Green House Adoption and Nursing Home Quality

Christopher C. Afendulis, Daryl J. Caudry, A. James O'Malley, Peter Kemper, and David C. Grabowski, for the THRIVE Research Collaborative^a

Objective. To evaluate the impact of the Green House (GH) model on nursing home resident-level quality of care measures.

Data Sources/Study Setting. Resident-level minimum data set (MDS) assessments merged with Medicare inpatient claims for the period 2005 through 2010.

Study Design. Using a difference-in-differences framework, we compared changes in care quality and outcomes in 15 nursing homes that adopted the GH model relative to changes over the same time period in 223 matched nursing homes that had not adopted the GH model.

Principal Findings. For individuals residing in GH homes, adoption of the model lowered readmissions and several MDS measures of poor quality, including bedfast residents, catheter use, and pressure ulcers, but these results were not present across the entire GH organization, suggesting possible offsetting effects for residents of non-GH “legacy” units within the GH organization.

Conclusions. GH adoption led to improvement in rehospitalizations and certain nursing home quality measures for individuals residing in a GH home. The absence of evidence of a decline in other clinical quality measures in GH nursing homes should reassure anyone concerned that GH might have sacrificed clinical quality for improved quality of life.

Key Words. Nursing homes, quality of care, Green House nursing home, culture change

The quality of care delivered to frail elders residing in nursing homes remains an important and perplexing issue in American health policy (Institute of Medicine 2001; U.S. Government Accountability Office 2005; Koren 2010). One innovation for improving quality in nursing homes is the Green House (GH) model, a culture change initiative that prescribes changes to the living and social environment and aims to increase the level of resident and staff

autonomy compared with traditional nursing homes. The GH model involves the creation of small buildings, designed to serve a maximum of 12 residents, which fit the style of the surrounding neighborhood. Its goal is to provide a more homelike environment that facilitates resident direction and maintains the dignity and independence of residents while providing a comparable level of skilled care. The GH organizational structure is intended to be much less hierarchical than traditional nursing homes, and more control over daily activities is given to residents and the *Shahbazim* (which is the GH term for the direct care staff in GH homes, similar to certified nursing assistants [CNAs] in traditional nursing homes).

Although a key emphasis of the GH model is to improve the quality of life for residents (Kane et al. 2007), the effect of the model on traditional quality of care measures is unclear. The objective of this study was to examine the relationship between adoption of the GH model and different minimum data set (MDS) quality and hospitalization measures. A number of providers have adopted GH and other culture change models over the past two decades, often supported by different government policies (Grabowski et al. 2014a). This study is the first to examine the longitudinal impact of the adoption of the GH model on quality of care and selected outcomes, which will provide important direction for potential adopters and government funding of this model.

BACKGROUND

Previous Literature

The GH model is part of the broader culture change movement in the nursing home sector (Koren 2010). Culture change encompasses a series of innovative care models that reconceptualize the structure, roles, and processes of nursing home care to transform nursing homes from health care institutions to person-centered homes offering long-term care services. Key elements of culture

Address correspondence to David C. Grabowski, Ph.D., Department of Health Care Policy, Harvard Medical School, 180 Longwood Avenue, Boston, MA 02115; e-mail: grabowski@med.harvard.edu. Christopher C. Afendulis, Ph.D., and Daryl J. Caudry, M.A., are with the Department of Health Care Policy, Harvard Medical School, Boston, MA. A. James O'Malley, Ph.D., is with the Geisel School of Medicine, The Dartmouth Institute, Dartmouth College, Lebanon, NH. Peter Kemper, Ph.D., is with The Pennsylvania State University, University Park, PA.
^a The THRIVE Research Collaborative includes Barbara Bowers, Patrick Brown, Lauren Cohen, David Grabowski, Susan Horn, Sandy Hudak, Kimberly Nolet, David Reed, and Sheryl Zimmerman.

change nursing homes can include resident direction, homelike atmosphere, close relationships, staff empowerment, collaborative decision making, and quality-improvement processes (Koren 2010). A relatively large health services literature has examined the association between culture change and quality of care. In a recent review of this literature, Shier et al. (2014) concluded that culture change was not generally associated with better or worse quality of care. The authors noted that the literature to date has been limited by small sample sizes, inconsistent measurement of culture change, and weak study designs. Grabowski et al. (2014b) addressed these concerns by using difference-in-difference panel regression methods to examine the effect of adopting culture change on the quality of care. They found that nursing homes that were identified as culture change adopters had 14.6 percent fewer health-related survey deficiency citations than comparable nonadopting nursing homes, but they found no significant change in nurse staffing or MDS quality measures. In another recent study, Miller et al. (2014) found that nursing homes with a high culture change practice implementation score experienced a significant decrease in the prevalence of restraints, tube feeding, and pressure ulcers; an increase in the proportion of residents on bladder training programs; and a small decrease in the average number of hospitalizations per resident year. This study also found that nursing homes with a lower practice implementation score experienced fewer deficiency citations.

The research on quality of care in GH homes in particular is somewhat sparse. In an early cross-sectional evaluation of the program, Kane et al. (2007) compared chart-based MDS outcomes in four GH homes relative to both other individuals residing in the traditional part of that same nursing home organization (known as the “legacy home”) and individuals in another local traditional nursing home under the same ownership. Relative to both comparison groups, few of the MDS outcomes were statistically significant, although those that were generally suggested better quality in the GH homes. Importantly, this study also found that self-reported dimensions of quality of life were better in the GH homes relative to one of the comparison groups or for some measures, both comparison groups. Using this same sample and study design, family members were also more satisfied with care under the GH model relative to the comparison groups (Lum et al. 2008). Important limitations associated with this previous GH evaluation included the cross-sectional study design, the possible introduction of selection bias due to the comparisons of GH homes with legacy units within the same GH organization, the small sample size, and the reliance of data from only a single GH organization.

