

Supplementary Appendix

This appendix has been provided by the authors to give readers additional information about their work.

Supplement to: Gozalo P, Plotzke M, Mor V, Miller SC, Teno JM. Changes in Medicare costs with the growth of hospice care in nursing homes. *N Engl J Med* 2015;372:1823-31. DOI: 10.1056/NEJMsa1408705

On-line Only Supplemental Material:

Increased Medicare Costs with the Growth of Hospice Care in Nursing Homes

Pedro Gozalo, PhD, MSc,

Michael Plotzke, PhD,

Vincent Mor, PhD,

Susan C. Miller, PhD,

Joan M. Teno, MD, MS,

Table of Contents:

Title	Page
Overall Methodological Strategy	2
Implementation of the Cross-Temporal Propensity Score Matching Estimation Strategy	2
Selection of Covariates for the Hospice Propensity Score Models	4
Adequacy of the Propensity Score Matching	4
Censoring Considerations	6
Sensitivity Analysis	7
Table S1. Comparison of Nursing Home Decedents in 2004 and 2009, by Hospice Election Status	8
Table S2. Propensity Score Models of Hospice Election for 2004 and 2009 Nursing Home Decedents	10
Table S3. Comparison of NH Decedent and NH Facility Characteristics by Matching Group	12
Table S4. Comparison of Nursing Home Decedents in 2004 and 2009, by Fee-For-Service (FFS) versus Medicare Advantage (HMO) Enrollment Status	14
Table S5. Association of Hospice and Total Medicare Expenditures in the Last Year of Life and End-of-Life Care Outcomes for Decedents in Nursing Homes with No Hospice in 2004 but using Hospice in 2009, All Diagnoses	15
Figure S1: Covariate Balance of Patient Characteristics for Group 2 Consisting of NH Decedents that Elected Hospice in 2009 and their 1:1 Matched Non-Hospice 2004 NH Decedents	16
References	16

Overall Methodological Strategy

Our approach to minimize the effect of selection bias which is always present in comparisons of hospice and non-hospice patient outcomes, takes advantage of the natural experiment created by the substantial growth in the use of Hospice services among nursing home (NH) residents between 2004 and 2009. Rather than matching hospice decedents in 2009 to similar non-hospice decedents in 2009 (cross-sectional matching), we matched hospice decedents in 2009 cross-temporally to similar non-hospice decedents in 2004 in accordance with the strategy outlined below. This cross-temporal matching has the clear advantage that we are more likely to find a non-hospice match for our 2009 hospice user that has similar characteristics (such as preferences for aggressive end of life care), thus reducing the selection bias.

A difference-in-differences (DID) approach is then used to account for changes over time (secular trends) in factors (other than the change in a patient's hospice status) which could impact differences in outcomes over time.

Implementation of the Cross-Temporal Propensity Score Matching Estimation Strategy

To explain the steps of our matching strategy, recall from the paper that we have the following three potential groups of individuals:

1. Group 1 (G1) consists of NH residents that used hospice in the 2004 baseline year and their matches are hospice users in the 2009 post-baseline year. These 2009 hospice users represent the 'traditional' hospice users in the sense that they would have elected hospice in the baseline year when exposure was lower. Because these individuals are only observed as electing hospice in both years, we cannot use them to infer what their outcomes would have been under no hospice use and thus they are not used in our hospice-outcomes analyses.
2. Group 2 (G2) is the critical group for identifying the relationship of hospice on expenditures. It consists of individuals that did not elect hospice in the 2004 baseline year matched to those electing hospice in the 2009 post-baseline year due to the larger access to hospice care (those with hospice in 2009 not in G1). For this group of "new" hospice users, the baseline year decedents provide the outcome under no exposure to hospice, while the post-baseline year decedents provide the outcome under exposure to hospice.
3. Group 3 (G3) consists of decedents that did not elect hospice in either year. The changes in outcomes between years for this group will be used to account for secular trends, other than the change over time in hospice exposure, which could have impacted the outcomes for Group 2 between these two years.

The steps required to implement the cross-temporal matching strategy are as follows (see Figure 1 in the paper for additional guidance):

1. **Separation of Post-Baseline Hospice Users into Group 1 and Group 2** (see Step 1 in Figure 1 of the paper):
 - 1.1. Estimate a propensity score for hospice use with our 2004 baseline cohort, p_{2004} .
 - 1.2. Use the p_{2004} model to predict the propensity score for the 2009 NH decedent cohort.
 - 1.3. Use these propensity scores to find 1:1 matches (with replacement – as described below) for the 2004 hospice decedents among the 2009 decedents ($G1_{2009}$ in Figure 1). Note: to further help comparability across time for this group, we stratified decedents in both years into quintiles of hospice length of stay and carried out the matching within strata. Sensitivity analyses were carried out with different numbers of hospice LOS strata and quintiles were chosen as they provided better balance of the resulting matched cohorts.
 - 1.4. The 2009 hospice decedents left unmatched when creating Group 1 become our predicted “new” hospice group in Group 2 ($G2_{2009}$ in Figure 1).
2. **Construction of Group 2:**
 - 2.1. Estimate a propensity score with the 2009 post-baseline cohort, p_{2009} .
 - 2.2. Use the p_{2009} model to predict the propensity score for the 2004 NH decedent cohort.
 - 2.3. Use these propensity scores to find 1:1 matches (with replacement) for the individuals in $G2_{2009}$ among 2004 non-hospice decedents ($G2_{2004}$ in Figure 1). For the sub-analysis involving NH facilities that had no NH decedent electing hospice in the 2004 baseline year we used a 1:2 nearest neighbor match within NH facility.
3. **Construction of Group 3:**
 - 3.1. Use p_{2009} to find 1:1 matches (with replacement) for the post-baseline 2009 non-hospice decedents ($G3_{2009}$ in Figure 1) among 2004 non-hospice decedents ($G3_{2004}$ in Figure 1).

