2 years, colonoscopies every 10 years, and pap tests every 3 years after three normal pap tests. We excluded women between ages 40 and 50 from the breast cancer screening analysis because of disagreements among clinicians about the appropriateness of routine screening in this age range. We modeled the probability that the cancer screening occurred for each eligible group, defining quality improvement as an increase in that probability.

Inpatient admissions excluded claims with a discharge status indicating an ongoing admission. We used the Billings ACS algorithm (Billings 2013a) to identify ACS admissions. Readmissions were identified by selecting admissions with a live discharge code, and looking for subsequent readmissions that occurred within 30 days. We excluded readmissions for chemotherapy or rehabilitation, assuming these were planned and appropriate. We also excluded readmissions for maternity claims; such readmissions are frequent due to preterm labor admissions and were presumed appropriate.

We modeled the probability of an inpatient admission. For those with an inpatient admission, we also modeled probabilities that the admission was ACS, and the admission was followed by a readmission within 30 days. We define quality improvement as a reduction in any of the three probabilities.

ED claims were identified using place of service codes. We used the Billings ED algorithm (Billings 2013b) to classify these visits according to type and necessity. This algorithm uses ICD-9 diagnosis codes to assign probabilities of falling into the one of four classifications: nonemergent; emergent but primary care treatable; emergent and ED needed, but preventable or avoidable; and emergent and ED needed, not preventable or avoidable. Following Dowd et al. (2014), we modified the algorithm to use the diagnosis code maximizing the probability the visit was emergent and not preventable or avoidable. This avoids penalizing clinic systems for differences in ordering of diagnosis codes. The modification made a substantive difference, raising the average probability of an ED visit being emergent and not preventable or avoidable from 20 to 28 percent. We developed a binary indicator of appropriateness by identifying ED visits with a positive probability (greater than 0 percent) of being emergent and not preventable or avoidable.

We modeled the probability of one or more ED visits in a year. For those with an ED visit, we modeled the probability that the ED visit was appropriate. We define quality improvement as a decline in the probability of ED use or an increase in the probability the visit was appropriate.

Econometric Models

Each observation of a quality indicator is a binary variable (y_{ijt}), where $y_{ijt} = 1$ is observed when the latent variable $y_{ijt}^* \geq 0$, and $y_{ijt} = 0$ is observed when $y_{ijt}^* < 0$. The differences-in-differences model for y_{ijt}^* is:

$$y_{ijt}^* = \beta_1' x_{ijt} + \sum_{j=1}^{12} \beta_{2j} \text{ClinicSystem}_j + \sum_{t=2006}^{2011} \beta_{3t} \text{Year}_t + \sum_{j=1}^3 \beta_{4j} \text{Treat}_{jt} + \varepsilon_{ijt}$$

The *ClinicSystem* indicators identify the enrollee's attributed clinic system; *Year* indicators identify year of observation; and *Treat* indicators identify whether the observation was from an acquired clinic system in a postacquisition year. Acquired clinic systems were identified by $j = 1-3$, and postacquisition years by $t = 2008-2011$. The three values of β_{4j} identify treatment effects, that is, the impact of acquisition on quality of care indicators. The vector x_{ijt} includes covariates that control for observed characteristics of the enrollees and their environments (Table 1). The error terms ε_{ijt} are assumed to be iid standard normal (probit model), with errors clustered within enrollee.

RESULTS*Characteristics of the Sample*

Summary statistics from the 796,962 person-years of data in our study are provided in Table 1. The population has an average age of 42, and 59 percent are female. Because the attribution method defining our sample relies on receipt of primary care, these attributed enrollees were a little older (1.3 years, on average) and more female (7 percentage points) than the insurer's full covered population, which includes enrollees who never sought care.

Nearly half (49 percent) the population is assigned a moderate health risk based on their diagnosis code history. On average, enrollees' neighborhoods are primarily white and non-Hispanic (84 percent), have residents who speak English only (88 percent), have residents with at least a high school degree (56 percent without a 4-year college degree, 36 percent with a 4-year degree), and have few households below the federal poverty limit (9 percent). There is a considerable amount of variation in these neighborhood effects across observations, as shown in Table 1.

The most common type of insurance coverage is a broad PPO network, but there is significant enrollment in a restricted-network plan that