In a later study, Sharkey et al. (2011) addressed these last two limitations of the earlier evaluation by comparing components of care across 27 GH and traditional nursing home units. The authors found that CNAs in GH homes spent 0.4 more hours per resident day (24 minutes) in direct care activities relative to CNAs in traditional units. Although a number of research studies have concluded that more direct care contact can equate to better quality of care (Harrington et al. 2000), this study also suffers from potential biases arising from a cross-sectional design and comparisons within the same GH organization.

Conceptual Framework

The potential benefit of the GH model, which emphasizes resident-centered care to improve quality of life, depends on the complementarity between quality of care and quality of life in nursing home care. On the one hand, under the GH model they might be complements; for example, by striving to offer person-centered care, the Shahbazim may be more engaged in meeting resident care needs, and consequently avoidable incidents such as pressure ulcers or infections may be less likely to occur. On the other hand, by allowing greater resident autonomy, residents in GH nursing homes may experience a higher fall rate, increased weight loss, or other adverse events.

In this study, we hypothesized that the GH model would improve nursing home quality of care, assuming based on previous research that quality of life is stronger in the GH setting and theorizing that quality of care and quality of life are complements. However, we recognized the alternate hypothesis that the GH model may have a null or negative relationship with quality of care might hold. This might be the case if either (1) quality of life and quality of care are complements, but contrary to previous research, GH does not improve quality of life; or (2) quality of care and quality of life are substitutes—that is, better quality of life comes at the expense of quality of care. Finally, it is also possible that some aspects of quality of care may be improved, while others decline.

For GH adopters that also operate a “legacy home,” the original nursing home that stays open alongside the new GH home(s), it is also important to recognize that GH adoption may have both direct and indirect impacts on quality of care. The direct impact would refer to quality within the GH units, while the indirect impact would refer to quality within the legacy units. The overall GH impact on the organization is the combined direct and indirect effects. The indirect effect of GH adoption for the legacy unit could run in the

same or opposite direction as the direct effect for the GH units. By definition, the legacy unit cannot adopt the “small house” model, but other tenets of the GH culture such as resident empowerment and staffing philosophy can “spillover” to the legacy part of the organization. If so, we would expect the legacy impact to run in the same direction as the GH impact, although attenuated relative to the direct effect. Alternatively, if behaviors adopted in the GH unit “crowd out” organizational resources, then we would expect the legacy results to run in the opposite direction of the main GH unit findings. For this reason, it is important to examine quality in the GH and legacy unit and then the combined results for the organization as a whole.

Our Contribution

Our primary methodological contributions to this literature include a “difference-in-differences” study design, adjustment for important differences using exact matching techniques, and analysis of a larger sample of GH homes. Rather than relying on a cross-sectional design comparing GH and non-GH nursing homes at a single point in time, we examine change over time following the adoption of GH relative to similar nursing homes not adopting the GH model over this same time period. It is well known that a cross-sectional identification strategy leads to misleading inferences if unobserved factors that affect quality are correlated with GH adoption. For example, if high-resource nursing homes adopt the GH model (Grabowski et al. 2014a), then a cross-sectional model might provide an overestimate of the association between GH and the quality of care. Our strategy has the advantage of balancing the analytic sample on observed risk factors at baseline prior to the adopting GH and “differencing out” time-invariant unobservable risk factors and secular trends. This balance is achieved by examining pre-post differences in nursing homes adopting the GH model relative to pre-post differences in nonadopters for a matched sample of nursing homes.

A potential challenge associated with a “difference-in-differences” approach is reliance on the assumption that nursing homes that ultimately adopt the GH model would have otherwise followed similar trends in the outcome across time to the comparison group (nonadopters). Previous culture change research has often generated a comparison group that may be very different from the culture change group (Shier et al. 2014). For example, in the evaluation of quality in GH homes described above, the authors constructed a comparison group of individuals residing either in

the traditional unit of the nursing home or a neighboring nursing home with the same owner (Kane et al. 2007). The comparison group was older and had more disabilities, thereby potentially biasing estimates of the program effect. To address this issue, we implement an exact matching approach to construct our comparison group. Thus, our research compares the change in quality for a nursing home adopting GH relative to the change for a *similar* nursing home not adopting GH; the rationale is that homes that have similar characteristics at baseline will be the most likely to have had similar trends across time in the absence of the GH intervention. This is particularly important in this study because the majority of nursing homes that adopted GH in this study were nonprofit, faith-based, and part of a continuing care retirement community (Cohen et al. 2016). Also of note, the earlier evaluation of GH quality was based on a single nursing home in Mississippi. Our treatment group includes the adoption of GH in 15 organizations, thereby increasing sample size and the generalizability of our results.

DATA AND METHODS

Data

Our analyses used three different data files, all maintained by the Centers for Medicare and Medicaid Services (CMS): resident-level assessment data from the MDS; nursing home-level data from the Online Survey, Certification, and Reporting (OSCAR) file; and beneficiary-level Medicare enrollment and claims data. We merged these administrative data with nursing home-level data on GH adoption and individual-level data indicating whether residents lived in the GH organization's legacy units or its new GH homes. Ultimately, the analytic file contained 645,191 assessments from 131,794 unique residents receiving care from 238 nursing homes. We describe each of these data sources in greater detail below:

MDS Data. We used assessment-level data from the MDS 2.0 for the period January 1, 2005 through September 30, 2010 to determine clinical care outcomes using eight quality measures: bedfast; incontinence, low risk; catheterization; pain; physical restraints; pressure ulcers, high risk; pressure ulcers, low risk; and urinary tract infection. The incontinence (low risk) and pressure ulcer (high or low risk) measures were only

collected for a subset of individuals with a particular risk profile, meaning the sample sizes were smaller in the analyses of these outcomes. The data's end date allowed us to restrict the analysis to assessments performed using the MDS 2.0 instrument (i.e., MDS 3.0 was adopted as of October 1, 2010). The MDS was collected at admission and then at least quarterly thereafter. Using these data, we constructed a nursing home assessment-level dataset.