The 1:1 matching was done using the nearest neighbor (closest) propensity score method, and it was done with replacement to allow the selection of the closest propensity score for each person to be done from the entire pool of potential matching candidates. Weights were then used to account for a candidate being selected multiple times. Because the matching is done with replacement, there were 2004 non-hospice decedents that were not selected into the $G2_{2004}$ or $G3_{2004}$ groups as best matches by any of the 2009 decedents in $G2_{2009}$ or $G3_{2009}$, and they are labeled as unmatched (UNM_{2004}) in Table S3.

Groups 2 and 3 are the two groups used to carry out the DID analysis with the following regression model

$$Y = \alpha_0 + \alpha_1 G2 + \alpha_2 Post + \delta G2 * Post + \beta X + \epsilon$$

where $G2$ is a 0-1 indicator of whether the individual belonged to Group 2 (either $G2_{2004}$ or $G2_{2009}$), $Post$ is a 0-1 indicator of the observation corresponding to the post-baseline year, X is the list of patient and facility covariates used in the propensity models and decedent hospital referral region fixed effects to adjust for residual imbalance in our matched Groups 2 and 3, and ϵ is the error term. The desired hospice-outcome relationship is given by the coefficient of the interaction term δ , which estimates the changes in outcomes over time for Group 2 minus the changes in outcomes over time for Group 3, $(G2_{2009}-G2_{2004}) - (G3_{2009}-G3_{2004})$.

The fact that individuals in Group 3 continue to not elect hospice when it is more widely available implies that these individuals are more likely to prefer aggressive care and will therefore have higher expenditures than would “new” hospice users, had hospice not been as readily available. As such, we may be overcorrecting our hospice cost estimates and the true effect may be larger than what we report in our results.

Selection of Covariates for the Hospice Propensity Score Models

We selected patient and NH facility characteristics (see Table S1) that have been found in the literature to be associated with health care utilization.¹ Continuous covariates, such as age, and the 0-28 range activities of daily living (ADL) physical impairment scale and the 0-6 range cognitive performance scale (CPS), were modeled and tested as linear, categorical and non-linear cubic splines. We opted for categorizing age at death, and the ADL scale into deciles, while the CPS measure was categorized into 7 categories, one for each value in the scale. We used cancer and Alzheimer’s/other-dementia to create a 2x2 major diagnosis categorization of NH decedents into patients with one but not the other, both, or none of the two because, cancer and Alzheimer’s/other-dementia are the most prevalent hospice diagnoses (71% of our 2004 and 2009 cohorts) and patients with these two diagnoses tend to have the highest (cancer) and lowest (dementia) utilization, expenditures and levels of aggressive care. Nursing facility characteristics were included in the propensity score construction since they can be associated with hospitalization rates and other aggressiveness of care outcomes. The quarter of the date of death was used to account for seasonality in utilization and expenditure patterns, such as the higher rates of hospitalizations in the winter months associated with influenza and other respiratory conditions. Table S1 provides the characteristics of the NH decedents in each year with their breakdown by hospice election status, before matching takes place. The last two columns provide the differences between non-hospice and hospice groups within each cohort year.

Adequacy of the Propensity Score Matching

We assessed the quality of the propensity score matching in two steps. First we checked to see whether the factors associated with the hospice election propensity scores changed in 2004 compared to the 2009 cohort and verified that the propensity scores were balanced across hospice and non-hospice subgroups for each of our three major groups

(Group 1- Group 3). Second, we verified that the covariates were properly balanced across the baseline and post-baseline subgroups for each of our matched groups.

The results of the logistic regressions used to construct the two propensity scores, p_{2004} and p_{2009} , are shown in Table S2. The magnitude of the changes across years in several odds ratios justifies our use of the baseline propensity score model, p_{2004} , for separating 2009 hospice users into Group 1 and Group 2. The c-statistics (area under the ROC curve) for both propensity score models are around 0.7 (see last row Table S2), indicating adequate sensitivity/specificity. We verified that the baseline and post-baseline matched subgroups for Groups 1-3 had common support and balanced propensity scores. Table S3 provides (row 5 from bottom) the 99th percentile of the differences in propensity score for each 1:1 matched pair in a given group. All are fairly small indicating that the vast majority of persons in each group were matched to an individual with a very similar propensity score.

