Original Research Article

Person-Centered Care as Facilitated by Kansas' PEAK 2.0 Medicaid Pay-for-Performance Program and Nursing Home Resident Clinical Outcomes

Linda Hermer, PhD,[1,][*,] Laci Cornelison, MS,[2] Migette L. Kaup, PhD,[3] Judith L. Poey, PhD,[1] Patrick N. Drake, MS,[1] Robyn I. Stone, DrPH,[1] and Gayle A. Doll, PhD[2]

[1]LeadingAge LTSS Center at UMass Boston, LeadingAge, Washington, District of Columbia. [2]Center on Aging and [3]Department of Apparel, Textiles and Interior Design, Kansas State University, Manhattan, KS.

*Address correspondence to: Linda Hermer, PhD, National Center for Health Statistics, 3311 Toledo Rd., Metro IV, Rm. 3535, Hyattsville, MD 20782. E-mail: lindalhermer@gmail.com

Abstract

Purpose of the Study: Person-centered care (PCC) is intended to improve nursing home residents' quality of life, but the closer bonds it engenders between residents and staff may also facilitate improvements to residents' clinical health. Findings on whether adoption ameliorates resident clinical outcomes are conflicting, with some evidence of harm as well as benefit. To provide clearer evidence, the present study made use of Kansas' PEAK 2.0 Medicaid pay-for-performance (P4P) program, which incents the adoption of PCC. The program is distinctive in training facilities' staff on adopting PCC through a series of well-defined stages and providing regular feedback about their progress.
Design and Methods: A retrospective cohort study was performed with 349 Kansas facilities spread across several well-defined PCC adoption stages, ranging from nonadoption to comprehensive adoption. The outcomes were thirteen 2014–2016 Nursing Home Compare long-stay resident clinical measures and a composite measure incorporating only nonimputed data for those 13 outcomes. Observed facility demographic differences were controlled for with propensity score adjustment. Treatment effect analyses were run with each outcome, with the predictor variable of program stage.
Results: Seven of the 13 clinical measures plus the composite measure indicated better health for residents in homes at higher program stages, relative to those in nonparticipating homes, including a 49% lower prevalence of major depressive symptoms in strongly adopting facilities.
Implications: The findings suggest that greater PCC adoption through PEAK participation is associated with better quality of care. Policymakers in other states may want to consider implementing a program modeled on PEAK 2.0.

Translational Significance: Unlike most other studies of person-centered care, or culture change, and pay-for-performance programs, this study found evidence of major health benefits for residents, including a 49% lower prevalence in major depressive symptoms in facilities most strongly participating in the program.

Keywords: Culture change, Value-based purchasing, Nursing homes

In a landmark mixed-method study in the early 1980s, U.S. nursing home residents emphasized that quality of care and quality of life were inextricably linked and that they considered both to be woefully deficient (Koren, 2010). Despite a long series of regulatory reforms over the past three decades, improving resident care in U.S. nursing homes has remained a vexing issue (Government Accountability Office, 2005). As the deficits became more widely acknowledged, several extraregulatory movements and initiatives emerged, including the "culture change" movement and the deployment of pay-for-performance (P4P) programs for nursing homes. The nursing home culture change movement aims primarily to improve residents' quality of life while not diminishing the quality of the care, whereas various nursing home P4P programs, typically funded by Medicaid, have targeted quality of life, quality of care, and other dimensions of nursing home operations (Arling, Job, & Cooke, 2009; Stone, Bryant, & Barbarotta, 2009; Weissert & Frederick, 2013). Recent reviews and large-scale studies suggest that some of these efforts have begun to improve residents' quality of life (Kim & Park, 2017) and satisfaction with their nursing homes (Poey et al., 2017). However, it remains unclear whether either culture change or P4P programs consistently improve residents' quality of care. The present study examines whether a current P4P program in the state of Kansas that incents culture change adoption is associated with improved quality of care.

Culture Change and Quality of Care

The culture change movement aims to deconstruct the medical model for nursing homes (White-Chu, Graves, Godfrey, Bonner, & Sloane, 2009), replacing it with a more homelike atmosphere in which nursing home operations are centered around the needs and desires of individual residents. Its main tenets are to make resident choices and preferences guide all aspects of daily life; to ensure that residents, staff members, families, and the broader community develop close bonds; to make the nursing home environment more like the homes residents used to live in as opposed to a hospital; and to empower frontline staff members to fulfill residents' needs while also having opportunities to advance in their careers; and finally, to continually improve the quality of care with the use of quality measures (Koren, 2010). "Culture change" is often used synonymously with "person-centered care" (PCC), as it will be here.

The evidence on whether adopting PCC improves resident clinical outcomes is mixed. Most previous studies of PCC had very small sample sizes or suffered from other serious methodological limitations and were collectively given a "D" grade in a recent systematic review (Shier, Khodyakov, Cohen, Zimmerman, & Saliba, 2014). The few moderate- to large-scale studies that have been performed have revealed either scant evidence of clinical improvements (Grabowski et al., 2014; Stone et al., 2002; Sullivan,

Shwartz, Stolzmann, Afable, & Burgess, 2018) or modest benefits to resident health (Afendulis et al., 2016; Miller, Lepore, Lima, Shield, & Tyler, 2014). For example, Sullivan and colleagues (2018) conducted a large, longitudinal study of the clinical effects of the Veterans' Administration's initiative to adopt PCC at all its nursing facilities, examining potential changes to a composite measure of health status incorporating numerous MDS 3.0 measures. They found no effect of adopting PCC on resident health overall and a weakly negative effect of it in the first seven quarters after adoption began. Similarly, although Grabowski and colleagues (2014) found a slight reduction in survey health deficiencies in nationally recognized PCC-adopting facilities compared with matched facilities that were not known to be strong adopters, they found no other evidence of clinical benefits. In contrast, some other studies have found modest benefits, such as Miller and colleagues (2014) who demonstrated that facilities thoroughly as opposed to partially adopting culture change experienced several clinical improvements. But as with Sullivan and colleagues (2018), Miller and colleagues (2014) also found evidence of harm—increases in urinary tract infections and hospitalizations—although only in the comparison group of partially adopting homes.

It is difficult to draw conclusions from these studies for a number of reasons beyond the mixed evidence itself. For instance, some studies' control groups included facilities without any measurement of their adoption levels (Grabowski et al., 2014) or employed a comparison group of homes spread across a wide range of partial adoption stages (Miller et al., 2014), potentially reducing the studies' sensitivity to effects of specific levels and areas of adoption. This is important because only 13% of U.S. facilities are thorough adopters, whereas 74% of U.S. facilities are spread along the vast continuum of partial adoption (Miller, Looze, Shield et al., 2014). Furthermore, the degree of PCC adoption was often not objectively determined. In the study by Sullivan and colleagues (2018), for example, facility administrators' judgments of PCC adoption were employed, as reflected in their Artifacts of Culture Change scores. This is a significant limitation because some members of facility administration, particularly directors of nursing, tend to over-report the extent of culture change adoption (Tyler et al., 2011), and adoption tends to be irregular, being tailored to each facility (Zimmerman et al., 2016).

P4P Programs and Quality of Care

To improve clinical quality, many states have turned to pay-for-performance (P4P) programs (Arling et al., 2009; Briesacher, Field, Baril, & Gurwitz, 2009; Weissert & Frederick, 2013; Werner, Konetzka, & Polsky, 2013). These programs are designed to reward quality of care, and they have long held the promise of improving health outcomes while lowering costs. Nursing homes should be an especially good target of P4P mechanisms because of

their low operating margins, their chronically ill population, and their performance of many routine tasks amenable to regular measurement and evaluation (Weissert & Frederick, 2013).