OSCAR Data. We used the OSCAR data in two ways. First, we used nursing home characteristics to find suitable non-GH matches for each GH organization, as described below. Second, in our assessment-level regression analyses, we included variables describing the percentage of residents in each nursing home whose payer was Medicaid (the omitted category), Medicare, or other (private-pay); these were used to capture the share of short-stay postacute residents (i.e., Medicare) and also the generosity of reimbursement associated with long-stay residents (i.e., private-pay residents are associated with greater revenue relative to Medicaid residents).

Medicare Enrollment and Claims Data. Using the resident's Social Security number and working through a CMS data contractor to ensure confidentiality, we merged the MDS file with Medicare enrollment data, using data from the month of nursing home admission. We dropped residents who were not entitled to Medicare Part A in the month of admission, because we will not have complete hospital claims data for these individuals. We also dropped assessments where the resident died during the month of admission, because Medicare hospitalization data for these cases will not be reflective of care received in the nursing home setting. We also excluded cases where the resident was enrolled in the Medicare Advantage program during the month of admission, because claims data were not available for them.

We created two set of measures from the merged claims file. First, we constructed a set of hospitalization and rehospitalization measures from the claims, as detailed below. Second, we used the merged data to assess whether the resident was entitled to Medicaid during the month of admission, using the "state buy-in" data available in the Medicare enrollment data. This measure is included as an independent variable in our regression models described below to capture dual eligibility status, serving as a proxy in the models for both socio-economic and payer status.

Green House Adopters and Resident Identifiers. We were provided a list of the exact date of GH adoption from The Green House® Project, a program administered by Capital Impact Partners that provides technical assistance to GH homes. Over the period 2005 through 2010, 18 organizations adopted new GH units. Given our estimation approach below, we required that the nursing home had been in operation prior to the adoption of The GH model. Ultimately, we identified 15 nursing homes that adopted The GH model over our study period and were in operation in the period preceding the adoption (two GHs adopted in 2006, five in 2007, three in 2008, two in 2008, and three in 2010). These adopters, which collectively built 72 GH homes, were located in Alabama, Arkansas, Kansas, Massachusetts, Michigan, Montana, Nebraska, New York, Pennsylvania, Tennessee, and Texas.

These 15 adopters predominantly consisted of organizations that built a GH home on an existing nursing home campus. In some instances, these nursing homes had to shut down existing licensed beds to build the GH beds under a state certificate-of-need law, while in other instances the nursing home could simply expand the number of beds. We included one GH “adopter” that established a new nursing home license for the GH homes, even though the traditional nursing home was located nearby. Thus, although this organization did not technically add the GH homes to an existing nursing home license, we concluded that this organization’s adoption of the GH model was consistent with the other adopters in our database.

As discussed below, we examined the overall impact of GH adoption in two ways. First, we examined the overall impact of GH adoption on all residents in the nursing home organization, including those in GH units and those in the legacy unit. This analysis used all 15 nursing homes, 10 of which included a mix of GH homes and legacy units, while the other 5 organizations had converted entirely to GH homes and had only GH residents. Second, we examined the impact specifically on those individuals residing in a GH home. Because we were not able to ascertain GH residence from the MDS, each of the participating GH homes was asked to provide data on the timing of GH residence for all individuals living at the nursing home. These identifiers were linked to the Medicare claims files by a CMS data contractor. Three of the nursing homes with a mix of GH and legacy residents did not provide resident identifiers and had to be excluded from this second set of analyses, leaving us with 12 GH adopters (5 with only GH residents and 7 with a mix of GH and legacy residents) for our analysis of Medicare utilization in GH homes and legacy units.

Finding Matched Controls

We used a two-stage process to construct our comparison group. First, because GH nursing homes likely differed in systematic ways from non-GH ones, we used facility-level matching methods to find controls for each GH nursing home. To find matches, we selected from the set of nursing homes that were in operation within the state in the year each GH came online. We excluded from the set of potential control nursing homes any of the nursing homes that eventually adopted the GH model. Then for each combination of state and year of GH adoption, we sought to find matched control nursing homes in that state and at that time period. All of the matches we found for each GH organization were included for the entire study period in the subsequent analysis.

Using the OSCAR data, we employed nearest neighbor matching at the nursing home level, which determined “nearest” based on the Mahalanobis distance, in which the weights were based on the inverse of the following 12 covariates’ variance-covariance matrix: nonprofit ownership, for-profit ownership, government ownership, chain status, small size (75 beds or fewer), medium size (76–125 beds), large size (126 or more beds), rural location, above median Medicaid share, above median Medicare share, above median private-pay share, and a nursing home-level aggregate activities of daily living (ADL) score (0 if less than 4 on a scale of 0–5, 1 otherwise). For our organization-level analysis, our approach yielded a total of 223 matched control nursing homes for our 15 GH treatment nursing homes in 11 states. In our analysis of residents living in GH and legacy units, we had a total of 178 matched control nursing homes for our 12 GH treatment nursing homes in 10 states. Reflecting the differences between many GH and the average nursing home, we found large differences in characteristics (e.g., for-profit status and high reliance on private payers) between GH nursing homes and other nursing homes in states where GH nursing homes are located. With one exception (rural location), the balance on all of the facility measures improved, often substantially (see Table 1). Although some meaningful differences remained after matching, the differences were not statistically different for any of the measures.