The most critical validation is to verify that after the propensity score matching the covariates are properly balanced within each of the three groups. Table S3 provides the averages of the covariates for each group in the first seven columns, while the last three columns display the standardized % differences, defined as the % difference of the sample means in the treated (original subgroup) and non-treated (matched subgroup) sub-samples of a given group, standardized by the square root of the average of the sample variances in the treated and non-treated sub-samples of a given group.² This is a better statistic for judging individual covariate balance than the t-test because is not influenced by sample size. For example, for Group 2, the treated group refers to the “new” hospice decedents in 2009, G_{2009} , while the non-treated are the entire group of 2004 non-hospice decedents from which we find the 1:1 matches G_{2004} . There is no agreement criterion as to what threshold of the standardized difference should be used to indicate imbalance.³ Some studies have used a standard difference less than 0.1 to indicate a negligible difference in the mean or prevalence of a covariate between different treatment groups.⁴ Figure S1 provides a graphical representation of the standardized % differences for the patient-level covariates of decedents in Group 2 after matching (previous to last column in Table S2) and before matching (unmatched sample) where the improvement in covariate balance after matching is apparent.

In addition to the evaluation of balance across individual covariates, the last four rows of Table S3 display additional statistics summarizing the overall quality of the match:

- Rows three and four from the bottom provide the mean (over all covariates) differences, an overall measure of covariate imbalance, in the unmatched and matched samples of each group. Of note is the large reduction in mean differences for Group 2 in the matched sample relative to the unmatched sample. For Groups 1 and 3 the improvement is considerably smaller since the treated and non-treated subgroups are both either hospice (Group 1) or (Group 3) non-hospice users, and is natural to observe similar aggregate differences before and after matching.

- The last two rows: Rubin's B statistic (the absolute standardized difference of the means of the linear index of the propensity score in the treated and (matched) non-treated subgroups) and Rubin's R statistic (the ratio of the variances of the treated to (matched) non-treated variances of the propensity score index).⁵ For the samples to be considered sufficiently balanced it is recommended that B be 25% or lower and that R be between 0.5 and 2. Our B-statistics are all at the 25% mark.

Residual imbalance after matching, such as that observed in the Cognitive Performance Scale, diabetes and depression, was controlled by including all the patient and facility matching model covariates in the DID regression model.

Censoring Considerations

Our analysis relies on Medicare claims as the source of Medicare expenditures in the last year of life. These claims are available for beneficiaries enrolled in the traditional Medicare fee-for-service (FFS) insurance plan, however, for beneficiaries that enroll in a Medicare Advantage HMO insurance plan we only observed their claims once they elect hospice, as the hospice benefit is the only benefit not covered by Medicare Advantage HMO insurance. Therefore, election of Medicare Advantage insurance introduces censoring of Medicare expenditures. In our decedent cohorts, the number of NH decedents age 67 or older that were enrolled in Medicare Advantage at some point during their last two years of life was 51,027 (10.6%) in 2004 and 78,483 (17.7%) in 2009. This increase mirrors similar increases during that period among community dwelling Medicare beneficiaries. We assessed the impact of this in our analysis in two ways. First, we examined the decedent characteristics to determine if the FFS cohort we used in our analyses exhibited significant differences relative to the full cohort that included both the FFS and the Medicare Advantage decedents. For the great majority of characteristics the differences were very small (see Table S4). Second, we carried out a sensitivity analysis where we formally control for Medicare Advantage censoring. This is done by calculating first a logistic model, separately for each year, where the outcome is Medicare Advantage enrollment, given the covariates in our table. These propensity scores are then used to create inverse probability of censoring (IPC) weights defined for each person as the inverse of their probability of not being censored. These annual-specific weights have the property of making the averages of the covariates in the FFS censored sample in a given year look like the corresponding uncensored population sample. These annual IPC weights were used to calculate the annual hospice logistic models used to create the propensity scores for each year, p_{2004} and p_{2009} , so that the matching is done as if we had no censoring in the cohorts of each year. The results using these censoring-corrected propensity scores yielded almost identical results, \$6,703 compared to the \$6,761 we estimated; a less than 1 percent difference that was not statistically significant.

Sensitivity Analysis

As a sensitivity analysis of our Medicare spending results shown in Table 2 of the paper, we replicated our analysis using N=52,547 decedents from 1,106 facilities that had ten or more decedents in both 2004 and 2009 but in which no decedent elected hospice in 2004. Hospice utilization was 18.9% in 2009 for this subgroup. Decedents in 2009 were matched 1:2 to 2004 decedents *within* the same NH facility. The resulting net increase in Medicare reimbursements for hospice users was \$6,934 (95% CI; \$5,444, \$8,424) (see Table S5), very similar to our overall \$6,761 estimate.