But as with P4P programs in other areas of health care (e.g., Emmert, Eijkenaar, Kemter, Esslinger, & Schoffski, 2012; Markovitz & Ryan, 2016; Mendelson et al., 2017), few nursing home P4P programs have realized their promise. Werner and colleagues (2013) analyzed nursing home clinical outcomes in eight states with P4P programs incenting improvements to the quality of care, comparing their outcomes to those in states without such programs. They found approximately 5% reductions in the prevalence of pressure ulcers, the use of physical restraints, and residents reporting moderate to severe pain. However, they also found significantly worse performance in the areas of excessive weight loss, the use of indwelling catheters, and survey deficiencies (Werner et al., 2013)—as many deleterious outcomes as beneficial ones. Systematic reviews and meta-analyses of nursing home P4P programs have found similarly mixed and unimpressive results (Briesacher et al., 2009; Weissert & Frederick, 2013).

The Centers for Medicare and Medicaid Services (CMS) recently undertook the largest nursing home P4P (or value-based purchasing) quantitative experiment attempted in the United States, the Nursing Home Value-Based Purchasing (NHVBP) demonstration, finding little evidence that it improved resident health, some evidence that it harmed residents, and no evidence that it lowered costs (Grabowski et al., 2017). A qualitative analysis of factors underpinning the results led to recommendations that to succeed, P4P programs should

1. not be designed with the assumption that facilities have the infrastructure and expertise to succeed in quality improvement on their own, and instead, incorporate training and guidance on this process;
2. provide regular feedback on facilities' progress in achieving the targeted outcomes;
3. include an incentive structure that is easy for facility leadership and managers to understand; and
4. include substantial financial incentives (Green, 2013).

The PEAK 2.0 Program Incenting PCC Adoption in Kansas Nursing Homes

The Promoting Excellent Alternatives in Kansas (PEAK 2.0) program is a state program that has incorporated recommendations from the NHVBP program and several other studies of P4P programs (e.g., Arling et al., 2009). Nursing homes are rigid organizations that typically resist change (Rahman, Applebaum, Schnelle, & Simmons, 2012), and PEAK's design also accords with recommendations for promoting innovation in such organizations (Rogers, 2003). Founded in 2012, PEAK 2.0 incents the adoption of PCC rather than directly incenting the improvement of clinical quality. It provides participants with explicit training on

best practices for implementing PCC, standardized definitions, structured education and training, objective and regular feedback on progress, and a sizable and straightforward, escalating Medicaid incentive that is tied to six graduated levels of PCC adoption. The program facilitates PCC implementation in four major areas: resident choice, homelike environment, meaningful life, and staff empowerment (Doll, Cornelison, Rath, & Syme, 2017).

These four domains are broken down into 12 cores that homes may address across their participation the program (see Supplementary Figure 1).

> Each of these 12 cores addresses a unique aspect of PCC and is subsequently broken down into supporting practices. For instance, the food core, which falls under the Resident Choice domain, includes three supporting practices: what to eat, when to eat, and where to eat. Expected outcomes in the food core include, but are not limited to, items such as resident input in menu development, enhanced dining options, expanded meal times, and food availability 24/7 (Doll et al., 2017, 309).

PEAK levels for each home are determined objectively each year by in-person or video-conferenced evaluations conducted by program staff. The program incents and trains participants on both the processes needed to achieve PCC in all areas of nursing home operations and the actual content areas of PCC, starting with a year of structured education, training, and strategic planning at what is called the Foundation Level. Homes at Level 1 strive to implement PCC in the first four cores they select during the Foundation year. Homes at Level 2 work to implement PCC in the remaining eight program cores, and homes at Levels 3–5 sustain adoption of PCC in all 12 program cores. The Medicaid incentive ranges from $0.50 per Medicaid resident per day for Foundation and Level 1 facilities to $4.00 per Medicaid resident per day for Level 5 facilities. (For more information on the PEAK program, see Doll and colleagues (2017) and visit http://www.kdads.ks.gov/commissions/survey-certification-and-credentialing-commission/peak.)

Conceptual Model

Although the chief aim of both the nursing home culture change movement and PEAK is to improve residents' quality of life (Burack, Weiner, Reinhardt, & Annunziato, 2012; Koren, 2010), they may also improve the quality of care, similar to arguments made by Grabowski and colleagues (2014). The adoption of PCC practices such as consistent assignment of staff to residents and other means of fostering close resident–staff relationships should lead to more attentive care. It should also result in greater empowerment of elders to communicate about their health status, which might additionally improve health outcomes.

Some outcomes might be expected to improve more than others under the adoption of PCC practices. For example, clinical depression might be diagnosed more promptly and

treated more aggressively and effectively. However, the thorough adoption of resident choice might result in putative harms as well. Facilities at high program stages might exhibit lower rates of influenza or pneumococcal vaccinations because residents are given the choice to reject them, for instance, or they might experience more falls because they feel empowered to move throughout the facility.

Quantifying PEAK 2.0's Impact

To investigate these possibilities, a retrospective cohort study was conducted of the effects of the program on facility-level, resident Nursing Home Compare health outcomes that included all certified facilities in Kansas during the PEAK 2.0 program years of 2014–2015 and 2015–2016. It was hypothesized that participation in the PEAK program, especially at higher stages, would be associated with better clinical outcomes overall, relative to nonparticipating facilities. The outcome measures for the study came from Nursing Home Compare. Because some outcomes were missing data, the hypothesis was investigated in two ways: (1) with separate analyses for each of the 13 outcome measures after imputation of missing data had taken place, and (2) with an analysis of a composite measure constructed only from extant (nonimputed) data from the 13 outcomes, with the composite analysis serving as a sensitivity analysis.

Method

The approach to this study entailed constructing a data set of covariates, outcomes, and the critical predictor variable Stage for each of the two program years studied. For the main analyses, multivariate imputation was performed with the covariates for each year, followed by the estimation of propensity scores with multinomial logistic regression for each facility each of the two studied program years, to control for all observed differences across homes at different stages. Finally, inverse probability-weighted, average treatment effect models were used to test for associations between program stage and each of the 13 resident health measures. For the sensitivity analysis, a composite measure was constructed using only nonimputed data for all 13 outcomes (important because the sentence just above refers to using imputed data), and an inverse probability-weighted treatment effect analysis of the relationship between program stage and the composite measure was performed.

Data sets

The data sets were derived from data from the Kansas Department of Aging and Disability Services (PEAK program data), CMS's Nursing Home Compare, CASPER survey reports, and the 2010 U.S. Census. The data sets included 349 certified Kansas nursing homes in existence both years serving approximately 18,230 residents. All data were at the facility level.

Variables

The predictor variable program stage (henceforth "Stage") was created as follows:

- Stage 0: Nonparticipating.
- Stage 1: Participating at the Foundation Level (undergoing structured education and training).
- Stage 2: Participating at Level 1 (adopting PCC in four program areas). A typical Stage 2 home was beginning to implement changes, often starting with cores from the Resident Choice domain such as Food.
- Stage 3: Participating at Level 2 (adopting PCC in 8-12 program cores) through Level 5 (Levels 3-5 were sustaining PCC adoption across all 12 program cores).

The number of facilities at Levels 3–5 was small ($N = 9$ in 2014–2015 and $N = 10$ in 2015–2016), and they differed considerably from homes at other levels demographically. In initial treatment effect models run with Levels 3–5 comprising their own stage, it was not possible to achieve a satisfactory balance of covariates. Thus, homes at Levels 3–5 were combined with homes at Level 2 (strong partial adopters) to make a larger Stage 3 group.