The second stage in constructing the comparison group consisted of propensity score weighting at the person level based on the inverse of the propensity score. For the organization-wide analysis with the 15 GH adopters, we calculated the conditional probability (propensity) of being in a GH organization using a logistic regression model. For the unit-level analysis with 12 GH adopters, we calculated the propensity of being in a GH home, legacy

Table 1: Comparison of Facility-Level Variables across Green House and Comparison Nursing Homes

<i>Variables</i>	<i>Green House</i>	<i>All Potential Controls</i>	<i>Matched Controls</i>
Nonprofit ownership	0.667	0.233	0.480
For-profit ownership	0.200	0.716	0.422
Government ownership	0.133	0.051	0.099
Chain membership	0.267	0.553	0.395
Small facility (<91 beds)	0.200	0.317	0.251
Medium facility (91–125 beds)	0.200	0.341	0.238
Large facility (>125 beds)	0.600	0.342	0.511
Rural facility	0.267	0.329	0.332
High average ADL score	0.667	0.536	0.641
High percentage Medicaid	0.133	0.494	0.274
High percentage Medicare	0.467	0.517	0.448
High percentage other payer	0.800	0.498	0.695
Number of facilities	15	5,209	223

Notes. These facility-level observations were taken at baseline in 2005. A control facility had to be located in the same state as the matched Green House and it had to be in operation during the year the Green House opened. Given the small number of Green House nursing homes, none of the differences were statistically significant ($p < .1$) between the Green House nursing homes (column 1) and either the potential (column 2) or matched (column 3) control nursing homes.

home, or non-GH organization using a multinomial logistic regression model (Stuart et al. 2014). The covariates in these models were gender, black, age (younger than 65, 65–74, 75–84, 85 or older), Medicaid enrollment, diabetes mellitus, congestive heart failure, hypertension, dementia, depression, chronic obstructive pulmonary disease, cancer, an ADL score, and a cognitive performance scale. The application of the propensity score weights improved the balance across the GH and non-GH organizations for all the variables except cognitive impairment and cancer (see Table 2). Given the large number of observations however, the majority of differences were still statistically significant following the application of the weights.

Regressions

Our regression model took the following form:

$$Y_{iht} = \alpha + \beta GHU_{iht} + \gamma L_{iht} + \delta X_{iht} + \eta_h + \theta_t + \varepsilon_{iht}$$

where Y_{iht} was the person-level outcome (MDS quality measure, hospitalization, rehospitalization) for individual i in nursing home h at time t , GHU_{iht} was residence in a GH unit, L_{iht} was residence in a legacy unit, X_{iht} was a vector of

Table 2: Covariate Balance before and after Weighting for the Green House (GH) and Comparison Nursing Home Residents

<i>Variable</i>	<i>Before Weighting</i>		<i>After Weighting</i>	
	<i>GH</i>	<i>Non-GH</i>	<i>GH</i>	<i>Non-GH</i>
Female	0.733	0.701***	0.721	0.702***
Black race	0.032	0.070***	0.058	0.068***
Age <65	0.021	0.052***	0.045	0.051*
Age 65–74	0.097	0.129***	0.106	0.127***
Age 75–84	0.341	0.368***	0.386	0.367***
Age 85+	0.542	0.451***	0.463	0.455
Dual eligible	0.203	0.250***	0.270	0.248***
Activities of daily living score	12.376	11.883***	12.078	11.902***
Cognitive impairment scale	1.809	1.788***	1.856	1.789***
Diabetes mellitus	0.267	0.300***	0.317	0.299***
Congestive heart failure	0.234	0.275***	0.305	0.274***
Hypertension	0.707	0.729***	0.739	0.728**
Dementia	0.262	0.318***	0.361	0.316**
Depression	0.413	0.435***	0.453	0.434***
Chronic obstructive pulmonary disease	0.178	0.205***	0.218	0.204***
Cancer	0.098	0.097	0.118	0.097***
Number of assessments		645,191		
Number of unique individuals		135,794		
Number of nursing homes		238		
Number of GH nursing homes		15		

Notes. The results in the “Before Weighting” columns do not take account of propensity score weights, while the results in the “After Weighting” columns do account for these weights.
*Statistically significant at 10% level; **Statistically significant at 5% level; ***Statistically significant at 1% level.

resident-level characteristics, and η_h and θ_t were nursing home and quarter-year fixed effects, respectively. This model has the advantage of decomposing the impact of GH adoption for individuals residing in GH homes and individuals residing in the legacy nursing home. Relative to the comparison nursing homes, this model identified the direct effect of GH adoption for individuals residing in GH homes and the indirect (or spillover) effect of residing in a legacy unit within a nursing home that implemented the GH program on the same nursing home campus.

The resident-level characteristics X were gender, black/non-black, age (younger than 65, 65–74, 75–84, 85 or older), an indicator for whether the resident was enrolled in Medicaid, and a series of chronic condition indicators from the MDS (i.e., diabetes mellitus, congestive heart failure, hypertension, dementia, depression, chronic obstructive pulmonary disease, and cancer).

We also included two measures to capture resident acuity: an ADL score and a cognitive performance scale (Morris et al. 1994). Finally, as described earlier, we included a nursing home-level measure of payer mix.

We ran these regressions for a series of hospitalization and MDS-based quality indicators as the outcome Y using a series of logistic regression models. For the MDS regressions, each variable was coded as 1 if it was present in the assessment record, 0 otherwise. We investigated four hospitalization measures. We assessed whether the resident experienced any hospitalization or an “avoidable” hospitalization for any one of several conditions considered to be sensitive to nursing home care over the time since the previous MDS assessment. We also examined whether the individuals experienced a rehospitalization that followed within 30 days of the discharge date of the first hospitalization and an avoidable rehospitalization for any of these conditions within 30 days of the first hospitalization. Following the CMS Nursing Home Value-Based Purchasing Demonstration (L&M Policy Research 2013), we identified avoidable hospitalizations using the following six conditions: anemia (long-stay only), congestive heart failure, electrolyte imbalance, respiratory infection, sepsis, and urinary tract infection. Under this definition about three-fourths of hospitalizations in our samples were considered avoidable.