Table S1. Comparison of Nursing Home Decedents in 2004 and 2009, by Hospice Election Status

Covariate	2004			2009			Differences between Non-Hospice and Hospice	
	No Hospice	Hospice	Total	No Hospice	Hospice	Total	2004	2009
N	308,418	117,858	426,276	216,658	143,394	360,052		
Age at Death	85.2	85.2	85.2	85.4	85.9	85.6	0.0	-0.6
Male	35.6	31.5	34.4	37.4	32.0	35.2	4.1	5.3
Nonwhite	10.8	9.5	10.4	12.3	10.3	11.5	1.3	2.0
Married	25.4	24.5	25.1	26.6	24.3	25.7	0.9	2.3
Major Diagnosis Group								
Cancer, no Dementia	13.4	18.6	14.9	13.7	14.5	14.1	-5.2	-0.8
Cancer and Dementia	8.1	12.5	9.4	8.3	12.4	9.9	-4.4	-4.1
Dementia, no Cancer	45.1	46.6	45.5	45.2	51.9	47.9	-1.4	-6.7
No Cancer and No Dementia	33.3	22.3	30.3	32.8	21.1	28.1	11.0	11.7
Other Important Diagnoses								
Emphysema/COPD	34.7	31.9	33.9	36.3	32.3	34.7	2.7	4.0
CHF	51.9	46.2	50.3	50.9	44.5	48.4	5.8	6.3
Diabetes	28.2	25.8	27.5	32.7	29.2	31.3	2.3	3.5
ASHD	16.0	14.7	15.6	16.9	15.5	16.3	1.3	1.4
PVD	13.5	12.9	13.3	13.9	13.4	13.7	0.6	0.5
CVA	22.4	21.8	22.2	19.5	19.5	19.5	0.6	0.0
Depression	36.9	42.8	38.5	40.5	48.3	43.6	-5.9	-7.8
Asthma	2.5	2.4	2.5	3.1	3.1	3.1	0.2	0.0
Resistant infection	4.2	3.1	3.9	4.6	3.4	4.1	1.1	1.2
Pneumonia	13.0	9.6	12.1	13.2	9.2	11.6	3.4	3.9
Respiratory infection	4.2	3.7	4.1	3.7	3.2	3.5	0.5	0.5
Septicemia	2.2	1.6	2.0	2.3	1.5	2.0	0.6	0.7
Cognitive Performance Scale (0-6 range; 0=intact, 6=very severe with eating problems)								
	3.3	3.5	3.4	3.1	3.5	3.2	-0.2	-0.4
ADL score (0-28 range; 0=independent, 28=totally dependent in all seven ADLs)								
	21.2	21.7	21.3	21.0	21.8	21.3	-0.6	-0.7
ADL worsening since last MDS								
DNH	17.5	30.4	21.1	14.5	27.0	19.5	-12.9	-12.6
DNR	6.7	8.7	7.2	7.4	9.2	8.2	-2.0	-1.8
Short NH stay (<30 days)	67.8	75.8	70.0	66.4	77.4	70.8	-8.0	-11.0
Long NH stay (>90 days)	17.4	8.7	15.0	16.7	6.9	12.8	8.6	9.8
Days from last MDS to Death	65.2	76.4	68.3	64.3	79.9	70.5	-11.2	-15.6
No. of Hospitalizations in the year prior to the year of death	26.7	33.9	28.7	25.2	34.3	28.9	-7.2	-9.1
Average Hospitalizations LOS	0.7	0.7	0.7	0.7	0.8	0.7	-0.1	0.0
NH FACILITY COVARIATES	2.1	2.3	2.2	2.1	2.2	2.1	-0.1	-0.1
% Medicaid NH residents	61.3	59.4	60.8	59.5	57.3	58.7	1.8	2.2
% Medicare NH residents	15.8	13.8	15.3	17.4	15.7	16.7	2.0	1.7
% Private Pay NH residents	22.9	26.8	24.0	23.1	26.9	24.6	-3.8	-3.8
NH has special care unit	26.9	31.1	28.1	24.3	27.1	25.4	-4.2	-2.8
For-profit Status	64.6	66.5	65.1	67.2	67.1	67.1	-1.9	0.1

Hospital-Based	7.5	3.4	6.4	5.6	2.9	4.5	4.1	2.8
Multisite facility (chain)	53.6	55.3	54.1	53.7	55.9	54.6	-1.6	-2.2
Last MDS in Quarter 1	27.1	25.1	26.5	27.4	25.6	26.7	1.9	1.9
Last MDS in Quarter 2	23.0	23.4	23.1	23.7	24.0	23.8	-0.4	-0.3
Last MDS in Quarter 3	23.0	25.1	23.6	23.3	25.0	24.0	-2.1	-1.7
Last MDS in Quarter 4	27.0	26.4	26.8	25.6	25.4	25.5	0.6	0.2

Table S2. Propensity Score Models of Hospice Election for 2004 and 2009 Nursing Home Decedents