Table 1: Summary of Data

	<i>Mean</i>	<i>Minimum</i>	<i>Maximum</i>
Enrollee demographics			
Age	41.7	18	107
Female	59.30%		
Prior-year health status			
No diagnosis history	8.39%		
Healthy users	11.61%		
Low health risk	14.36%		
Moderate health risk	48.58%		
High health risk	13.08%		
Very high health risk	3.99%		
Neighborhood effects			
% Less than high school/GED	7.39%	0.00%	64.30%
% High school/GED but no Bachelor's degree	56.32%	0.00%	85.80%
% Bachelor's degree or greater	36.29%	0.40%	95.10%
% White, non-Hispanic	84.16%	0.00%	100.00%
% Speaking English only	87.89%	0.00%	100.00%
% Households with income below the federal poverty limit	8.71%	0.00%	92.40%
Product			
Broad PPO plan	66.06%		
Restricted network with medical home designation	16.50%		
Managed Medicaid plan	17.44%		
		Person-years	
Attributed clinic system			
Treatment groups			
Acquired clinic system (Acq1A)		55,368	
Acquired clinic system (Acq2)		38,740	
Acquired clinic system (Acq1B)		20,283	
Control groups			
Clinic system in acquiring integrated delivery system (IDS1)		322,378	
Clinic system in acquiring integrated delivery system (IDS2)		144,505	
Clinic system in integrated delivery system (IDS3)		220,142	
Clinic system in integrated delivery system (IDS4)		73,229	
Clinic system in integrated delivery system (IDS5)		39,919	
Physician-owned multispecialty clinic system (MS1)		30,955	
Physician-owned multispecialty clinic system (MS2)		20,239	
Physician-owned primary care clinic system (PC1)		32,204	
Physician-owned primary care clinic system (PC2)		14,278	

Continued

Table 1. *Continued*

	<i>Mean</i>	<i>Minimum</i>	<i>Maximum</i>
Calendar years			
2006		180,001	
2007		174,194	
2008		168,958	
2009		160,750	
2010		163,405	
2011		164,932	

requires the designation of a medical home (18 percent), and in a managed Medicaid product (16 percent). The distribution of observations by attributed clinic system shows considerable concentration of care delivery in clinics owned by the three largest IDSs (IDS1, IDS2, and IDS3). Together, these three clinic systems represent 68 percent of the attributed enrollee-years in our data.

The following discussion of regression results focuses on statistically significant ($p < .05$) and marginally significant ($.05 < p < .10$) acquisition effects. In addition to parameter estimates, marginal effects (MEs) were computed to provide more intuitive measures of the acquisition impacts. These MEs estimate the average percentage-point (pp) change in probability of the outcome when the acquisition effect variable ($Treat_{jt}$) changes from 0 to 1.

Cancer Screening Results

Breast Cancer Screening. The impact of acquisition on the probability of breast cancer screening (Table 2) varied by clinic system. There was a statistically significant decrease in screening probability after acquisition in Acq1A (ME = -5.3 pp), and an increase in Acq2 (ME = +3.5 pp), indicating an improvement in only one of the three acquisitions. These changes are relative to a 59.5 percent average probability that a woman in the eligible population received a mammogram in any 1 year, averaged across the 6 years in our study.

There was an inverted u-shaped relation between breast cancer screening probabilities and health status, with moderate-risk enrollees having the highest probability of screening. Probabilities increased in highly educated neighborhoods and declined in neighborhoods with higher fractions of the population below the federal poverty limit. There

Table 2: Cancer Screening Regressions: Parameter Estimates and Acquisition Marginal Effects

	<i>BC Screening</i>		<i>CRC Screening</i>		<i>CC Screening</i>	
	<i>Estimate</i>	<i>p-value</i>	<i>Estimate</i>	<i>p-value</i>	<i>Estimate</i>	<i>p-value</i>
Parameter estimates						
Acquisition effect						
Acq1A	-0.1370	.0000	0.1254	.0000	0.0680	.0000
Acq1B	0.0114	.8310	0.0910	.0610	0.0061	.8420
Acq2	0.0933	.0130	-0.0520	.1500	0.0565	.0060
Age						
Age 21-22	-		-		Omitted	
Age 23-30	-		-		0.0877	.0000
Age 30-45	-		-		-0.1423	.0000
Age 45-60	Omitted		Omitted		-0.4003	.0000
Age 60+	0.0140	.1200	-0.0618	.0000	-0.6998	.0000
Female	-		-0.0594	.0000	-	
Prior-year health status						
No valid diagnosis	-0.1857	.0000	0.0009	.9600	-0.0867	.0000
Healthy user	Omitted		Omitted		Omitted	
Low risk	0.0137	.4250	-0.0542	.0010	-0.0560	.0000
Moderate risk	0.1175	.0000	-0.0705	.0000	-0.1336	.0000
High risk	0.0865	.0000	-0.0613	.0000	-0.1889	.0000
Very high risk	-0.0744	.0000	-0.0517	.0040	-0.3762	.0000
Neighborhood effects						
% White, non-Hispanic	0.0001	.9330	-0.0010	.0340	-0.0035	.0000
% High school only	0.0074	.0000	0.0006	.6130	0.0069	.0000
% Bachelors or higher	0.0092	.0000	0.0014	.1750	0.0130	.0000
% Income below FPL	-0.0032	.0000	-0.0005	.4410	-0.0029	.0000
% Speaking English only	-0.0014	.1150	0.0012	.0910	-0.0054	.0000
Product						
Broad PPO	Omitted		Omitted		Omitted	
Narrow network	0.0119	.3020	0.0193	.0230	0.0054	.3690
Medicaid	-0.4576	.0000	-0.0742	.0000	-0.2173	.0000
Attributed system						
Acq 1A	0.1202	.0000	-0.0509	.0490	-0.0129	.3680
Acq 1B	-0.0929	.0300	-0.0451	.2540	-0.0194	.4360
Acq 2	-0.1852	.0000	0.1123	.0000	-0.0069	.7000
IDS 1	Omitted		Omitted		Omitted	
IDS 2	-0.1085	.0000	0.0636	.0000	-0.1983	.0000