A set of potential confounder variables (Table 1) was chosen because they had previously been shown to distinguish culture change adopters from nonadopters and to be associated with nursing home quality of care (as discussed at length in Hermer et al., 2017). They were included in all analysis models as potential confounders.

The outcomes for the main analyses included all 13 Nursing Home Compare measures for long-stay residents and a description of each is presented in Table 2. The Nursing Home Compare data are derived from CMS's Minimum Dataset 3.0. For the composite analysis, a single combined measure was derived from the 13 Nursing Home Compare outcomes (see Data Preparation).

Data Preparation

As can be seen in Supplementary Tables 1 and 2, there was a small amount of missing data for the covariates and a considerable amount of missing data for the outcomes, mostly deriving from facilities' having too small a denominator for a given measure, which typically correlates with facility size. Thus, for the main analyses, single multivariate imputation using chained equations in R (van Buuren & Groothuis-Oudshourn, 2010) was performed to prevent the loss of observations during these analyses due to missing outcome or covariate data. Next, propensity scores were generated from covariate data for performing propensity score adjustment, estimated with multinomial logistic regression. See Supplemental Material for more details on these procedures.

For the sensitivity analysis, a composite measure was constructed from nonimputed data only from all 13 outcomes. This measure was z-score-based. For more information on its construction, see Supplementary Material.

Table 1. Study Covariates Considered to Be Potential Confounders

Covariate	Source	Description
Profit status	Nursing Home Compare	Nonprofit, for profit, or government owned (dummy-coded for analysis)
CCRC affiliation	Nursing Home Compare	CCRC affiliated or not
Chain ownership	Nursing Home Compare	Chain owned or not
Certified beds	Nursing Home Compare	Number of Medicaid, Medicare, or both types of certified beds, as a measure of size
% Occupancy	Nursing Home Compare	Number of residents divided by the number of beds × 100
% Medicaid residents	Nursing Home Compare	% of residents whose primary payer is Medicaid
% Private-pay/-insurance residents	Nursing Home Compare	% of residents whose primary payer is not Medicaid or Medicare
Aide HRD	Nursing Home Compare	Actual aide staffing HRD
LPN HRD	Nursing Home Compare	Actual LPN staffing HRD
RN HRD	Nursing Home Compare	Actual RN staffing HRD
Exp. Tot. Hrs.	Nursing Home Compare	Expected total aide, LPN, and RN HRD given case mix index (as in Hermer et al., 2017)
HHI	Derived from CASPER survey reports	County-based measure of market concentration, ranging from 0 = perfectly competitive to 1 = completely monopolized
Rurality	2010 U.S. Census	Percentage, ranging from lower to higher population density

Note: CCRC: Continuing Care Retirement Community (LifePlan community); Exp. Tot. Hrs. = expected total hours of RN, LPN, and aide staffing given facilities' case mix (as a proxy for case-mix index); HHI = Herfindahl–Hirschman Index; HRD = hours per resident day; LPN = licensed practical nurse; RN = registered nurse.

Table 2. Long-Stay Resident Outcome Measures from Nursing Home Compare Used in the Study (RTI, 2014) and the Composite Outcome Measure Created from Them

Outcome Variable	Description
Pressure ulcers in high-risk residents	% of residents with impaired bed mobility or who are malnourished with stages II–IV or unstageable pressure ulcers during the reporting period
Physical restraint use	% of residents physically restrained on a daily basis
Major depressive symptoms	% of residents feeling depressed or hopeless, and/or feeling little interest in doing things half or more of the days in the last 2 weeks, or whose total severity score indicates major depression
Antipsychotic use	% of residents on antipsychotics during the current assessment period who did not have a diagnosis of schizophrenia, Tourette's syndrome, or Huntington's disease
Indwelling catheter use	% of residents with an indwelling catheter in the last 7 days
Urinary tract infections	% of residents with a urinary tract infection in the last 30 days
Incontinent episodes in low-risk residents	% of residents experiencing one or more incontinent episode among residents with few transferring or other mobility issues
Excessive weight loss	% of residents who had a weight loss of 5% or more in the last month or 10% or more in the last two quarters who were not on a physician-prescribed weight loss regimen
Increased loss of ADLs	% of residents whose need for help with ADLs has increased over the last reporting period
Moderate to severe pain	% of residents reporting either (1) almost constant or frequent moderate to severe pain in the last 5 days or (2) any very severe/horrible pain in the last 5 days
Falls with major injury	% of residents experiencing one or more falls with major injury in current reporting or look-back period
Pneumococcal vaccinations	% of residents who medically appropriately received pneumococcal vaccinations during the last 12 months
Influenza vaccinations	% of residents medically appropriately given the influenza vaccination during the most recent influenza season
Composite measure	z-score-based measure created from the above 13 outcomes

Note: ADLs = activities of daily living. All the individual resident outcome measures were prevalence rates.

Analyses

Analyses were conducted using Stata v.15 (StataCorp, College Station, TX). First, to examine whether facilities exhibited any demographic differences as a function of program stage, chi-square or Fisher's exact tests were performed on categorical data and one-way ANOVAs were performed on continuous data.

Next, to conduct the main analyses and the composite analysis to test the hypothesis that better clinical outcomes (either taken separately or in composite form) would be

associated with higher stages of the PEAK 2.0 program when compared with Stage 0, linear average treatment effect models were run with inverse probability weighting of covariates using the propensity scores. All potential confounders and each case's inverse probability weight were included in the models. A case's weight is equal to the inverse of the probability of the treatment they received. Treatment effect models approximate randomization on observed variables (StataCorp, 2017). As such, they account more robustly for missing data than multivariate models by modeling a counterfactual control group, comprised a random and balanced subset of individuals in the study (including those actually receiving the control condition and those not), and (in this study's case) three counterfactual treatment groups, also each comprised a random and balanced subset of individuals in the study. The analysis uses the overlap, however small, between actual participants in different treatment groups to model what would happen in a balanced and randomized version of the study. The overlap is determined by the cases' propensity scores, that is, each case's probability of being in the specified treatment group, which in this study's case was Stage 1 (though any level of Stage could have been used).

More formally, let Y_1 denote a case's outcome if treated, and Y_0, the case's outcome in the counterfactual untreated condition. Then the estimate τ of the average treatment effect (ATE) is

$$\tau = E(Y_1) - E(Y_0).$$

The relationship of the outcomes observed in the actual experiment to Y_1 and Y_0 can be expressed. If Y is the response actually observed,

$$Y = Y_1 T + (1 - T)Y_0.$$

Finally, the ATE can be defined as

$$\tau = 1/n \sum_{i=1}^{n} \frac{T_i Y_i}{ps_i} - 1/n \sum_{i=1}^{n} \frac{(1 - Z_i)Y_i}{(1 - ps_i)}$$

(Lunceford & Davidian, 2004). To see how the covariate vector **X** is treated in this analysis, see Lunceford and Davidian (2004).

Each model was run on data on two program years of cases, with a given long-stay resident clinical outcome (or in the final analysis, the composite measure) as the dependent variable, the corresponding inverse-probability-weighted covariates as potential confounders, and the program facility stages as the predictor variables. There was no Year or Year × Stage term included in the analysis because Year and Stage were highly collinear, necessitating an analysis strategy that did not use both terms in the model. Both years' data were included in the analyses, however, and to account for clustering within Year, robust clustering variance adjustment was performed.