Covariate	2004		2009		OR Difference*
	Odds Ratio	p-value	Odds Ratio	p-value	
Age at Death					
67-74	1.33	0	1.13	0	-0.20
75-78	1.25	0	1.12	0	-0.13
79-81	1.20	0	1.10	0	-0.11
82-83	1.20	0	1.10	0	-0.10
84-85	1.15	0	1.08	0	-0.07
86-87	1.15	0	1.08	0	-0.07
88-89	1.13	0	1.05	0.003	-0.08
90-91	1.09	0	1.02	0.23	-0.07
92-94	1.08	0	1.02	0.14	-0.06
95 and older	Ref		Ref		
Male	0.84	0	0.88	0	0.04
Nonwhite	1.04	0.005	0.96	0.002	-0.07
Married	1.05	0	1.00	0.97	-0.05
Major Diagnosis Group					
Cancer, no Dementia	2.36	0	1.94	0	-0.42
Cancer and Dementia	1.97	0	1.84	0	-0.13
Dementia, no Cancer	1.21	0	1.24	0	0.04
No Cancer and No Dementia	Ref		Ref		
Other Important Diagnoses					
Emphysema/COPD	1.02	0.007	1.01	0.24	-0.01
CHF	0.92	0	0.93	0	0.01
Diabetes	0.93	0	0.91	0	-0.02
ASHD	0.94	0	0.95	0	0.00
PVD	0.97	0.005	0.95	0	-0.02
CVA	0.94	0	0.95	0	0.00
Depression	1.11	0	1.12	0	0.01
Asthma	0.92	0.001	1.00	0.91	0.08
Resistant infection	0.82	0	0.84	0	0.02
Pneumonia	0.84	0	0.83	0	-0.01
Respiratory infection	0.92	0	0.93	0	0.01
Septicemia	0.92	0.005	0.87	0	-0.05
Cognitive Performance Scale					
CPS = 0 (intact)	Ref		Ref		
CPS = 1	1.05	0.01	1.09	0	0.05
CPS = 2	1.10	0	1.14	0	0.04
CPS = 3	1.10	0	1.23	0	0.14
CPS = 4	1.06	0.002	1.23	0	0.17
CPS = 5	1.14	0	1.40	0	0.26
CPS = 6 (totally dependent)	1.03	0.07	1.30	0	0.26
ADL score (0-28 range)					
0-12	Ref		Ref		
13-16	1.08	0	1.07	0	-0.01
17-18	1.04	0.04	1.08	0	0.04
19-20	1.13	0	1.15	0	0.02
21-21	1.18	0	1.22	0	0.04
22-23	1.18	0	1.25	0	0.07
24-25	1.29	0	1.35	0	0.05
26-27	1.39	0	1.44	0	0.05
28	1.53	0	1.53	0	0.00
ADL worsening since last MDS	2.09	0	2.24	0	0.15
DNH	1.15	0	1.05	0	-0.10
DNR	1.28	0	1.39	0	0.11
Short NH stay (<30 days)	0.63	0	0.64	0	0.01

Long NH stay (>90 days)	1.22	0	1.40	0	0.19
Days from last MDS to Death	1.01	0	1.01	0	0.00
No. of Hospitalizations in the year prior to the year of death	1.06	0	1.06	0	0.00
Average Hospitalizations LOS	1.00	0.31	1.00	0.12	0.00
NH FACILITY COVARIATES					
% Medicare NH residents	1.00	0	1.00	0	0.00
% Private Pay NH residents	1.01	0	1.02	0	0.00
NH has special care unit	1.16	0	1.07	0	-0.08
For-profit Status	1.20	0	1.17	0	-0.04
Hospital-Based	0.55	0	0.59	0	0.04
Multisite facility (chain)	1.09	0	1.13	0	0.04
Last MDS in Quarter 1	Ref		Ref		
Last MDS in Quarter 2	1.08	0	1.08	0	0.00
Last MDS in Quarter 3	1.12	0	1.09	0	-0.03
Last MDS in Quarter 4	1.00	0.67	1.01	0.21	0.01
	2004		2009		
C-statistic	0.70		0.71		

* Bold font used to highlight differences above 0.1 in absolute value.

Table S3. Comparison of NH Decedent and NH Facility Characteristics by Matching Group