Continued

Table 2. Continued

	<i>BC Screening</i>		<i>CRC Screening</i>		<i>CC Screening</i>	
	<i>Estimate</i>	<i>p-value</i>	<i>Estimate</i>	<i>p-value</i>	<i>Estimate</i>	<i>p-value</i>
IDS 3	0.0765	.0000	0.1472	.0000	0.0242	.0000
IDS 4	0.0922	.0000	-0.0862	.0000	0.0960	.0000
IDS 5	-0.2296	.0000	-0.0041	.8200	-0.0605	.0000
PC 1	-0.2979	.0000	-0.0448	.0140	-0.1443	.0000
PC 2	0.1720	.0000	0.0047	.8480	0.2227	.0000
MS 1	-0.0076	.8090	0.0527	.0280	-0.3118	.0000
MS 2	-0.0905	.0040	-0.0705	.0040	0.2341	.0000
Calendar year						
2006	Omitted		Omitted		Omitted	
2007	0.0773	.0000	0.0451	.0000	0.0014	.8260
2008	0.0309	.0080	0.1254	.0000	-0.0296	.0000
2009	0.0786	.0000	0.1448	.0000	-0.0342	.0000
2010	0.0167	.1640	0.0776	.0000	-0.1724	.0000
2011	0.0456	.0000	0.0956	.0000	-0.3280	.0000
Constant	-0.4279	.0010	-1.3590	.0000	0.3151	.0000
Acquisition marginal effects						
Acq 1A	-0.0525	.0000	0.0245	.0000	0.0257	.0000
Acq 1B	0.0043	.8310	0.0174	.0760	0.0023	.8420
Acq 2	0.0348	.0110	-0.0091	.1370	0.0213	.0060
N	133,738		246,249		435,874	
Avq indicator	0.5945		0.1059		0.4555	

was a significantly lower probability of breast cancer screening among Medicaid enrollees.

Colorectal Cancer Screening. Acquisition was associated with an increase in the probability of colorectal cancer screening (Table 2) in Acq1A (ME = +2.5 pp) and a marginally significant increase in Acq1B (ME = +1.7 pp). These percentage-point increases (about 20 percent) are large relative to the 10.6 percent average annual probability of a colonoscopy. Note that both Acq1A and Acq1B were acquired by the same integrated delivery system (IDS1). These increased probabilities of colorectal cancer screening after acquisition are consistent with an acquisition-related improvement in quality of care.

Enrollees with higher health risk were less likely to have colorectal cancer screening than low-risk enrollees, as were the oldest enrollees and women.

There was a slight increase in the screening probability in the narrow-network product and a decreased probability for Medicaid enrollees. Screening probabilities declined in neighborhoods with a greater fraction of white and non-Hispanic residents. There was a general, though not monotonic, increase in screening probabilities over time.

Cervical Cancer Screening. The probability of cervical cancer screening (Table 2) increased after acquisition for Acq1A (ME = +2.6 pp) and Acq2 (ME = +2.1 pp), indicating an improvement in this quality indicator associated with the acquisition. These changes are approximately 5 percent of the average 45.6 percent probability an eligible woman received cervical cancer screening in any 1 year. Cervical cancer screening probabilities decreased with age and illness burden. Screening probabilities increased in more highly educated neighborhoods and decreased in neighborhoods with a higher fraction of white and non-Hispanic, and a higher fraction speaking English only. Screening probabilities declined for residents of lower income neighborhoods, consistent with the decline for Medicaid enrollees. There was a steady decline over time in cervical cancer screening probabilities, reflecting diffusion of a change in ICSI guidelines to screen every 3 years, rather than annually.

Appropriate Health Care Use

Ambulatory Care–Sensitive Admissions. Given the occurrence of an admission, acquisition was estimated to increase the probability an admission was ACS (Table 3) at Acq2 (ME = +4.0 pp), relative to the overall 10.6 percent ACS probability. Interestingly, the decline in this quality measure occurred in the acquisition that caused the greatest disruption in inpatient admission patterns, suggesting that stability of hospital–clinic relations may be important in preventing ACS admissions.

The conditional probability an admission was ACS increased with age and illness burden, and for Medicaid enrollees. Women and those living in highly educated neighborhoods were less likely to have an ACS admission. The probability an admission was ACS declined over time, particularly in the most recent years.

