Final

Evaluation of the Medicare Prior Authorization Model for Repetitive Scheduled Non-Emergent Ambulance Transport (RSNAT): Second Interim Evaluation Report Appendices

HHSM-500-2014-00034I/HHSM-500-T0004

September 2020

Andrew Asher, Kara Contreary, and Jared Coopersmith

Submitted to:

Steve Blackwell
U.S. Department of Health and Human Services
Centers for Medicare & Medicaid Services
7500 Security Blvd.
Baltimore, MD 21244

Submitted by:

Mathematica
1100 1st Street, NE, 12th Floor
Washington, DC 20002-4221
(202) 484-9220
(202) 863-1763

This project was funded by the Centers for Medicare & Medicaid Services under contract no. HHSM-500-2014-00034I, Task Order HHSM-500-T0004. The statements contained in this report are solely those of the authors and do not necessarily reflect the views or policies of the Centers for Medicare & Medicaid Services. Mathematica assumes responsibility for the accuracy and completeness of the information contained in this report.

This page has been left blank for double-sided copying.

CONTENTS

This page has been left blank for double-sided copying.

TABLES

FIGURES

This page has been left blank for double-sided copying

Appendix A

Comparison group selection

This page has been left blank for double-sided copying.

The evaluation uses a difference-in-difference design that estimates the effect of the prior authorization model as the difference between the change in outcomes in the model states versus the change in outcomes for a set of comparison states. An important part of the design is the choice of the latter (generally referred to as the counterfactual), which permits the comparison with the model states. Because a well-chosen counterfactual reduces the need for the analysis to depend as critically as it otherwise would on the multivariate analyses' modeling specifications, it is important to choose a set of comparison states as similar as possible to the model states along a range of characteristics, including baseline values of the outcomes, assumed to be related to the model and its measured outcomes at follow-up. Such an approach minimizes the risk that confounding factors will produce misleading results. The purpose of this appendix is to describe the challenge we faced in selecting a counterfactual, our approach to selecting a set of matched-comparison states, and the results of our approach.

A. Choosing comparison states

To maximize the internal validity of the difference-in-difference analytic approach, we chose a set of states as similar as possible to the model states in the period before implementation of prior authorization. A simple comparison of the model states to all other non-model states could be misleading, particularly as the Centers for Medicare and Medicaid Services (CMS) chose the states with the highest service utilization of repetitive, scheduled, non-emergent ambulance transports (RSNAT) for the initial implementation of prior authorization. For example, in the year immediately before the prior authorization implementation, average RSNAT utilization and the proportion of beneficiaries frequently using RSNAT services were roughly six times higher in the Year 1 model states than in the rest of the United States (Table A.1). In contrast, the increase in RSNAT utilization between 2012 and 2014 was 47 times larger in the remaining states than in the Year 1 model states, and beneficiaries in the remaining states were about twice as likely to live in rural areas as in the model states. Thus, based on numerous characteristics, the model states differ substantially from the rest of the United States as a whole.

Table A.1. Comparing Year 1 model states with all other states in 2014

Mean characteristic	Year 1 model states (New Jersey, Pennsylvania, South Carolina)	Rest of United States[a]
RSNAT utilization (RSNAT trips per 100,000 Medicare beneficiaries)	3446.80	565.20
Change in RSNAT utilization between 2012 and 2014	+1.90	+90.10
Proportion frequently using RSNAT services (proportion of all beneficiaries with more than 40 RSNAT trips)	0.19	0.03
Percentage of beneficiaries living in rural areas	16.50	36.70

[a]Excludes Alaska and Hawaii.

To address the challenge associated with differences in the states, we used a statistical technique that is designed to select a group of states as similar as possible to the model states on a range of characteristics (described below). One complicating factor in this approach was the expansion of the prior authorization model to six additional states in Year 2 of the model (2016). To avoid the complication of matching some expansion states to the Year 1 states, which would then require re-matching in Year 2, we adopted a strategy whereby we matched Year 1 and Year 2 model states simultaneously with comparison states, making no distinction between the Year 1 and Year 2 model states when matching. This approach optimized balance for the analysis of the combined Year 1 and Year 2 model states and avoided the need to re-match.

The above approach of matching at the state level essentially limits the subsequent comparison analysis at the supplier and beneficiary levels to those suppliers located or those beneficiaries residing in the matched comparison states. An alternative approach would largely ignore any state-level matching and instead simply match suppliers and beneficiaries in the model states to suppliers and beneficiaries in any other state regardless of location. We rejected such an approach for two reasons. First, it is commonly accepted that matched-comparison selection (at least primary unit selection) should be performed at the same level at which the actual selection was made. In this case, individual ambulance suppliers and Medicare beneficiaries did not select themselves into the prior authorization model; rather, the entire state in which they are located or live was selected into the model. Therefore, a similar selection process is warranted when selecting comparison units. Second, state-level matching may control for some unobserved confounding factors if these factors are correlated with the observed characteristics used in the matching.

B. Matching approach

To select a set of comparison states as similar as possible to the model states, we first needed to identify a set of variables measuring the characteristics we expect to be related to the intervention and the outcomes. In Table A.2, we describe the set of variables we identified as potential matching variables. Unless otherwise noted in the table, all measures were observed in each of the years 2012, 2013, and 2014 for all states.