In each regression, we accounted for the clustering of observations within nursing homes by using Huber–White robust standard errors. We recognize that we are making multiple comparisons by examining eight MDS-based quality measures. As such, when interpreting our MDS quality results, we adjust p -values using the Holm–Bonferroni method (Holm 1979). Thus, the asterisks reporting statistical significance take into account the multiple comparison issues. We did not adjust the hospitalization results because we examined a primary outcome (overall hospitalizations) and then a series of subanalyses (avoidable, readmissions). All analyses were performed using STATA statistical software.

RESULTS

In the full sample of all residents in GH organizations, we examined 645,191 assessments of 135,794 unique residents in 11 states (see Table 3). After applying the weights based on the inverse of the propensity score, 50.3 percent of the overall sample resided in a nursing home that had implemented a GH home. For the analyses examining residents of GH homes and legacy units, the sample included assessments of residents in 10 states. Of this smaller

Table 3: Descriptive Statistics: Full Sample and Sample Identifying Residence in Green House (GH) Homes or Legacy Units

Variable	Full Sample		Sample with GH Identifiers	
	Mean	SD	Mean	SD
Independent variables				
Resident in GH organization	0.503	0.003	—	—
Resident in GH home	—	—	0.292	0.455
Resident in “legacy” unit	—	—	0.320	0.466
Female	0.712	0.453	0.700	0.458
Black race	0.063	0.243	0.064	0.245
Age <65	0.048	0.214	0.047	0.211
Age 65–74	0.117	0.321	0.139	0.346
Age 75–84	0.377	0.485	0.346	0.476
Age 85+	0.459	0.498	0.469	0.499
Dual eligible	0.259	0.438	0.269	0.444
Activities of daily living score	11.991	4.174	11.969	4.233
Cognitive impairment scale	1.822	0.706	1.799	0.663
Diabetes mellitus	0.308	0.462	0.288	0.453
Congestive heart failure	0.290	0.454	0.243	0.429
Hypertension	0.733	0.442	0.737	0.440
Dementia	0.339	0.473	0.306	0.461
Depression	0.443	0.497	0.437	0.496
Chronic obstructive pulmonary disease	0.211	0.408	0.204	0.403
Cancer	0.107	0.310	0.083	0.275
Dependent variables				
Hospitalization	0.092	0.290	0.094	0.292
Avoidable hospitalization	0.054	0.225	0.049	0.217
30-day rehospitalization	0.226	0.418	0.177	0.382
Avoidable 30-day rehospitalization	0.165	0.371	0.129	0.336
Bedfast	0.021	0.142	0.019	0.138
Incontinence, low risk	0.412	0.492	0.455	0.498
Catheter	0.097	0.296	0.091	0.287
Pain	0.145	0.352	0.159	0.366
Physical restraints	0.023	0.148	0.014	0.117
Pressure ulcers, high risk	0.155	0.362	0.158	0.364
Pressure ulcers, low risk	0.054	0.226	0.050	0.218
Urinary tract infection	0.152	0.359	0.144	0.351
Number of assessments	645,191		510,349	
Number of unique individuals	135,794		109,484	
Number of nursing homes	238		190	
Number of GH nursing homes	15		12	

Notes. The “Full Sample” was used to generate the results presented in Tables 4 comparing residents in GH organizations with residents in non-GH nursing homes. The “Sample with GH Identifiers” was used to generate the results presented in Tables 5 and 6 comparing individuals residing in GH homes and “legacy” units with residents of non-GH nursing homes. The “legacy” nursing home was the original nursing home that remained open alongside the GH homes. The smaller *N* in the GH and legacy unit sample was because three nursing homes (and their matched comparisons) were dropped because these nursing homes did not provide data on which residents lived in the GH homes.

weighted sample, 29.2 percent resided in a GH home, while 32 percent lived in a legacy home.

In an unadjusted comparison using the full sample, overall hospitalizations declined by 1.3 percentage points in nursing homes that adopted GH relative to 0.9 percent point decline in the comparison nursing homes (see Table 4). Thus, an unadjusted difference-in-difference estimate suggests the adoption of GH lowered hospitalizations by -0.3 percentage point ($= -0.0125 - (-0.0092)$) relative to the comparison nursing homes. All of the MDS QMs suggested relative quality improvements in GH organizations relative to the comparison nursing homes.

We examined the impact of GH adoption on residents of GH homes and those residing in legacy units (see Tables 5 and 6). Importantly, we excluded three GH nursing homes and their matched comparison nursing homes because we lacked data on whether residents were in the GH home or legacy unit. In contrast to the organization-level findings, we did not find any statistically significant impacts of GH adoption on either all hospitalizations or avoidable hospitalizations for residents in GH units. We did find a 5.5 percentage point decline in all 30-day readmissions; however, we observed the offsetting effect of a 7.2 percentage point increase for residents in the legacy unit. In addition, we found a significant 3.9 percentage point decline in avoidable

Table 4: Unadjusted Comparison of Pre-Post Quality across Green House (GH) and Comparison Nursing Homes ($N = 645,191$ Assessments)

	(1) Green House Pre-Post Difference	(2) Comparison Pre-Post Difference	(3) GH versus Comparison Difference
Hospitalization	-0.013	-0.009	-0.003
Avoidable hospitalization	-0.012	-0.008	-0.003
30-day rehospitalization	0.008	-0.005	0.013
Avoidable 30-day rehospitalization	0.002	-0.007	0.009
Bedfast	-0.013	-0.003	-0.010
Incontinence, low risk	0.022	0.058	-0.035
Catheter	-0.031	-0.013	-0.018
Pain	-0.056	-0.035	-0.021
Physical restraints	-0.011	-0.009	-0.003
Pressure ulcers, high risk	-0.062	-0.045	-0.016
Pressure ulcers, low risk	-0.019	-0.017	-0.002
Urinary tract infection	-0.024	-0.019	-0.005

Notes. Column 3 represents the unadjusted difference-in-difference estimates presenting the pre-post difference in Green House nursing homes minus the pre-post difference in the comparison nursing homes.