Readmissions. Conditional on initial admission, the probability of readmission within 30 days (Table 3) declined at Acq1A (ME = –1.6 pp), an improved

Table 3: Appropriate Use Regressions: Parameter Estimates and Acquisition Marginal Effects

	<i>ACS\Admission</i>		<i>Future Readmit\Admission</i>		<i>ED Appropriate</i>	
	<i>Estimate</i>	<i>p-value</i>	<i>Estimate</i>	<i>p-value</i>	<i>Estimate</i>	<i>p-value</i>
Parameter estimates						
Acquisition effect						
Acq1A	-0.0096	.8740	-0.1143	.1120	0.0163	.7240
Acq1B	-0.0212	.8480	0.0207	.8530	0.1192	.0840
Acq2	0.2054	.0040	0.0534	.4730	0.1286	.0080
Age						
Age 18–22	Omitted		Omitted		Omitted	
Age 23–30	-0.0475	.2350	-0.1080	.0030	0.0717	.0000
Age 30–45	0.1942	.0000	0.0732	.0260	0.2170	.0000
Age 45–60	0.3872	.0000	0.1196	.0000	0.3128	.0000
Age 60+	0.5742	.0000	-0.1448	.0000	0.3059	.0000
Female	-0.1336	.0000	-0.2218	.0000	-0.0200	.0920
Prior-year health status						
No valid diagnosis	0.1367	.0010	0.1544	.0000	0.0344	.2000
Healthy user	Omitted		Omitted		Omitted	
Low risk	0.0252	.5100	0.0821	.0420	-0.0234	.3620
Moderate risk	0.0168	.5700	0.0899	.0050	0.0686	.0000
High risk	0.1084	.0010	0.2068	.0000	0.1531	.0000
Very high risk	0.4275	.0000	0.4907	.0000	0.3094	.0000
Neighborhood effects						
% White, non-Hispanic	0.0001	.9450	0.0009	.2940	-0.0003	.5220
% High school only	-0.0046	.0390	-0.0043	.0600	0.0047	.0020
% Bachelors or higher	-0.0049	.0110	-0.0038	.0480	0.0049	.0000
% Income below FPL	0.0007	.5780	0.0017	.1850	-0.0014	.0850
% Speaking English only	0.0024	.0520	0.0029	.0320	-0.0024	.0030
Product						
Broad PPO	Omitted		Omitted		Omitted	
Narrow network	-0.0433	.0550	-0.1038	.0000	0.0762	.0000
Medicaid	0.1382	.0000	-0.0387	.0580	0.0040	.7520
Attributed system						
Acq 1A	0.0225	.6370	0.0100	.8660	-0.0589	.1140
Acq 1B	0.1480	.0690	-0.0889	.2950	-0.1978	.0000
Acq 2	0.0387	.4890	-0.0359	.5260	-0.0685	.0960
IDS 1	Omitted		Omitted		Omitted	

Continued

Table 3. Continued

	<i>ACS\Admission</i>		<i>Future Readmit\Admission</i>		<i>ED Appropriate</i>	
	<i>Estimate</i>	<i>p-value</i>	<i>Estimate</i>	<i>p-value</i>	<i>Estimate</i>	<i>p-value</i>
IDS 2	0.0144	.5350	-0.0951	.0000	-0.2142	.0000
IDS 3	0.0392	.0780	-0.0313	.1780	-0.0915	.0000
IDS 4	0.0284	.2860	0.0408	.1440	0.0222	.2670
IDS 5	0.0595	.1320	0.0026	.9430	-0.1777	.0000
PC 1	-0.0233	.6350	-0.0308	.5720	-0.1200	.0010
PC 2	-0.2076	.0130	-0.2818	.0060	-0.1886	.0020
MS 1	0.1579	.0000	-0.0122	.7220	-0.2481	.0000
MS 2	0.0254	.6940	-0.1492	.0190	-0.1207	.0080
Calendar year						
2006	Omitted		Omitted		Omitted	
2007	0.0098	.6400	-0.0258	.2830	-0.0196	.2050
2008	0.0130	.6120	0.0078	.7800	-0.0847	.0000
2009	-0.0305	.1810	0.0272	.2650	0.0224	.1740
2010	-0.0473	.0440	0.0053	.8340	0.0680	.0000
2011	-0.0643	.0060	0.0127	.6160	0.0620	.0000
Constant	-1.4164	.0000	-1.3667	.0000	0.8114	.0000
Acquisition marginal effects						
Acq1A	-0.0017	.8730	-0.0163	.0870	0.0031	.7220
Acq1B	-0.0036	.8460	0.0032	.8550	0.0212	.0630
Acq2	0.0401	.0090	0.0085	.4870	0.0227	.0040
<i>N</i>	85,286		89,180		177,196	
Avg indicator	0.1063		0.0865		0.8859	

quality indicator in that clinic system. However, as discussed below in the Robustness of Results section, this may be a false-positive result. The conditional probability of readmission peaked in middle age and increased with illness burden. Women were less likely to have a readmission, as were those in narrow-network products.

Appropriateness of Emergency Department Visit. For those who had an ED visit (excluding visits due to injury, psychiatric diagnoses, or alcohol/substance abuse), we modeled the probability that the ED visit was appropriate (Table 3). Appropriate ED use was more likely after acquisition in Acq2 (ME = +2.3 pp) and Acq1B (marginally significant ME = +2.1 pp), consistent with improvements in quality of care. While these increases are small, we believe they are significant given that our generous definition of “appropriate” resulted in an average 88.6 percent of ED visits deemed

appropriate, leaving little room for improvement. The probability that ED use was appropriate increased with age and illness burden. Appropriate use was more likely in highly educated neighborhoods and less likely in neighborhoods with a greater fraction speaking English only. Appropriate use was more likely for those enrolled in a narrow-network plan. Appropriate ED use increased in recent years.