The average treatment effect is the effect of a given level of a treatment on all members of the study, given what was actually observed and in light of the study's imbalances (StataCorp, 2017). For each model, the average treatment effect of each higher stage compared with Stage 0 was estimated, with significance evaluated in terms of z-scores. The potential outcome means for each treatment level, that is, the means modeled to be observed in the randomized and balanced study for each treatment group (StataCorp, 2017)—given what was observed in the actual study—were also estimated.

Multiple-Test Correction

Three tests of statistical significance were conducted for each of the 13 separate clinical outcomes and for the composite outcome, a considerable number of tests, a small fraction of which might be falsely significant by chance. Thus, to mitigate the risk of Type I errors, the Benjamini–Hochberg false discovery rate (FDR) correction was applied to each p value (Benjamini & Hochberg, 1995; Benjamini & Yekutieli, 2001). See Supplementary Material for more details. In the Results section where the treatment effects analysis findings are presented, z-scores with FDR-corrected p values are provided.

Results

Table 3 presents the demographic characteristics of the included facilities in the baseline program year of 2014–2015. It can be seen that at baseline, and before propensity score adjustment, homes at different program stages differed significantly in their profit status, CCRC affiliation, chain ownership, occupancy, expected total hours of aide, LPN and RN staffing (case-mix index) and HHI. In that program year, 318 of 349 facilities either participated in PEAK ($N = 190$) or were Stage 0 nonparticipants ($N = 128$); the remaining 31 facilities either withdrew from the program or were dropped because of insufficient participation. Program year 2015–2016 was also included in the analyses. In 2015–2016, 328 of 349 facilities either participated ($N = 234$) or were Stage 0 nonparticipants ($N = 94$), with 21 withdrawing or being dropped. Thus, the main analyses included 318 homes for 2014–2015 and 328 homes for 2015–2016.

Figure 1a shows the potential outcome means for residents with clinically significant depressive symptoms at each stage. The analysis revealed a significantly lower prevalence of these symptoms in Stages 1–3, up to a 49% reduction in significant depressive symptoms for Stage 3 residents compared with residents in Stage 0 (nonparticipating) facilities. The average treatment effect of Stage 1 compared with Stage 0 was significant ($z = -2.80$, FDR-corrected $p = .0123$), as were Stage 2 compared with Stage 0 ($z = -5.55$, FDR-corrected $p = .00000$) and Stage 3 compared with Stage 0 ($z = -6.32$, FDR-corrected $p = .00000$). This outcome displayed a dose–response relationship, with successively higher adoption stages having a lower prevalence of residents with major depressive symptoms. Supplementary Material contains the rest of the analysis output from this analysis and all others.

Table 3. Demographic Characteristics of Kansas Facilities by Program Stage (Stage 0: Nonparticipants, Stage 1: Foundation Level, Stage 2: Level 1, Stage 3: Levels 2–5) for the Baseline Year of 2014–2015, With Covariate Associations With Stage

	Stage 0	Stage 1	Stage 2	Stage 3
N	128	95	34	61
Not-for-profit	27%***	39%	35%	59%
CCRC affiliated	13%***	21%	29%	39%
Chain owned	52%*	49%	38%	41%
Certified beds	63 (39)	65 (30)	72 (41)	74 (39)
Occupancy	76% (18%)***	81% (17%)	84% (11%)	88% (9%)
Medicaid residents	60% (38%)	53% (17%)	57% (19%)	53% (20%)
Private-pay-/insurance residents	35% (24%)	39% (18%)	31% (15%)	38% (19%)
Aide HRD	2.6 (0.73)	2.6 (0.64)	2.6 (0.58)	2.8 (0.66)
LPN HRD	0.64 (0.32)	0.58 (0.26)	0.74 (0.61)	0.60 (0.27)
RN HRD	0.90 (0.67)	0.74 (0.27)	0.78 (0.35)	0.78 (0.23)
Exp. Tot. Hrs.	3.9 (0.49)*	3.8 (0.36)	3.8 (0.22)	3.7 (0.35)
HHI	0.33 (0.31)*	0.37 (0.28)	0.31 (0.28)	0.22 (0.15)
Urban/rural	56% (43%)	48% (42%)	56% (43%)	59% (40%)

Notes: Exp. Tot. Hrs. = expected total hours of RN, LPN, and aide staffing given facilities' case mix (as a proxy for case-mix index); HHI = Herfindahl–Hirschman Index; HRD = hours per resident day; LPN = licensed practical nurse; RN = registered nurse. Percentages in the table represent the percentages of ones that facilities had for covariates with dummy coding.

*$p < .05$. ***$p < .005$.

Figure 1b presents the potential outcomes means for catheter use. Residents in Stage 1 facilities had a marginally lower prevalence than Stage 0 residents of indwelling catheters ($z = -2.01$, FDR-corrected $p = .08$), and their prevalence was significantly lower in residents of Stage 2 and Stage 3 facilities (respectively, $z = -7.07$, FDR-corrected $p = .00000$ and $z = -3.92$, $p = .00000$). This outcome also displayed an approximate dose–response relationship.

Figure 1c shows the potential outcomes means for high-risk residents with pressure ulcers. Their prevalence was not lower in Stage 1 residents compared with Stage 0 residents ($z = -1.11$, FDR-corrected $p = .1.53$). However, the prevalence of residents with pressure sores was significantly lower in Stage 2 and Stage 3 facilities compared with Stage 0 facilities (respectively, $z = -11.89$, $p = .00000$ and $z = -10.9$, $p = .00000$). This outcome also displayed an approximate dose–response relationship with PEAK adoption stages.

Figure 1d shows the potential outcomes means for residents being bound by physical restraints. The prevalence of restraint use was lower in Stage 1 facilities compared with Stage 0 facilities ($z = -3.86$, FDR-corrected $p = .00000$), as it was in Stage 3 facilities (respectively, $z = 8.78$, FDR-corrected $p = .00000$). Stage 2 facilities, however, did not differ in their use of physical restraints ($z = -1.11$, FDR-corrected $p = .153$). This outcome displayed a rough dose–response relationship with the exception of the Stage 2 results.

Figure 2a and b show the respective potential outcome means for residents receiving timely pneumococcal and influenza vaccinations. For pneumococcal vaccinations, compared with Stage 0 residents, those in Stage 3 homes were much more likely to have received timely vaccinations ($z = 26.1$ and FDR-corrected $p = .00000$); in contrast, Stage 1 and Stage 2 residents did not differ significantly from