Table 5: Estimated Difference-in-Difference Regression Coefficient Estimates Comparing Individuals Living in Green House Homes and Legacy Units to Those Not Residing in a Green House Organization

	(1) Hospitalizations	(2) Avoidable Hospitalizations	(3) Rehospitalizations	(4) Avoidable Rehospitalizations
Green House home	0.004 (0.015)	−0.007 (0.005)	−0.055** (0.024)	−0.039** (0.019)
Legacy Unit	−0.014 (0.016)	−0.005 (0.010)	0.072** (0.033)	0.023 (0.017)
N	510, 349	510, 347	53, 609	53, 560

Notes. The “legacy” nursing home is the original nursing home that remains open alongside the Green House homes. Robust standard errors clustered at the nursing home level are presented in parentheses. All regressions include covariates in Table 3, nursing home-level payer mix measures, and time (quarter) and nursing home fixed effects.

*Statistically significant at 10% level; **Statistically significant at 5% level; ***Statistically significant at 1% level.

readmissions in GH units, but the corresponding increase in legacy units was not significant.

In regard to the MDS outcomes, GH units experienced a 0.3 percentage point (15.8 percent) decline in bedfast residents, a 4.1 percentage point (45 percent) decline in catheterized residents, and 1.9 percentage point (38 percent) decline in low-risk residents with pressure ulcers. We found no statistically significant impact of GH unit residence on incontinence (low-risk residents only), restraints, pressure ulcers (high-risk residents only), pain, and urinary tract infections, but all of these measures except the pain result were in the direction of higher GH quality. When we created a composite MDS quality measure using the bedfast, catheter, pain, physical restraint, and urinary tract infection measures (following the approach used in Anderson 2008), we found that quality improved following GH adoption, but the result was not statistically significant (data not shown). In terms of the results for residents of legacy units, none of the results were statistically significant, but the majority of the point estimates suggested lower quality following the adoption of a GH unit.

To investigate the overall impact of GH adoption, we ran the model on the entire universe of residents in a GH organization (results available upon request from the authors). In this specification, we modeled GH adoption based on the share of residents in a GH organization. The only statistically significant impact of GH improving overall quality was for the bedfast measure. Because the findings for GH unit residents using the readmission, catheter, and low-risk pressure ulcer measures were not present across the entire GH

Table 6: Estimated Difference-in-Difference Regression Coefficient Estimates for Minimum Data Set Quality Measures: Comparison of Individuals Living in Green House (GH) Homes and Legacy Units Relative to Those Not Residing in a GH Organization

	(1) <i>Bedfast</i>	(2) <i>Incontinence Low Risk</i>	(3) <i>Catheter</i>	(4) <i>Pain</i>
GH home	−0.003*** (0.0003)	−0.037 (0.063)	−0.041*** (0.010)	0.006 (0.024)
Legacy home	−0.001 (0.002)	−0.129 (0.107)	−0.019 (0.009)	0.002 (0.021)
<i>N</i>	374,713	284,809	378,025	510,193
	(5) <i>Restraints</i>	(6) <i>Ulcers High Risk</i>	(7) <i>Ulcers Low Risk</i>	(8) <i>Urinary Tract Infection</i>
GH home	−0.003 (0.001)	−0.012 (0.013)	−0.019*** (0.003)	−0.029 (0.014)
Legacy home	0.001 (0.003)	0.008 (0.012)	0.018 (0.012)	0.001 (0.014)
<i>N</i>	326,987	223,144	155,232	378,814

Notes. The “legacy” nursing home is the original nursing home that remains open alongside the GH homes. Robust standard errors are presented in parentheses. Robust standard errors clustered at the nursing home level are presented in parentheses. All regressions include covariates in Table 3, nursing home-level payer mix measures, and time (quarter) and nursing home fixed effects.

*Statistically significant at 10% level; **Statistically significant at 5% level; ***Statistically significant at 1% level. The asterisks reporting statistical significance take into account the multiple comparison issues.

organization, some possible offsetting effects may exist for residents of legacy units within the GH organization.

DISCUSSION

Overall, we found some evidence that the GH model improved nursing home quality of care. For residents living in GH homes, we found the adoption of the model led to fewer 30-day readmissions (overall and avoidable), bedfast residents, catheterized residents, and low-risk residents with pressure ulcers.

The finding that adopting the model reduced readmissions for residents in the GH unit is supported by other evidence in this special issue. Bowers et al. (2016) found that the GH model’s consistent assignment of universal worker direct care staff and small homes built around a central living area resulted in familiarity with residents and the opportunity for frequent interactions among direct care workers, nurses, and physicians. This potential increase in communication and multidisciplinary collaboration can, in turn, lead to early identification and intervention in response to resident change of

medical condition, a vital step in high-quality care. Although some GH homes took advantage of these opportunities to improve quality, others did not. Grabowski et al. (2016) found lower overall Medicare spending for residents of GH homes, as measured by hospitalizations, institutional postacute care, and hospice.

The potential to reduce readmissions is also consistent with previous nursing home research. Nursing home readmissions have been found to occur frequently and often for preventable reasons (Mor et al. 2010). Models that emphasize care management in the nursing home such as the INTERACT (Ouslander 1988) and Evercare (Kane et al. 2003) programs have been found to be associated with fewer hospital transfers. Although the GH model differs from both of these models, the close contact between the direct care staff and residents in GH homes coupled with a person-centered model of care might have led to better postacute care management and ultimately fewer rehospitalizations.