General Utilization Levels

Inpatient Admissions. The probability of an inpatient admission (Table 4) declined in Acq1A after the acquisition (ME = -0.7 pp), relative to the 8.4 percent overall average probability of an inpatient admission. While this decline in the probability of an inpatient admission suggests improved quality of care, these results were not robust in subanalyses discussed in the Robustness of Results section below.

The probability of having one or more inpatient admissions in a year varies by age in a nonmonotonic pattern, with the highest probabilities at ages 23–30 (child-bearing years) and over age 60. Women were more likely to have an admission, as were those with a higher illness burden. Admission probabilities declined in neighborhoods with higher fractions of white and non-Hispanic residents and more highly educated neighborhoods. Admission probabilities increased in lower income neighborhoods and neighborhoods that had a higher fraction speaking English only. Those covered by narrow-network plans were less likely and Medicaid enrollees more likely to have an admission. Admission probabilities declined steadily over time, with a possibly recession-related large drop in 2008.

Emergency Department Visits. The impact of acquisition on the probability of one or more ED visits (Table 4) varied by acquired system, decreasing in Acq1A (ME = -1.0 pp) and increasing in Acq2 (ME = $+1.2$ pp). As discussed below, acquisition-related changes in ED utilization were not robust.

The probability of an ED visit declined with age, but increased with illness burden. Women and those in highly educated neighborhoods were less likely to have an ED visit. Probabilities increased in poorer neighborhoods and neighborhoods with a greater fraction speaking English only. There was a slight decline in ED visit probability for those in narrow-network products and a large increase for Medicaid enrollees. ED use declined by a small amount in the most recent years.

Table 4: General Utilization Regressions: Parameter Estimates and Acquisition Marginal Effects

	<i>1 + IP Admits</i>		<i>1 + ED Visits</i>	
	<i>Estimate</i>	<i>p-value</i>	<i>Estimate</i>	<i>p-value</i>
Parameter estimates				
Acquisition effect				
Acq1A	-0.0464	.0130	-0.0456	.0040
Acq1B	-0.0384	.2620	0.0277	.2940
Acq2	-0.0158	.4740	0.0506	.0060
Age				
Age 18-22	Omitted		Omitted	
Age 23-30	0.1327	.0000	-0.0934	.0000
Age 30-45	-0.0661	.0000	-0.1629	.0000
Age 45-60	-0.1268	.0000	-0.1931	.0000
Age 60+	0.0706	.0000	-0.1915	.0000
Female	0.1918	.0000	-0.0337	.0000
Prior-year health status				
No valid diagnosis	0.0964	.0000	0.0727	.0000
Healthy user	Omitted		Omitted	
Low risk	-0.0618	.0000	0.0469	.0000
Moderate risk	0.2208	.0000	0.2282	.0000
High risk	0.5698	.0000	0.5235	.0000
Very high risk	1.0097	.0000	0.9065	.0000
Neighborhood effects				
% White, non-Hispanic	-0.0010	.0000	-0.0006	.0220
% High school only	-0.0017	.0210	-0.0042	.0000
% Bachelors or higher	-0.0048	.0000	-0.0070	.0000
% Income below FPL	0.0010	.0120	0.0035	.0000
% Speaking English only	0.0014	.0010	0.0037	.0000
Product				
Broad PPO	Omitted		Omitted	
Narrow network	-0.0665	.0000	-0.0186	.0010
Medicaid	0.2767	.0000	0.5650	.0000
Attributed system				
Acq 1A	-0.0121	.4190	-0.0042	.7480
Acq 1B	-0.1435	.0000	-0.0469	.0280
Acq 2	-0.0690	.0000	-0.0308	.0470
IDS 1	Omitted		Omitted	
IDS 2	-0.0507	.0000	0.0686	.0000
IDS 3	0.0253	.0000	-0.0260	.0000
IDS 4	0.0557	.0000	0.0299	.0000
IDS 5	0.0325	.0050	0.0110	.2870
PC 1	-0.1456	.0000	-0.1238	.0000
PC 2	-0.2602	.0000	-0.0890	.0000

Continued

Table 4. *Continued*

	<i>1 + IPAdmits</i>		<i>1 + ED Visits</i>	
	<i>Estimate</i>	<i>p-value</i>	<i>Estimate</i>	<i>p-value</i>
MS 1	0.3220	.0000	0.3092	.0000
MS 2	-0.0375	.0250	-0.0290	.0430
Calendar year				
2006	Omitted		Omitted	
2007	-0.0108	.1170	-0.0065	.2570
2008	-0.1704	.0000	0.0020	.7310
2009	-0.0470	.0000	-0.0200	.0010
2010	-0.0816	.0000	-0.0500	.0000
2011	-0.1131	.0000	-0.0201	.0010
Constant	-1.4938	.0000	-0.9901	.0000
Acquisition marginal effects				
Acq1A	-0.0065	.0110	-0.0102	.0040
Acq1B	-0.0054	.2500	0.0064	.3000
Acq2	-0.0022	.4690	0.0118	.0070
N	796,962		796,962	
Avg indicator	0.0842		0.1628	