Stage 0 residents ($z = 1.74$, FDR-corrected $p = .1479$, and $z = -0.33$, FDR-corrected $p = .94$). For influenza vaccinations, compared with residents in Stage 0 homes, residents in Stage 2 facilities were much likelier to have received them ($z = 3.62$, FDR-corrected $p = .00000$). Those in Stage 1 and Stage 3 facilities, however, did not differ from those in Stage 0 homes ($z = 1.72$, FDR-corrected $p = .15$ and $z = 0.53$, FDR-corrected $p = .63$). These outcomes may not have displayed a dose–response relationship with Stage because some homes may have given elders a choice to reject the vaccinations.
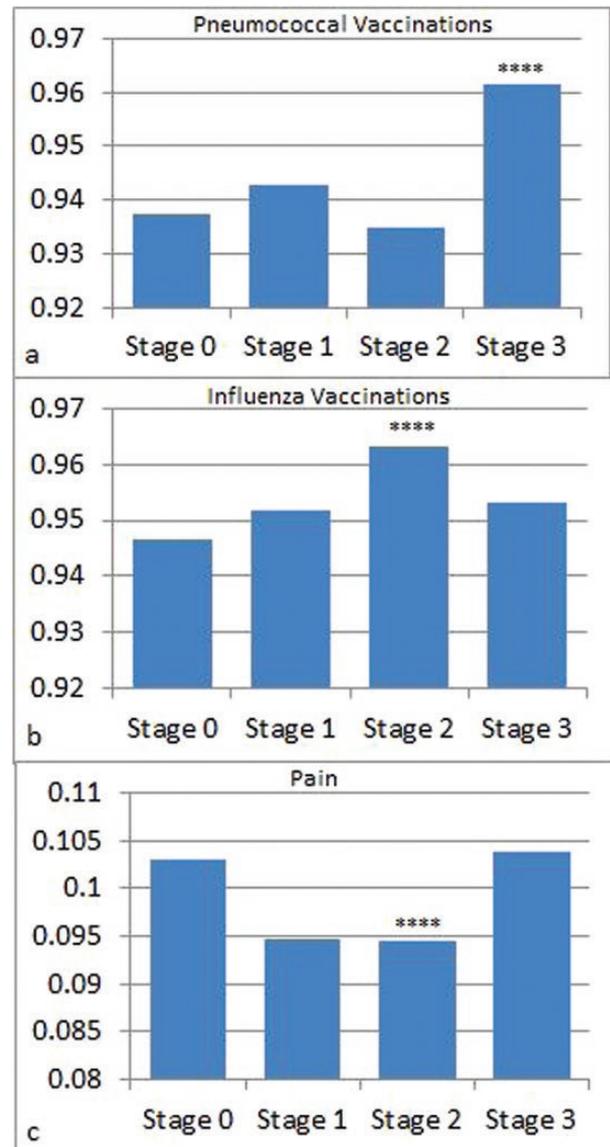
Figure 2c shows the respective potential outcome means for residents reporting moderate to severe pain. The prevalence of this complaint was lower in Stage 2 facilities compared with Stage 0 homes ($z = -2.41$, FDR-corrected $p = .037$). Those in Stage 1 and Stage 3 homes, however, were no less likely than those in Stage 0 homes to report this (respectively, $z = -1.35$, $p = .152$, and $z = 0.21$, $p = .96$). It is unclear why there was no semblance of a dose-response relationship with Stage.

In addition to the abovementioned findings of better clinical outcomes in higher-stage facilities, there were three potentially deleterious outcomes, concerning falls with major injury, activities of daily living (ADL) losses, and incontinent episodes in low-risk residents. As depicted in Figure 3a, residents in Stage 1 facilities fell more often than those in Stage 0 facilities ($z = 9.20$, $p = .00000$), and those in Stage 2 homes were marginally more likely to have a major fall ($z = 1.54$, FDR-corrected $p = .0756$). Stage 3 residents, however, were no more likely than Stage 0 residents to experience a major fall ($z = 1.04$, FDR-corrected $p = .42$). Furthermore, ADL losses were greater among residents in all three higher-stage homes compared with Stage 0
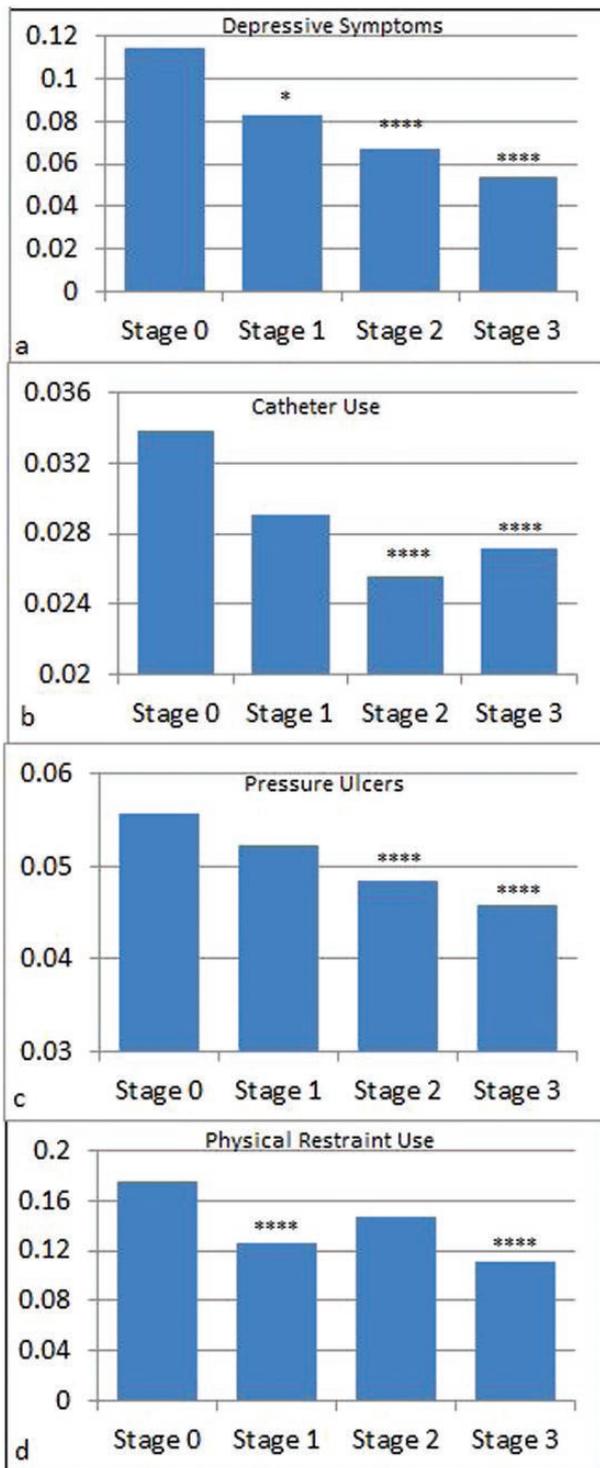
Figure 1. Potential outcome means by Stage for proportions of residents with major depressive symptoms (a), an indwelling catheter (b), pressure ulcers (c), and physical restraints (d), by program stage. Asterisks denote significance of average treatment effects of higher stages relative to Stage 0, as reported in the text. *False discovery rate (FDR) corrected $p < .05$ and ****FDR $p < .0001$.

Figure 2. Potential outcome means by Stage for proportions of residents receiving a timely pneumococcal vaccination (a), receiving a timely influenza vaccination (b), and reporting moderate to severe pain (c). Asterisks denote significance of average treatment effects of higher-level stages relative to Stage 0, as reported in the text. **** False discovery rate (FDR)-corrected $p < .0001$.

residents (respectively, $z = 7.05$, FDR-corrected $p = .00000$; $z = 3.52$, FDR-corrected $p = .00000$; and $z = 4.05$, FDR-corrected $p = .00000$; Figure 3b). Finally, low-risk residents

in Stage 2 homes were more likely to lose control of their bladder or bowels ($z = 4.05$, FDR-corrected $p = .00000$), although those in Stage 1 and Stage 3 homes did not differ from those in Stage 0 homes (respectively, $z = 0.41$, FDR-corrected $p = .63$ and $z = 0.24$, FDR-corrected $p = .93$; Figure 3c).