The findings suggest that adoption of the GH model did improve certain quality measures for residents living in a GH home, but the absence of evidence of a broader statistically significant impact on MDS-based quality of care measures is consistent with an earlier evaluation of the GH model (Kane et al. 2007), the culture change literature more broadly (Shier et al. 2014), and a number of other evaluations of long-term care programs that emphasize quality of life, including home- and community-based services (Grabowski 2006) and consumer-directed home care (Carlson et al. 2007).

In this current study, we have to acknowledge its limited statistical precision. Taken as a whole based on what we were able to learn from this study, it is unclear whether better quality of care and better quality of life are complementary outputs in the production of care under this model. Specialization is fairly common in health care (Detsky, Gauthier, and Fuchs 2012), and it is not uncommon for providers to excel in certain dimensions of care while being roughly “average” in others. Similar to how certain nursing homes specialize in outcomes for short-stay rehabilitation care (Li et al. 2012), GH adopters may specialize in improvements to the quality of life for their residents without a major change in quality of care. Put alternatively, GH may offer a more homelike, person-centered model without sacrificing quality of care relative to a traditional nursing home setting.

Our analysis is subject to several limitations. First, the nursing homes that adopted the GH model differed from nonadopters (Grabowski et al. 2014a). Although we constructed a comparison group by matching on observable facility and individual characteristics, we cannot control for unobserv-

ables associated with GH adoption by organizations or unobservables in the attributes of residents cared for in the GH homes relative to the legacy units. The gold standard study would randomize nursing homes and residents at baseline to receive GH or not and then follow them over time. However, the required funding and complicated logistics of such a study makes a randomized intervention unlikely. Ultimately, we will have to rely on observational studies such as this one.

Second, statistical power in the study is somewhat limited, especially in the model that analyzed residents of GH units. The number of individuals in many GH organizations was relatively small. By design, the GH homes have 8–12 elders living together. Many organizations have multiple GH units, but the overall number of residents was still relatively small at the nursing home level, ranging from 3 to 174 total residents at any given time during our study period. We excluded three GH organizations in the analysis of residents living in GH homes due to missing data, which further diminished our sample size. Moreover, the inclusion of nursing home-level fixed effects further inflated the size of our standard errors. Although the inclusion of random effects (rather than fixed effects) generates smaller standard errors, we opted not to use this approach because it rests on the rather strict assumption that the random effects are uncorrelated with all the observable variables included in the regression model.

To evaluate the degree of precision in our estimates, we multiplied the standard error estimate by ± 1.96 to get an estimate of what effect size would have counterfactually been significant. When we examined the impact of GH adoption on avoidable hospitalizations for GH home residents, we would have to obtain an effect size of a 1.0 percentage point decline ($= -1.96 \times 0.005$) to achieve statistical significance. Relative to the dependent variable mean, a 1.0 percentage point decline would constitute a 20 percent decline in avoidable hospitalizations; yet most would consider a range of values less than 1.0 to be clinically significant. Given this issue of limited precision, we recommend continued analysis of the impact of the GH model using larger samples of adopters over longer periods of study.

Third, we acknowledge that several of the quality of care outcomes may also be related to quality of life within a nursing home. For example, whether a resident is bedfast or in pain would have important implications for the quality of life experienced by this resident. This inability to distinguish between the constructs of quality of care and quality of life does not bias our empirical estimates, but it may have implications for whether we have identified a clean test of the hypotheses outlined in the conceptual framework.

Finally, our estimates are of average effects across all GH adopters. As other research in this issue found, GH organizations varied in their implementation (Cohen et al. 2016) and sustainability (Bowers et al. 2016) of the model. Qualitative analysis suggested some GH adopters were able to take advantage of opportunities to achieve quality improvements, while others were not (Bowers et al. 2016). Future research should test whether GH adopters with greater model fidelity achieve improvements in quality relative to adopters that depart from the model. Once again, testing heterogeneous effects across adopters and residents will require the application of larger samples of adopters over longer periods of study.

The absence of evidence of universally better quality, on the one hand, may disappoint those who view quality of life and quality of care as strict complements and may have hoped that the GH model would lead to measurable “across-the-board” improvements in quality of clinical care and reduced hospitalization rates. On the other hand, for those who view quality of life and quality of care as substitutes and may have been concerned about lower quality, the absence of reductions in quality of care should be reassuring and the fact that GH did lower hospitalization readmissions and improve certain quality indicators should be encouraging. In any case, for GHs and culture change more broadly, the findings underscore the challenge of improving quality of care and highlight the need for adopters to continue to focus on clinical care as they implement efforts to improve quality of life.

ACKNOWLEDGMENTS

Joint Acknowledgment/Disclosure Statement: This project was supported by the Robert Wood Johnson Foundation. The authors thank the participating subjects for their time and effort to promote better care and quality of life for those receiving support in these and related settings.

Disclosures: None.

Disclaimers: None.