Robustness of Results

Our assessment of robustness of results focused on the following issues: (1) accumulation of Type 1 error; (2) appropriateness of the control groups; (3) testing placebo interventions to assess the strength of the differences-in-differences framework; (4) falsification tests to assess the strength of the differences-in-differences framework; and (5) the impact of the unbalanced nature of our panel data. Each of these is discussed below. Table 5 summarizes the baseline marginal effects discussed above, the tests for accumulation of Type 1 error, and provides representative examples of the control group and placebo treatment group testing.

Accumulation of Type 1 Error. Because we are estimating multiple equations, it is possible that the positive results are due to chance. We define statistically significant results as $p < .05$ and marginally significant results as $.05 < p < .10$. Testing the statistical significance of three treatment effects in each of eight regressions means, we are testing a total of $n = 24$ hypotheses. A simple way to adjust for multiple tests is to use a Bonferroni correction, basing critical values on $p = .05/24$ rather than $p = .05$.

Table 5: Summary of Results and Robustness Testing

	BC Screening		CRC Screening		CC Screening		ACS/Admission	
	Marg Eff	p-value	Marg Eff	p-value	Marg Eff	p-value	Marg Eff	p-value
Baseline marginal effects— δP (outcome)								
Acq1A	-0.0525	.0000	0.0245	.0000	0.0257	.0000	-0.0017	.8730
Acq1B	0.0043	.8310	0.0174	.0760	0.0023	.8420	-0.0036	.8460
Acq2	0.0348	.0110	-0.0091	.1370	0.0213	.0060	0.0401	.0090
Outcome	Mixed		Favorable		Favorable		Negative	
Passed Bonferroni Test?	Yes		Yes		Yes		Yes	
Testing alternate control groups								
Compare only to those not vertically integrated (MS1&2, PC1&2)	Consistent results		Consistent results; loss of precision		Consistent results; large loss of precision		Consistent results; loss of precision	
Compare only to those not vertically integrated and multispecialty (MS1&2)	Consistent results		Consistent results; large loss of precision		Consistent results; loss of precision		Consistent results; loss of precision	
Compare only to those always vertically integrated (IDS1-5)	Consistent results		Consistent results		Consistent results		Consistent results	
Compare only to those always vertically integrated and involved in acquisitions (IDS1&2)	Consistent results		Consistent results		Consistent results		Consistent results	
Compare only to IDS1 (always vertically integrated and involved in acquisition)	Consistent results; loss of precision		Consistent results		Consistent results		Consistent results	
Compare only to IDS2 (always vertically integrated and involved in acquisitions)	Consistent results		Consistent results		Opposite results		Consistent results	

Continued

Table 5. Continued

	BC Screening		CRC Screening		CC Screening		ACS/Admission	
	Marg Eff	p-value	Marg Eff	p-value	Marg Eff	p-value	Marg Eff	p-value
Testing placebo treatment effects								
IDS1 as acq; IDS2&3 as control	Incr prob		Not stat sig by Bonferroni test		Not stat sig by Bonferroni test		Not stat sig by Bonferroni test	
PC1 as acq; MS1 as control	Not stat sig by Bonferroni test		Not stat sig by Bonferroni test		Not stat sig by Bonferroni test		Not stat sig by Bonferroni test	
MS1 as acq; IDS5 as control	Incr prob		Not stat sig by Bonferroni test		Not stat sig by Bonferroni test		Not stat sig by Bonferroni test	
	<i>Future Readmit/Admission</i>		<i>ED Appropriate</i>		<i>1 + IP Admits</i>		<i>1 + ED Visits</i>	
	Marg Eff	p-value	Marg Eff	p-value	Marg Eff	p-value	Marg Eff	p-value
Baseline marginal effects—DP (outcome)								
Acq1A	-0.0163	.0870	0.0031	.7220	-0.0065	.0110	-0.0102	.0040
Acq1B	0.0032	.8550	0.0212	.0630	-0.0054	.2500	0.0064	.3000
Acq2	0.0085	.4870	0.0227	.0040	-0.0022	.4690	0.0118	.0070
Outcome	None		Favorable		Favorable		Mixed	
Passed Bonferroni Test?	No		Yes		No		Yes	
Testing alternate control groups								
Compare only to those not vertically integrated (MS1&2, PC1&2)	N/A		Consistent results; increased precision		N/A		Consistent results; loss of precision	
Compare only to those not vertically integrated and multispecialty (MS1&2)	N/A		Consistent results; increased precision		N/A		Consistent results for Acq1A; increased precision	