Group	2004				2009			Standardized % Differences*		
	G1 ₂₀₀₄	G2 ₂₀₀₄	G3 ₂₀₀₄	UNM ₂₀₀₄ †	G1 ₂₀₀₉	G2 ₂₀₀₉	G3 ₂₀₀₉	G1 ₂₀₀₄ vs. G1 ₂₀₀₉	G2 ₂₀₀₉ vs. G2 ₂₀₀₄	G3 ₂₀₀₉ vs. G3 ₂₀₀₄
	Yes	No	No	No	Yes	Yes	No			
Hospice										
N (unique persons)	117,858	58,475	139,351	143,440	71,003	72,391	216,658			
N (weighted for matching)	117,858	72,391	216,658	143,440	117,858	72,391	216,658			
Age at Death	85.2	85.9	85.2	85.1	85.9	86.0	85.4	-9.0	1.2	2.1
Male (%)	31.5	30.8	36.0	36.1	31.8	32.0	37.4	-0.8	2.4	2.9
Non-White (%)	9.5	8.9	11.0	11.0	10.3	10.4	12.3	-2.6	4.6	4.1
Married (%)	24.5	23.1	25.4	25.8	24.4	24.2	26.6	0.3	2.7	2.7
Cancer, no Dementia (%)	18.6	13.5	13.4	13.4	15.3	14.2	13.7	9.1	2.1	1.0
Cancer and Dementia (%)	12.5	11.2	7.9	7.8	12.8	12.2	8.3	-0.9	3.2	1.5
Dementia, no Cancer	46.6	52.3	44.7	44.2	51.3	52.1	45.2	-9.5	-0.4	1.0
No Cancer and No Dementia	22.3	23.0	34.0	34.6	20.6	21.5	32.8	4.1	-3.5	-2.7
Emphysema/COPD	31.9	30.5	35.1	35.1	32.3	32.2	36.3	-0.7	3.6	2.5
CHF	46.2	46.7	52.6	52.5	44.1	44.9	50.9	4.1	-3.6	-3.5
Diabetes	25.8	25.3	28.4	28.6	29.0	29.4	32.7	-7.2	9.1	9.4
ASHD	14.7	15.3	16.1	16.0	15.4	15.6	16.9	-1.9	0.7	2.0
PVD	12.9	13.1	13.5	13.5	13.4	13.5	13.9	-1.5	1.1	1.2
CVA	21.8	22.5	22.3	22.4	19.4	19.6	19.5	5.9	-7.0	-7.0
Depression	42.8	42.9	36.2	36.3	48.6	48.1	40.5	-11.8	10.4	8.9
Asthma	2.4	2.5	2.5	2.5	3.1	3.1	3.1	-4.4	3.4	3.9
Resistant infection	3.1	3.1	4.4	4.3	3.5	3.3	4.6	-1.8	1.0	1.1
Pneumonia	9.6	9.6	13.2	13.5	9.3	9.3	13.2	1.3	-1.1	-0.2
Respiratory infection	3.7	3.9	4.3	4.3	3.2	3.2	3.7	2.8	3.5	-3.1
Septicemia	1.6	1.5	2.2	2.2	1.5	1.6	2.3	0.6	0.5	0.3
CPS Score (0-6 range)	3.50	3.66	3.29	3.27	3.46	3.46	3.09	1.9	-11.4	-10.8
ADL score (0-28 range)	21.7	21.9	21.1	21.1	21.9	21.7	21.0	-2.4	-2.9	-0.5
ADL worsening since last MDS	30.4	27.4	16.4	16.5	28.5	26.2	14.5	4.1	-2.9	-5.3
DNH	8.7	8.1	6.5	6.5	9.5	9.2	7.4	-2.8	4.1	3.8
DNR	75.8	77.0	66.9	66.6	77.4	77.4	66.4	-3.8	0.8	-1.1
Short NH stay (<30 days)	8.7	7.3	17.9	18.8	6.8	6.9	16.7	7.2	-1.1	-3.2
Long NH stay (>90 days)	76.4	80.4	64.0	63.3	79.9	79.8	64.3	-8.4	-1.3	0.7
Days from last MDS to Death	33.9	34.6	25.7	26.0	34.7	34.1	25.2	-2.9	-2.0	-1.7
Hospitalizations in the year prior to year of death (n)	0.75	0.72	0.68	0.68	0.77	0.76	0.72	-2.1	3.8	3.3

Average Hospitalizations LOS	2.28	2.20	2.15	2.1	2.17	2.17	2.10	2.6	-0.8	-1.3
NH FACILITY COVARIATES										
% Medicaid NH residents	59.4	60.4	61.7	61.1	57.0	57.5	59.5	2.9	12.9	8.2
% Medicare NH residents	13.8	13.2	15.8	16.3	15.7	15.8	17.4	-15.5	18.3	10.4
% Private Pay NH residents	26.8	26.4	22.5	22.6	27.3	26.7	23.1	-3.0	2.2	3.7
NH has special care unit	31.1	29.5	26.4	26.8	27.3	27.0	24.3	8.2	-5.6	-4.9
For-profit Status	66.5	64.8	64.6	64.5	67.1	67.1	67.2	-1.2	4.7	5.4
Hospital-Based Multisite facility (chain)	3.4	3.9	7.5	8.3	3.0	2.9	5.6	2.6	-4.6	-7.5
Last MDS in Quarter 1	55.3	54.6	53.4	53.7	55.8	55.9	53.7	-1.2	2.6	0.7
Last MDS in Quarter 2	25.1	25.2	27.1	27.3	25.4	25.7	27.4	-0.6	1.1	0.5
Last MDS in Quarter 3	23.4	23.5	22.9	22.9	24.2	23.8	23.7	-2.0	0.5	1.8
Last MDS in Quarter 4	25.1	24.3	22.9	22.8	25.2	24.9	23.3	-0.3	1.4	1.0
99th percentile of Difference in propensity scores between matched pairs								1.5 e-04	2.1 e-05	9.5 e-06
Unmatched Mean Differences								4.0	10.9	3.4
Matched Mean Differences								3.7	3.2	3.1
Rubin's B-statistic								25.2	25.6	25.4
Rubin's R-statistic								1.00	0.99	1.01

* Standardized % Differences are defined as $100 * (\text{Mean1} - \text{Mean0}) / \sqrt{(\text{Variance1} + \text{Variance0}) / 2}$, where 1 is the treatment subgroup and 0 is the (matched) non-treated subgroup. For example, for Group2, G2₂₀₀₉ is the treatment subgroup and G2₂₀₀₄ is the (matched) non-treated subgroup.

† UNM₂₀₀₄ = Non-Hospice decedents in 2004 left unmatched.