The remaining three individual clinical measures—antipsychotic prescribing, urinary tract infections (UTIs), and unexplained weight loss—showed no statistical relationship to PEAK program stages during the studied time. For antipsychotic use, relative to Stage 0, Stage 1 $z = 0.23$ and FDR-corrected $p = .94$, Stage 2 $z = -0.36$ and FDR-corrected $p = .86$, and Stage 3 $z = -1.09$ and FDR-corrected

p = .152. For UTIs, also relative to Stage 0, Stage 1 z = −1.53 and FDR-corrected p = .152, Stage 2 z = 1.57 and FDR-corrected p = .15, Stage 3 z = 0.24 and FDR-corrected p = .89. Finally, for unexplained weight loss, relative to Stage 0, Stage 1 z = 0.17 and FDR-corrected p = .96, Stage 2 z = −0.03 and p = .99, and Stage 3 z = 0.94 and p = .455.

Regarding the composite analysis that made use only of extant (nonimputed) outcome data, facilities were included if they had data for seven or more separate outcomes either or both program years. Supplementary Table 6 presents the demographics of the 217 facilities satisfying this condition in 2014–2015 and the 214 facilities satisfying this condition in 2015–2016. As shown in Figure 4, the treatment effects analysis of the composite measure revealed that relative to residents in Stage 0 homes, residents in Stage 1 homes had marginally greater overall health (as measured by the composite of 13 Nursing Home Compare outcomes; z = −1.94, FDR-corrected p = .099), whereas residents in Stage 2 and Stage 3 homes had significantly greater overall health (respectively, z = −5.53, FDR-corrected p = .00000 and z = −2.22, FDR-corrected p = .047). With the exception of the Stage 3 results, there was an apparent dose–response relationship between the Stage variable and residents' overall health.

Discussion

Despite the fact that PEAK 2.0 incents the adoption of PCC, not clinical improvements per se, and is intended to improve quality of life, not quality of care, the adoption of PCC as facilitated by the program was associated with better clinical outcomes for 7 of 13 MDS 3.0 long-stay resident health measures. Moreover, the composite analysis revealed better clinical outcomes at higher stages of the program, and thus with greater PCC adoption. The gains were clinically as well as statistically significant—suggesting practical importance for individual nursing homes and residents—for major depressive symptoms, pressure sores, and the use of physical restraints and indwelling catheters. There were three findings of negative outcomes, but one of them, falls, occurred only early in adoption and two of them, ADL losses and incontinent episodes, had subsided or were declining by late-stage adoption. They could have resulted from the initial tumult that may occur as the deep organizational change that PEAK requires is taking place. In addition, further study is needed because the main comparisons used here were cross-sectional. Overall, however, the finding of multiple significant improvements, present
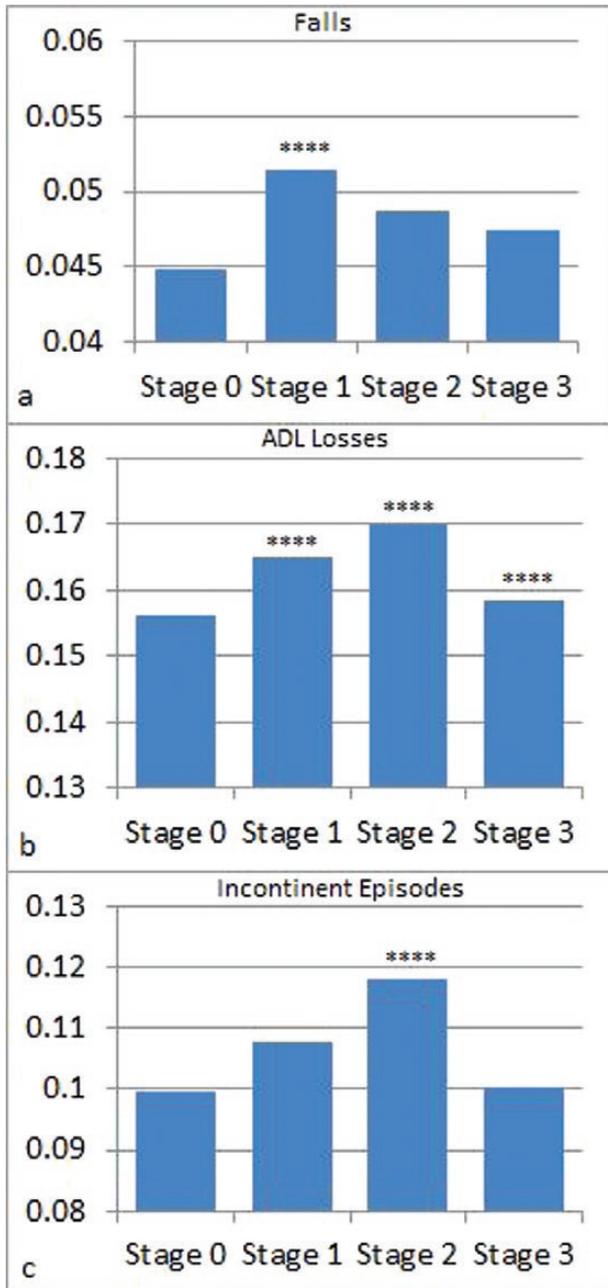
Figure 3. Potential outcome means by Stage for proportions of residents experiencing a fall with major injury (a), residents experiencing loss of one or more activities of daily living (ADLs; b), and low-risk residents experiencing incontinence of the bladder or bowels (c). Asterisks denote significance of average treatment effects of higher-level stages relative to Stage 0, as reported in the text. ****False discovery rate (FDR)-corrected p < .0001.
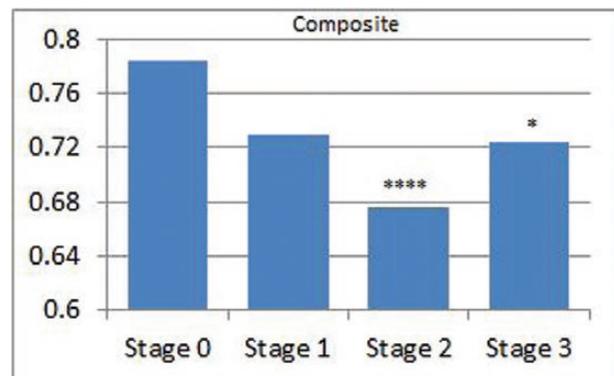
Figure 4. Potential outcome means by Stage for the composite measure of overall health. Asterisks denote significance of average treatment effects of higher-level stages relative to Stage 0, as reported in the text. *False discovery rate (FDR)-corrected p < .05 and ****FDR p < .0001.

when both imputed and unimputed data sets were used, and only a few potential drawbacks suggest that PEAK holds considerable promise as a means of improving the quality of care for nursing home residents.

Contributions to the Evidence Base

In contrast to past studies, the PEAK program allowed the examination of the clinical correlates of adopting PCC across well-defined, increasing levels of adoption, with participating homes having a standardized and operationalized definition of PCC, and program levels being determined by objective, outside evaluators. This study of PCC and clinical outcomes therefore probably had greater sensitivity than past studies to detect clinical benefits associated with varying degrees of adoption. To the authors' knowledge, it is the first large-scale study to find improvements in a majority of analyzed health outcomes. It is also the first study to the authors' knowledge to find a clear reduction in symptoms of clinical depression, a major scourge of nursing home life, with several major studies reporting the prevalence of symptoms of major depression or major depression itself at 40% or higher (e.g., Teresi et al., 2001).

Together with previous large-scale studies, patterns regarding which health outcomes PCC adoption may improve and which it does not more generally are beginning to emerge. For example, as with several other large-scale studies (Afendulis et al., 2016; Miller, Lepore, Lima, et al., 2014), this study found a lower prevalence of pressure ulcers in high-risk residents and in the use of physical restraints.