REFERENCES

- Anderson, M. L. 2008. “Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and

- Early Training Projects." *Journal of the American Statistical Association* 103 (484): 481–1495.
- Bowers, B., K. Nolet, and N. Jacobson, for the THRIVE Research Collaborative. 2016. "Sustaining Culture Change: Experiences in the Green House Model." *Health Services Research* 51 (1 Pt 2): 398–417.
- Bowers, B., T. Roberts, K. Nolet, and B. Ryther, for the THRIVE Research Collaborative. 2016. "Inside the Green House "Black Box": Opportunities for High Quality Clinical Decision Making." *Health Services Research* 51 (1 Pt 2): 378–397.
- Carlson, B. L., L. Foster, S. B. Dale, and R. Brown. 2007. "Effects of Cash and Counseling on Personal Care and Well-Being." *Health Services Research* 42 (1 Pt 2): 467–87.
- Cohen, L. W., S. Zimmerman, D. Reed, P. Brown, B. J. Bowers, K. Nolet, S. Hudak, and S. D. Horn for the THRIVE Research Collaborative. 2016. "The Green House Model of Nursing Home Care in Design and Implementation." *Health Services Research* 51 (1 Pt 2): 352–77.
- Detsky, A. S., S. R. Gauthier, and V. R. Fuchs. 2012. "Specialization in Medicine: How Much is Appropriate?" *Journal of the American Medical Association* 307 (5): 463–4.
- Grabowski, D. C. 2006. "The Cost-Effectiveness of Noninstitutional Long-Term Care Services: Review and Synthesis of the Most Recent Evidence." *Medical Care Research and Review* 63 (1): 3–28.
- Grabowski, D. C., A. Elliot, B. Leitzell, L. W. Cohen, and S. Zimmerman. 2014a. "Who Are the Innovators? Nursing Homes Implementing Culture Change." *The Gerontologist* 54 (Suppl 1): S65–75.
- Grabowski, D. C., A. J. O'Malley, C. C. Afendulis, D. J. Caudry, A. Elliot, and S. Zimmerman. 2014b. "Culture Change and Nursing Home Quality of Care." *The Gerontologist* 54 (Suppl 1): S35–45.
- Grabowski, D. C., C. C. Afendulis, D. J. Caudry, A. J. O'Malley, and P. Kemper for the THRIVE Research Collaborative. 2016. "The Impact of Green House Adoption on Medicare Spending and Utilization." *Health Services Research* 51 (1 Pt 2): 433–53.
- Harrington, C., C. Kovner, M. Mezey, J. Kayser-Jones, S. Burger, M. Mohler, R. Burke, and D. Zimmerman. 2000. "Experts Recommend Minimum Nurse Staffing Standards for Nursing Facilities in the United States." *The Gerontologist* 40 (1): 5–16.
- Holm, S. 1979. "A Simple Sequentially Rejective Multiple Test Procedure." *Scandinavian Journal of Statistics* 6: 65–70.
- Institute of Medicine. 2001. *Improving the Quality of Long-Term Care*. Washington, DC: National Academy Press.
- Kane, R. L., G. Keckhafer, S. Flood, B. Bershadsky, and M. S. Siadat. 2003. "The Effect of Evercare on Hospital Use." *Journal of the American Geriatrics Society* 51 (10): 1427–34.
- Kane, R. A., T. Y. Lum, L. J. Cutler, H. B. Degenholtz, and T. C. Yu. 2007. "Resident Outcomes in Small-House Nursing Homes: A Longitudinal Evaluation of the Initial Green House Program." *Journal of the American Geriatrics Society* 55 (6): 832–9.

- Koren, M. J. 2010. "Person-Centered Care for Nursing Home Residents: The Culture-Change Movement." *Health Affairs (Millwood)* 29 (2): 312–7.
- L&M Policy Research. 2013. "Evaluation of the Nursing Home Value-Based Purchasing Demonstration: Year 3 and Final Evaluation Report" [accessed on November 12, 2013]. Available at http://innovation.cms.gov/Files/reports/NursingHomeVBP_EvalReport.pdf
- Li, Y., X. Cai, J. Yin, L. G. Glance, and D. B. Mukamel. 2012. "Is Higher Volume of Postacute Care Patients Associated with a Lower Rehospitalization Rate in Skilled Nursing Facilities?" *Medical Care Research and Review: MCRR* 69 (1): 103–18.
- Lum, T. Y., R. A. Kane, L. J. Cutler, and T. C. Yu. 2008. "Effects of Green House Nursing Homes on Residents' Families." *Health Care Financing Review* 30 (2): 35–51.
- Miller, S. C., M. Lepore, J. C. Lima, R. Shield, and D. A. Tyler. 2014. "Does the Introduction of Nursing Home Culture Change Practices Improve Quality?" *Journal of the American Geriatrics Society* 62 (9): 1675–82.
- Mor, V., O. Intrator, Z. Feng, and D. C. Grabowski. 2010. "The Revolving Door of Rehospitalization from Skilled Nursing Facilities." *Health Affairs (Millwood)* 29 (1): 57–64.
- Morris, J. N., B. E. Fries, D. R. Mehr, C. Hawes, C. Phillips, V. Mor, and L. A. Lipsitz. 1994. "MDS Cognitive Performance Scale." *Journal of Gerontology* 49 (4): M174–82.
- Ouslander, J. G. 1988. "Reducing the Hospitalization of Nursing Home Residents." *Journal of the American Geriatrics Society* 36 (2): 171–3.
- Sharkey, S. S., S. Hudak, S. D. Horn, B. James, and J. Howes. 2011. "Frontline Caregiver Daily Practices: A Comparison Study of Traditional Nursing Homes and the Green House Project Sites." *Journal of the American Geriatrics Society* 59 (1): 126–31.
- Shier, V., D. Khodyakov, L. W. Cohen, S. Zimmerman, and D. Saliba. 2014. "What Does the Evidence Really Say about Culture Change in Nursing Homes?" *The Gerontologist* 54 (Suppl 1): S6–16.
- Stuart, E. A., H. A. Huskamp, K. Duckworth, J. Simmons, Z. Song, M. Chernew, and C. L. Barry. 2014. "Using Propensity Scores in Difference-in-Differences Models to Estimate the Effects of a Policy Change." *Health Services & Outcomes Research Methodology* 14 (4): 166–82.
- U.S. Government Accountability Office. 2005. "Nursing Homes: Despite Increased Oversight, Challenges Remain in Ensuring High-Quality Care and Resident Safety". GAO-06-117.

SUPPORTING INFORMATION

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

Appendix SA1: Author Matrix.