Table S4. Comparison of Nursing Home Decedents in 2004 and 2009, by Fee-For-Service (FFS) versus Medicare Advantage (HMO) Enrollment Status

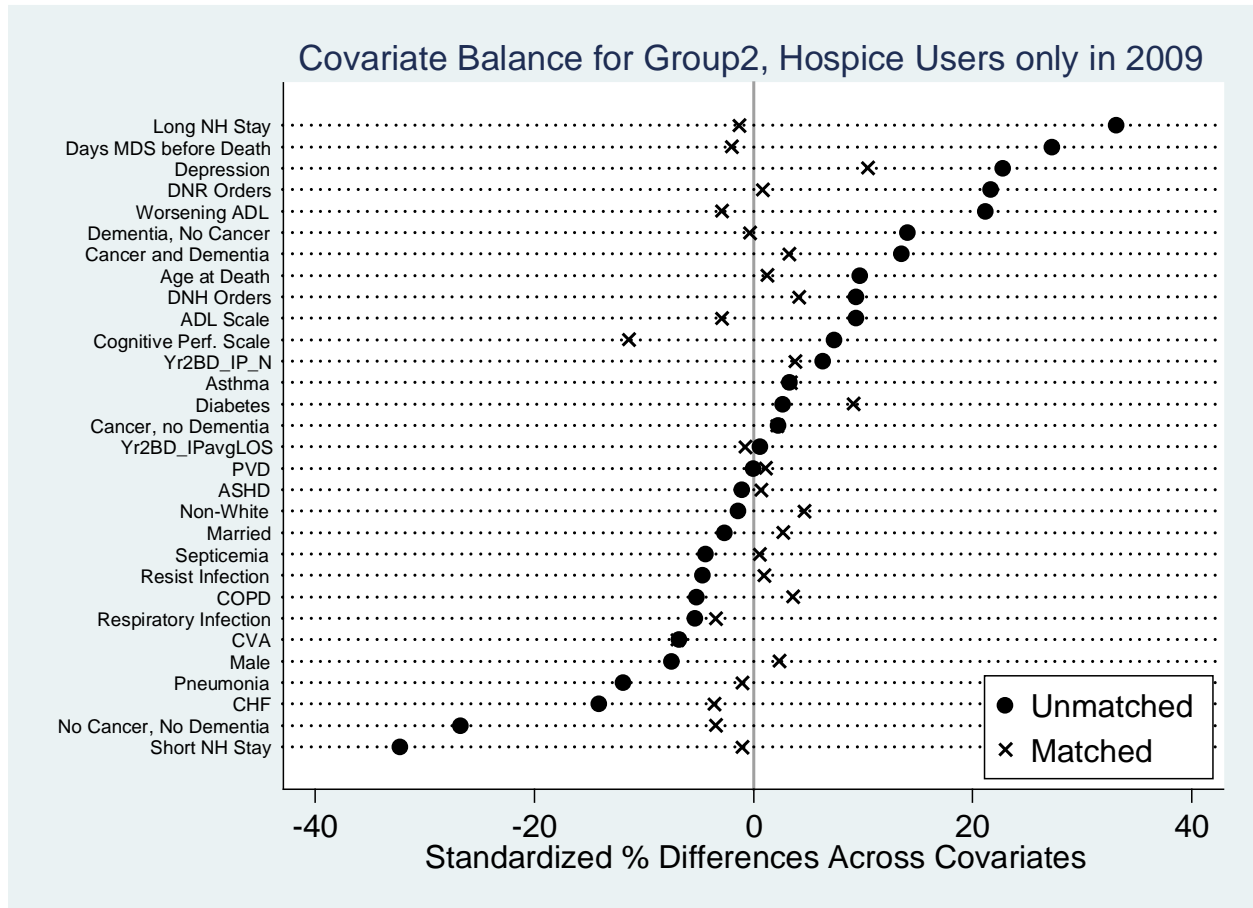
Covariate	2004				2009			
	FFS	HMO	Total	Total minus FFS	FFS	HMO	Total	Total minus FFS
N	429,883	51,027	480,910		364,847	78,483	443,330	
Age at Death	85.2	84.9	85.2	0.0	85.6	85.1	85.5	-0.1
Male	34.5	37.9	34.9	0.4	35.4	38.6	36.0	0.6
Nonwhite	10.5	12.8	10.7	0.3	11.5	15.2	12.2	0.6
Married	25.2	28.8	25.6	0.4	25.8	29.3	26.4	0.6
Major Diagnosis Group								
Cancer, no Dementia	14.9	12.7	14.7	-0.2	14.2	12.7	13.9	-0.3
Cancer and Dementia	9.3	7.4	9.1	-0.2	9.9	7.5	9.5	-0.4
Dementia, no Cancer	45.4	45.8	45.4	0.0	47.6	46.1	47.3	-0.3
No Cancer and No Dementia	30.4	34.1	30.8	0.4	28.3	33.7	29.2	1.0
Other Important Diagnoses								
Emphysema/COPD	33.9	25.4	33.0	-0.9	34.7	27.1	33.4	-1.4
CHF	50.3	36.7	48.9	-1.4	48.3	36.3	46.2	-2.1
Diabetes	27.5	26.6	27.4	-0.1	31.2	32.3	31.4	0.2
ASHD	15.6	12.8	15.3	-0.3	16.3	15.8	16.2	-0.1
PVD	13.3	11.6	13.1	-0.2	13.7	14.1	13.7	0.1
CVA	22.3	21.8	22.2	0.0	19.5	19.4	19.5	0.0
Depression	38.4	37.1	38.2	-0.1	43.4	42.4	43.2	-0.2
Asthma	2.5	2.3	2.5	0.0	3.1	3.1	3.1	0.0
Resistant infection	3.9	3.7	3.9	0.0	4.1	4.0	4.1	0.0
Pneumonia	12.1	12.3	12.1	0.0	11.7	11.0	11.5	-0.1
Respiratory infection	4.1	3.5	4.0	-0.1	3.5	3.1	3.4	-0.1
Septicemia	2.0	1.9	2.0	0.0	2.0	1.8	2.0	0.0
Cognitive Performance Scale (0-6 range; 0=intact, 6=very severe with eating problems)								
	3.4	3.3	3.4	0.0	3.2	3.2	3.2	0.0
ADL score (0-28 range; 0=independent, 28=totally dependent in all seven ADLs)								
	21.3	21.0	21.3	0.0	21.3	21.2	21.3	0.0
ADL worsening since last MDS								
	21.0	33.2	22.3	1.3	19.4	27.7	20.9	1.5
DNH	7.2	11.1	7.6	0.4	8.2	10.5	8.6	0.4
DNR	70.0	69.2	69.9	-0.1	70.7	71.0	70.8	0.0
Short NH stay (<30 days)	15.4	12.5	15.1	-0.3	13.4	11.7	13.1	-0.3
Long NH stay (>90 days)	68.0	68.1	68.0	0.0	69.9	70.0	69.9	0.0
Days from last MDS to Death	28.5	33.6	29.0	0.5	28.7	32.1	29.3	0.6
No. of Hospitalizations in the year prior to year of death								
	0.7	0.2	0.6	-0.1	0.7	0.3	0.7	-0.1
Average Hospitalizations LOS								
	2.2	0.7	2.0	-0.2	2.1	0.8	1.9	-0.2
NH FACILITY FACTORS								
% Medicaid NH residents	60.7	60.7	60.7	0.0	58.6	60.0	58.9	0.3
% Medicare NH residents	15.3	13.3	15.1	-0.2	16.8	14.3	16.4	-0.4
% Private Pay NH residents	24.0	26.1	24.2	0.2	24.6	25.6	24.8	0.2
NH has special care unit	28.1	31.3	28.4	0.3	25.4	27.5	25.8	0.4
For-profit Status	65.1	66.9	65.3	0.2	67.2	64.3	66.7	-0.5
Hospital-Based	6.4	4.2	6.2	-0.2	4.6	4.3	4.5	0.0
Multisite facility (chain)	54.1	54.5	54.2	0.0	54.8	53.9	54.6	-0.2
Last MDS in Quarter 1	26.6	26.4	26.5	0.0	26.7	26.1	26.6	-0.1
Last MDS in Quarter 2	23.1	23.0	23.1	0.0	23.8	23.9	23.8	0.0
Last MDS in Quarter 3	23.6	23.7	23.6	0.0	24.0	24.2	24.0	0.0
Last MDS in Quarter 4	26.8	26.9	26.8	0.0	25.5	25.8	25.6	0.1