The finding of a greater number of falls early in PCC adoption merits comment as well, as it coheres with at least two earlier studies (Chang et al., 2013; Chenoweth et al., 2009) and the findings of a recent systematic review (Brownie & Nancarrow, 2014). In the present study, falls occurred far more often in Stage 1 relative to Stage 0, but were not elevated in Stage 2 or Stage 3. This pattern along with the results of earlier studies suggests that there is either initial disorganization resulting from early culture change adoption or that elders are given more freedom early in culture change adoption without compensatory measures to prevent falls. Either way, if nursing home management is aware of these issues they can mitigate the risk of falls, without interfering with any newfound freedoms that elders deserve.

The PEAK program's higher stages' association with better clinical outcomes also contrasts with most studies of nursing home P4P programs to improve quality, in which the few observed gains were usually too small to be clinically significant and as many apparently negative effects were found as benefits (Briesacher et al., 2009; Grabowski et al., 2017; Weissert & Frederick, 2013; Werner et al., 2013). Among the reasons for PEAK's apparent success may be its design in accordance with recommendations made after study of failed P4P programs, including:

- incorporation of training and guidance for facility management and staff on how to undertake quality improvement, rather than assuming they know how to do it on their own;
- the provision of regular feedback on facilities' progress on achieving targeted outcomes; and
- the inclusion of a readily understood and substantial financial incentive (Green, 2013).

Policy Implications

This study of the PEAK 2.0 program has important implications for the future of nursing home policy at the federal and state levels. The recently issued Final Rule by CMS requires the adoption of PCC in select areas, many of which overlap with the PEAK criteria. The research findings here support the new regulatory emphasis on PCC, but suggest it does not go far enough.

PEAK 2.0 may be of great interest to states experimenting with value-based payment reform in their Medicaid programs. As Congress and the federal executive branch explore a greater role for states through block granting or capping Medicaid, state and federal policymakers may see the value of incorporating a P4P incentive program modeled on PEAK 2.0. State and federal surveyors and regulators should consider investing in a technical assistance approach similar to the one used in Kansas, where a team trains care providers and long-term care administrators in the fundamentals of PCC. Then, over time, it should introduce additional elements of PCC implementation and assist providers having difficulty with one or more phases of PCC adoption. In addition, homes should be provided with objective feedback on their progress rather than relying on their self-assessments, which could be done by relatively inexpensive means, for example, video conferences. Finally, policymakers should offer financial incentives that are sizable enough to attract for-profit nursing homes, as the Kansas program has succeeded in doing (Hermer et al., 2017).

Although each state varies in regulatory environment and in state agency motivation to improve resident outcomes, the guidelines for PEAK are easily adapted either at the individual or at the state policy level. PEAK team leaders have worked to get buy-in from state nursing home associations, advocacy groups, and other important stakeholders, which is a key element in garnering support for adoption of a program such as this.

Limitations

This study improved on prior studies of PCC and resident health in several ways, but it also had several limitations. First, although the study was technically longitudinal, its main comparisons were cross-sectional, and further study will be needed to determine causality between culture change adoption at increasing PEAK stages and improved clinical

outcomes. Second, the study was observational. Though advanced methods were used to control for observed differences across facilities at different stages, the lack of true randomization left the possibility that facilities differed in unobserved ways, such as in their profit margins or turnover rates. Third, the Nursing Home Compare data used here were derived from CMS's Minimum Dataset 3.0, and CMS had converted some indicators to moving averages, potentially reducing their timeliness (Abt Associates, 2016). Fourth, some of the Nursing Home Compare outcome data were high in missing data, but the composite analysis using only extant (nonimputed) data strongly supported the main analyses of individual measures after multivariate imputation. Fifth, the aggregation of homes in Levels 2 through 5 (in Stage 3) could have obscured a more nuanced understanding of the differences that occur during the process of PCC adoption as well as the demographic differences found in homes at later stages. Homes at higher levels, particularly Level 5, have also been practicing PCC for a longer period of time, and the implication of this could not be captured. A related issue is that not all the outcomes displayed a dose–response relationship. However, it should be noted that in separate analyses, Level 3–5 homes almost always showed the best clinical outcomes, though there was confounding with some demographic variables. Finally, facilities participating in PEAK adopted PCC in a semiregular, externally guided, and supervised manner and may therefore not have represented the full range of culture change/PCC adopters across the United States. However, given that substantial benefits of adopting PCC as facilitated by PEAK that were found, this should be considered a strength because PEAK's standardized methods and materials can become a blueprint for other facilities and states.

Summary

This large-scale study was among the very first to reveal that a majority of nursing home resident quality indicators were better in facilities adopting PCC, aided by the guidance of PEAK program staff and objective measures of PCC adoption. The results suggest that when facility staff understand PCC and hold an accurate view of their progress with implementation, PCC may markedly improve resident care. Recently, it was also shown that higher PEAK stages were associated with greater resident satisfaction with their quality of life and care (Poey et al., 2017). With the PEAK program's successes, it could serve as a model for other states.

Supplementary Material

Supplementary data are available at *Innovation in Aging* online.

Funding

Acknowledgments

Conflict of Interest

None reported.

References

Abt Associates. (2016). *Nursing home compare quality measure technical specifications*. Baltimore, MD: CMS.

Afendulis, C. C., Caudry, D. J., O'Malley, A. J., Kemper, P., Grabowski, D. C., & Collaborative, T. R. (2016). Green house adoption and nursing home quality. *Health Services Research*, **51** (Suppl 1), 454–474. doi:10.1111/1475–6773.12436

Arling, G., Job, C., & Cooke, V. (2009). Medicaid nursing home pay for performance: Where do we stand? *The Gerontologist*, **49**, 587–595. doi:10.1093/geront/gnp044

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B*, **57**, 289–300. doi:10.3102/10769986025001060

Benjamini, Y., & Yekutieli, D. (2001). The control of the false discovery rate in multiple testing under dependency. *Journal of Mathematical Statistics*, **29**, 1169–1188. doi:10.1214/aos/1013699998

Briesacher, B. A., Field, T. S., Baril, J., & Gurwitz, J. H. (2009). Pay-for-performance in nursing homes. *Health Care Financing Review*, **30**, 1–13. doi:10.1111/j.1532-5415.2008.01920.x

Brownie, S. & Nancarrow, S. (2014). Effects of person-centered care on residents and staff in aged care facilities: A systematic review. *Clinical Interventions in Aging* 2013, 1–10. doi:10.2147/CIA.S38589

Burack, O. R., Weiner, A. S., Reinhardt, J. P., & Annunziato, R. A. (2012). What matters most to nursing home elders: Quality of life in the nursing home. *Journal of the American Medical Directors Association*, **13**, 48–53. doi:10.1016/j.jamda.2010.08.002

Chang, Y. P., Li, J., & Porock, D. (2013). The effect on nursing home resident outcomes of creating a household within a traditional structure. *Journal of the American Medical Directors Association*, **14**, 293–299. doi:10.1016/j.jamda.2013.01.013

Chenoweth, L., King, M. T., Yun-Hee, J., Henry Brodaty, D., Stein-Parbury, J., Norman, R. M., … Luscombe, G. (2009). Caring for Aged Dementia Care Resident Study (CADRES) of person-centred care, dementia-care mapping, and usual care in dementia: A cluster-randomised trial. *The Lancet Neurology*, **8**, 317–325. doi:10.1016/S1474-4422(09)70045–6

Doll, G. A., Cornelison, L. J., Rath, H., & Syme, M. L. (2017). Actualizing culture change: The promoting excellent alternatives in Kansas nursing homes (PEAK 2.0) program. *Psychological Services*, **14**, 307–315. doi:10.1037/ser0000142

Emmert, M., Eijkenaar, F., Kemter, H., Esslinger, A. S., & Schoffski, O. (2012). Economic evaluation of pay-for-performance in health care: A systematic review. *The European Journal of Health Economics*, **13**, 755–767. doi:10.1007/s10198-011-0329-8

Government Accountability Office. (2005). *Nursing homes: Despite increased oversight, challenges remain in ensuring high-quality care and resident safety*. Washington, DC: Government Accountability Office.