Continued

Table 5. Continued

	Future Readmit/ Admission		ED Appropriate		1 + IP Admits		1 + ED Visits	
	Marg Eff	p-value	Marg Eff	p-value	Marg Eff	p-value	Marg Eff	p-value
Compare only to those always vertically integrated (IDS1-5)	N/A		Consistent results; loss of precision		N/A		Consistent results	
Compare only to those always vertically integrated and involved in acquisitions (IDS1&2)	N/A		Consistent results; loss of precision		N/A		Consistent results	
Compare only to IDS1 (always vertically integrated and involved in acquisition)	N/A		Consistent results		N/A		Consistent results	
Compare only to IDS2 (always vertically integrated and involved in acquisitions)	N/A		Consistent results		N/A		Consistent results	
Testing placebo treatment effects IDS1 as acq; IDS2&3 as control	N/A		Not stat sig by Bonferroni test		N/A		Decr prob	
PC1 as acq; MS1 as control	N/A		Not stat sig by Bonferroni test		N/A		Not stat sig by Bonferroni test	
MS1 as acq; IDS5 as control	N/A		Not stat sig by Bonferroni test		N/A		Decr prob	

However, when the number of hypotheses gets large, this correction creates a standard so conservative that it substitutes the risk of Type 2 error (false negatives) for the risk of Type 1 error (false positives). To balance these two risks, we applied the Bonferroni correction to a likelihood ratio test of the null hypothesis “ H_0 : All three treatment parameters are equal to zero.” This reduced the number of hypotheses tested to eight, leading to critical values based on $p = .05/8 = .0063$. Of course, this hypothesis tests for the presence of any treatment effect, rather than the system-specific treatment effects discussed, but it provides some assurance that accumulation of Type 1 error is not leading to false-positive results. Using this standard, the acquisitions did not have a statistically significant impact on inpatient admission rates or readmission rates. Because of this, we retain only the remaining six regressions in our continued robustness testing.

Appropriateness of Control Groups. It is possible that we did not use the appropriate counterfactual for acquisitions in our choice of controls. In particular, because classic control-group design compares an intervention group with a group that did not receive the intervention, observations for enrollees in the “always acquired” clinics (IDS1—IDS5) may be inappropriate comparators. Therefore, we reestimated the six remaining equations using a variety of alternative control groups. These include comparing the acquired clinics to the “never acquired” controls (MS1, MS2, PC1, PC2) in total and just the never acquired multispecialty groups (MS1, MS2), in addition to several permutations of “always acquired” controls. Though the reduced sample size often resulted in loss of precision, results of this test were generally very consistent with the results discussed above.

Placebo Interventions. We selected clinic systems that were not acquired, treating them as if they had been acquired. To increase the stability of the trends against which the placebo acquisition effect is measured, we selected the control groups that were at least as large as the number of enrollees attributed the placebo treatment group. We display representative examples of an IDS against other IDSs (IDS1 vs. IDS2 and IDS3), of an independent clinic system against another independent clinic system neither of which was acquired (PC1 vs. MS1), and an independent clinic system that was not acquired against an IDS that was not involved in an acquisition

(MS1 vs. IDS5). These tests throw suspicion on the robustness of the regressions modeling breast cancer screening and general ED utilization levels, but the other results remain robust.

Falsification Tests. An additional test of the strength of our differences-in-differences model is to estimate the impact of acquisition on quality measures that should not be affected by the acquisition. Falsification tests should focus on quality measures that are unlikely to be impacted by improved clinic infrastructure or improved communication across inpatient and outpatient care settings. One often-cited measure of hospital quality is inpatient mortality, a measure that we find is unrelated to the acquisitions. However, this negative finding may be due to lack of statistical power. Another inpatient-only measure with higher prevalence is maternity admissions. Reduced probability that a maternity admission results in a Cesarean section is a good candidate for a quality measure that should not be related to the acquisitions. Indeed, we found no connection between the acquisitions and C-section rates.

Other Tests. We also checked the impact of our unbalanced panel, confirming results were consistent in a subset of observations that have a minimum of 12 months of exposure before and 12 months after the acquisitions. We also confirmed that acquisition effects were not a continuation of a preacquisition difference in trends.

CONCLUSION

We examined changes in quality of care measures in three large, multispecialty clinics that were acquired by two hospital-owned IDSs in the Minneapolis–St. Paul area. We compared changes in quality indicators for the acquired clinics to nine control groups, using a differences-in-differences model. While the acquisition effects were small and, at least in these early postacquisition years, limited to cancer screening and appropriateness of ED use, our results suggest that integration of a clinic system into an IDS has the potential to improve quality of care. However, we also found an increased probability of ACS admissions when the acquisition caused disruption to existing physician–hospital admitting relations.

ACKNOWLEDGMENTS

Joint Acknowledgment/Disclosure Statement: The authors gratefully acknowledge funding for this work from The Commonwealth Fund (grant number 20120304), and to their staff for helpful editorial suggestions. We also appreciate the perspective provided by ICSI staff and Dr. John Wilkinson, Mayo Clinic, on historical changes in ICSI guidelines for adult preventive services. Finally, we appreciate excellent suggestions from the editors and two anonymous reviewers.