Table S5. Association of Hospice and Total Medicare Expenditures in the Last Year of Life and End-of-Life Care Outcomes for Decedents in Nursing Homes with No Hospice in 2004 but using Hospice in 2009, All Diagnoses

Outcome	Hospice Group (Hospice only in 2009)		Non-Hospice Group (No Hospice Both Years)		Unadjusted Difference in Differences (95% CI)	Adjusted Difference in Differences* (95% CI)
	Year 2004	Year 2009	Year 2004	Year 2009		
	N=4,826	N=4,826	N=20,759	N=20,759		
Average Total Medicare Expenditures in the Last Year of Life	\$31,604	\$40,821	\$35,276	\$37,496	\$6,997 (\$5,342, \$8,652)	\$6,934 (\$5,444, \$8,424)

* Adjusted for person and NH characteristics. Confidence intervals adjusted for clustering of decedents within nursing home facilities. All expenditure prices in 2007 dollars.

Figure S1: Covariate Balance of Patient Characteristics for Group 2 Consisting of NH Decedents that Elected Hospice in 2009 and their 1:1 Matched Non-Hospice 2004 NH Decedents



REFERENCES

1. Gozalo PL, Miller SC, Intrator O, Barber JP, Mor V. Hospice effect on government expenditures among nursing home residents. *Health Services Research* 2008;43:134-53.
2. Rosenbaum PR, Rubin DB. Constructing a Control-Group Using Multivariate Matched Sampling Methods That Incorporate the Propensity Score. *Am Stat* 1985;39:33-8.
3. Austin PC. Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. *Stat Med* 2009;28:3083-107.
4. Normand ST, Landrum MB, Guadagnoli E, et al. Validating recommendations for coronary angiography following acute myocardial infarction in the elderly: a matched analysis using propensity scores. *J Clin Epidemiol* 2001;54:387-98.
5. Rubin DB. Using Propensity Scores to Help Design Observational Studies: Application to the Tobacco Litigation. *Health Services and Outcomes Research Methodology* 2001;2:169-88.