Grabowski, D. C., O'Malley, A. J., Afendulis, C. C., Caudry, D. J., Elliot, A., & Zimmerman, S. (2014). Culture change and nursing home quality of care. *The Gerontologist*, **54** (Suppl 1), S35–S45. doi:10.1093/geront/gnt143

Grabowski, D. C., Stevenson, D. G., Caudry, D. J., O'Malley, A. J., Green, L. H., Doherty, J. A., & Frank, R. G. (2017). The impact of nursing home pay-for-performance on quality and medicare spending: Results from the nursing home value-based purchasing demonstration. *Health Services Research*, **52**, 1387–1408. doi:10.1111/1475-6773.12538

Green, L. H. (2013). *Evaluation of the nursing home value-based purchasing demonstration*. Washington, DC: Centers for Medicare and Medicaid Services.

Hermer, L., Cornelison, L., Kaup, M. L., Poey, J. L., Stone, R. I., & Doll, G. (2017). The Kansas PEAK 2.0 program facilitates the diffusion of culture change innovation to unlikely adopters. *The Gerontologist*. 58: 530–539. doi:10.1093/geront/gnw210

Kim, S. K., & Park, M. (2017). Effectiveness of person-centered care on people with dementia: A systematic review and meta-analysis. *Clinical Interventions in Aging*, **12**, 381–397. doi:10.2147/CIA.S117637

Koren, M. J. (2010). Person-centered care for nursing home residents: The culture-change movement. *Health Affairs*, **29**, 312–317. doi:10.1377/hlthaff.2009.0966

Lunceford, J. K., & Davidian, M. (2004). Stratification and weighting via the propensity score in estimation of causal treatment effects: A comparative study. *Statistics in Medicine*, **23**, 2937–2960. doi:10.1002/sim.1903

Markovitz, A. A., & Ryan, A. M. (2016). Pay-for-performance: Disappointing results or masked heterogeneity? *Medical Care Research and Review*, **74**, 3–78. doi:10.1177/1077558715619282

Mendelson, A., Kondo, K., Damberg, C., Low, A., Motúapuaka, M., Freeman, M., …Kansagara, D. (2017). The effects of pay-for-performance programs on health, health care use, and processes of care: A systematic review. *Annals of Internal Medicine*, **166**, 341–353. doi:10.7326/M16-1881

Miller, S. C., Lepore, M., Lima, J. C., Shield, R., & Tyler, D. A. (2014). Does the introduction of nursing home culture change practices improve quality? *Journal of the American Geriatrics Society*, **62**, 1675–1682. doi:10.1111/jgs.12987

Miller, S. C., Looze, J., Shield, R., Clark, M. A., Lepore, M., Tyler, D., …Mor, V. (2014). Culture change practice in U.S. nursing homes: Prevalence and variation by state Medicaid reimbursement policies. *The Gerontologist*, **54**, 434–445. doi:10.1093/geront/gnt020

Poey, J. L., Hermer, L., Cornelison, L., Kaup, M. L., Drake, P., Stone, R. I., & Doll, G. (2017). Does person-centered care improve residents' satisfaction with nursing home quality? *Journal of the American Medical Directors Association*, **18**, 974–979. doi:10.1016/j.jamda.2017.06.007

Rahman, A. N., Applebaum, R. A., Schnelle, J. F., & Simmons, S. F. (2012). Translating research into practice in nursing homes: Can we close the gap? *The Gerontologist*, **52**, 597–606. doi:10.1093/geront/gnr157

Rogers, E. M. (2003). *Diffusion of innovations* (5th ed.). New York, NY: Free Press.

RTI (2014). *MDS 3.0 Quality Measures: User's Manual v.9*. Washington, DC: Centers for Medicare and Medicaid Services.

Shier, V., Khodyakov, D., Cohen, L. W., Zimmerman, S., & Saliba, D. (2014). What does the evidence really say about culture change in nursing homes? *The Gerontologist*, **54** (Suppl 1), S6–S16. doi:10.1093/geront/gnt147

StataCorp. (2017). *Treatment effects reference manual*. College Station, TX: Stata Press.

Stone, R. I., Bryant, N., & Barbarotta, L. (2009). Supporting culture change: Working toward smarter state nursing home regulation. *Issue Brief (Commonwealth Fund)*, **68**, 1–10. Retrieved from https://www.researchgate.net/profile/Robyn_Stone/publication/38020250_Supporting_culture_change_working_toward_smarter_state_nursing_home_regulation/links/00b4951894f72ec4cb000000.pdf

Stone, R.I., Reinhard, S.C., Bowers, B., Zimmerman, D., Phillips, C.D., Hawes, C., Fielding, J.A., & Jacobsen, N (2002). *Evaluation of the Wellspring Model for Nursing Home Quality*. New York, NY: The Commonwealth Fund, August 2002. Retrieved from www.cmwf.org

Sullivan, J. L., Shwartz, M., Stolzmann, K., Afable, M. K., & Burgess, J. F. (2018). A longitudinal assessment of the effect of resident-centered care on quality in veterans health administration community living centers. *Health Services Research*, **53**, 1819–1833. doi:10.1111/1475-6773.12688

Teresi, J., Abrams, R., Holmes, D., Ramirez, M., & Eimicke, J. (2001). Prevalence of depression and depression recognition in nursing homes. *Social Psychiatry and Psychiatric Epidemiology*, **36**, 613–620. doi:10.1007/s127-001-8202-7

Tyler, D. A., Shield, R. R., Rosenthal, M., Miller, S. C., Wetle, T., & Clark, M. A. (2011). How valid are the responses to nursing home survey questions? Some issues and concerns. *The Gerontologist*, **51**, 201–211. doi:10.1093/geront/gnq095

van Buuren, S. & Groothuis-Oudshourn K. (2010). MICE: Multivariate imputation with chained equations in R. *Journal of Statistical Software*, **20**.

Weissert, W. G., & Frederick, L. F. (2013). Pay for performance: Can it help improve nursing home quality? *Public Administration Review*, **73** (Suppl 1), 11. doi:10.1111/puar.12074

Werner, R. M., Konetzka, R. T., & Polsky, D. (2013). The effect of pay-for-performance in nursing homes: Evidence from state medicaid programs. *Health Services Research*, **48**, 1393–1414. doi:10.1111/1475-6773.12035

White-Chu, E. F., Graves, W. J., Godfrey, S. M., Bonner, A., & Sloane, P. (2009). Beyond the medical model: The culture change revolution in long-term care. *Journal of the American Medical Directors Association*, **10**, 370–378. doi:10.1016/j.jamda.2009.04.004

Zimmerman, S., Bowers, B. J., Cohen, L. W., Grabowski, D. C., Horn, S. D., Kemper, P., & Collaborative, T. R. (2016). New evidence on the green house model of nursing home care: Synthesis of findings and implications for policy, practice, and research. *Health Services Research*, **51** (Suppl 1), 475–496. doi:10.1111/1475–6773.12430