Disclosures: None.

Disclaimers: None.

NOTES

1. The acquired clinic systems studied here made new EMR investments, but only one clinic system shared the timing of that investment: within 18 months of acquisition.
2. That is, medical groups that work together and are jointly owned.
3. This consortium does not include joint ownership.
4. No significant clinic acquisitions occurred in the area from 2006 to 2011 except the three we study.
5. We used prior-year health status to avoid issues of endogeneity, restricting our population to enrollees who had at least 2 years of enrollment. This restriction has little impact on the summary statistics for the population. The only meaningful difference is a modest increase in average age (41.6–42.1 years).
6. Note that our claims-based indicators of quality do not include clinical measures of quality available in medical records.

REFERENCES

- Billings, J. C. 2013a. "ACS Codes" [accessed on January 18, 2013]. Available at <http://wagner.nyu.edu/faculty/billings/acs-algorithm>
- . 2013b. "ED Utilization Background" [accessed on January 18, 2013]. Available at <http://wagner.nyu.edu/faculty/billings/nyued-background>
- Burns, L. R., and R. W. Muller. 2008. "Hospital-Physician Collaboration: Landscape of Economic Integration and Impact on Clinical Integration." *Milbank Quarterly* 86 (3): 375–434.
- Christianson, J., C. Carlin, and L. H. Warrick. 2014. "The Dynamics of Health Care Consolidation: Acquisition of Physician Practices." *Milbank Quarterly* 92 (3): 542–67.

- Dowd, B., M. Karmarker, T. Swenson, S. Parashuram, R. Kane, R. Coulam, and M. M. Jeffery. 2014. "Emergency Department Utilization as a Measure of Physician Performance." *American Journal of Medical Quality* 29 (2): 135–43.
- Gal-Or, E. 1999. "The Profitability of Vertical Mergers between Hospitals and Physician Practices." *Journal of Health Economics* 18 (5): 623–54.
- Hwang, W., J. Chang, M. Laclair, and H. Paz. 2013. "Effects of Integrated Delivery System on Cost and Quality." *American Journal of Managed Care* 19 (5): e175–84.
- ICSI. 2013. "Preventive Services for Adults" [accessed on November 7, 2013]. Available at https://www.icsi.org/guidelines__more/catalog_guidelines_and_more/catalog_guidelines/catalog_prevention__screening_guidelines/preventive_services_for_adults/
- Kocher, R., and N. R. Sahni. 2011. "Hospitals' Race to Employ Physicians—The Logic behind a Money-Losing Proposition." *New England Journal of Medicine* 364 (19): 1790–3.
- Mehrotra, A., A. M. Epstein, and M. B. Rosenthal. 2006. "Do Integrated Medical Groups Provide Higher-Quality Medical Care Than Individual Practice Associations?" *Annals of Internal Medicine* 145 (11): 826–33.
- Mukamel, D. B., and A. I. Mushlin. 1998. "Quality of Care Information Makes a Difference: An Analysis of Market Share and Price Changes after Publication of the New York State Cardiac Surgery Mortality Reports." *Medical Care* 36 (7): 945–54.
- Rittenhouse, D. R., S. M. Shortell, R. R. Gillies, L. P. Casalino, J. C. Robinson, R. K. McCurdy, and J. Siddique. 2010. "Improving Chronic Illness Care: Findings from a National Study of Care Management Processes in Large Physician Practices." *Medical Care Research and Review: MCRR* 67 (3): 301–20.
- Robert Wood Johnson Foundation. 2013. "Comparing Health Care Quality: A National Directory" [accessed on November 4, 2013]. Available at <http://www.rwjf.org/en/research-publications/find-rwjf-research/2013/09/national-directory.html>
- Shortell, S. M., J. Schmittiel, M. C. Wang, R. Li, R. R. Gillies, L. P. Casalino, T. Bodenheimer, and T. G. Rundall. 2005. "An Empirical Assessment of High-Performing Medical Groups: Results from a National Study." *Medical Care Research and Review: MCRR* 62 (4): 407–34.
- Trybou, J., P. Gemmel, and L. Annemans. 2011. "The Ties That Bind: An Integrative Framework of Physician-Hospital Alignment." *BMC Health Services Research* 11: 36.
- U.S. Census Bureau. 2012. "American Community Survey" [accessed on October 23, 2012]. Available at <http://www.census.gov/acs/www/>
- Weeks, W. B., D. J. Gottlieb, D. E. Nyweide, J. M. Sutherland, J. Bynum, L. P. Casalino, R. R. Gillies, S. M. Shortell, and E. S. Fisher. 2010. "Higher Health Care Quality and Bigger Savings Found at Large Multispecialty Medical Groups." *Health Affairs (Millwood)* 29 (5): 991–7.

Weiner, J. P., B. H. Starfield, and R. N. Lieberman. 1992. "Johns Hopkins Ambulatory Care Groups (ACGs). A Case-Mix System for UR, QA and Capitation Adjustment." *HMO Practice* 6 (1): 13–9.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article:

Appendix SA1: Author Matrix.