Table A.2. Measures identified for potential use in matching

Name	Definition
RSNAT service utilization	Number of RSNAT[a] service trips per 100,000 Medicare beneficiaries
Change in RSNAT service utilization since 2012	Percentage change in RSNAT service utilization using 2012 as base measure
Availability of ambulance suppliers	Number of unique ambulance suppliers with RSNAT services per 100,000 beneficiaries
Percentage using RSNAT services	Percentage of beneficiaries with at least one RSNAT trip during the year
Proportion frequently using RSNAT services	Proportion of beneficiaries with at least 40 RSNAT trips during the year
Proportion with ESRD	Proportion of beneficiaries with ESRD
Medicare improper payment rate	Bayesian shrinkage estimates of improper payment rates using CERT data (pooled across years)
Mean age	Average age of beneficiaries

Name	Definition
Percentage of Medicare beneficiaries living in nursing homes	Percentage of beneficiaries living in nursing homes in 2012 (http://www.dartmouthatlas.org/data/table.aspx?ind=337)
Percentage rural	Percentage of beneficiaries living in rural areas, defined as beneficiary zip codes outside MSAs

[a]RSNAT service trips were defined by identifying claim lines with a Healthcare Common Procedure Coding System (HCPCS) code value of A0426 or A0428 occurring at least six times in a single 10-day period or at least twice per week for at least three weeks.

CERT = Comprehensive Error Rate Testing; ESRD = end-stage renal disease; MSA = metropolitan statistical area.

To select the matched-comparison states, we used the statistical technique called optimal matching as implemented in the R package Optmatch (Hansen and Klopfer 2006). When forming matches, the technique examines the balance on covariates between individual matches and across the full matched sample, forming, rejecting, and reforming matches until achieving a minimum distance within and across matched sets. Hence, the algorithm minimizes both local and global imbalance, making it much more flexible than commonly used greedy[1] matching techniques.

Given the high RSNAT utilization in the model states and the expectation that prior authorization may differentially affect beneficiaries in rural versus urban locales, we prioritized these two measures when selecting the set of comparison states. We tested several matching specifications by using (1) the Mahalanobis distance based on various combinations of measures and (2) calipers[2] on several measures that disallowed matches if the distance on a specific measure exceeded a certain threshold. Even though the estimation of a propensity score is a common method for collapsing multidimensional data into a single distance measure, the small number of observations, particularly among the model states, could result in unstable parameter estimates, making the estimated propensity scores suspect. Our goal was to select up to two comparison states per model state, balancing on as many of the characteristics from Table A.2 as possible but prioritizing balance on RSNAT utilization and the percentage of beneficiaries living in rural areas. We also excluded Alaska, Hawaii, North Dakota, Rhode Island, South Dakota, Vermont, and Wyoming from the comparison group, because these states had unique geography and features (for example, very small populations) that could call into question the comparability of these states to the model states.

Our final matching specification used the Mahalanobis distance for RSNAT utilization, the change in RSNAT utilization since 2012, the availability of ambulance suppliers, and the percentage of beneficiaries living in rural areas as well as calipers on RSNAT utilization and the

[1] A greedy matching algorithm is frequently used to match cases to controls in observational studies. In a greedy algorithm, a set of treated units is matched to a set of control units. The comparison unit with the lowest distance is selected in sequential order for each treated unit, and once a match is made for a given treated unit, it is not reconsidered. As a result, the pool of potential comparisons for treated units matched last will be smaller than for those matched first, potentially leading to low quality matches.

[2] Calipers represent the maximum tolerated difference between matched beneficiaries on one or more measures.

growth in RSNAT utilization since 2012. In the matching, we used the 2013 observations of each of these measures to avoid the possibility of any anticipatory effects among the model states in 2014. The measures in 2013 were highly correlated with those in 2012.[3] After finalizing the matched sets, we defined state-level matching weights that equally weight the model states and the comparison states within each matched set. In Table A.3, we list the model states and the selected matched-comparison states; in Table A.4, we present the balance on the key measures before and after matching.[4]

Table A.3. Model and matched-comparison states

Year 2 model states	Year 2 matched-comparison states
Delaware Maryland North Carolina Virginia West Virginia Washington, DC	Alabama Florida Kentucky Louisiana Massachusetts Montana Nebraska Ohio Texas Washington
Year 1 model states	**Year 1 matched-comparison states**
New Jersey Pennsylvania South Carolina	Georgia Indiana Tennessee

[3] No changes were made to the state-level matching approach for this report. State-level matches are identical to those used in previous interim reports.

[4] Beneficiaries from all comparison states are included in each analysis. The year considered post-implementation is determined by the state-level match to either a Year 1 or Year 2 model state.

Table A.4. Pre- and post-matching balance on key measures, 2012–2014

Measure	Model states	Comparison, pre-matching, all states other than the model states and those listed below[a]		Comparison post-matching (the final matched-state list provided above)[b]	
	Mean	Mean	Standardized difference	Mean	Standardized difference
All model states					
RSNAT utilization (trips per 100,000 Medicare beneficiaries)[c]	1,716.00	525.00	1.33	1,356.50	0.40
Change in RSNAT utilization since 2012 (trips per 100,000 Medicare beneficiaries)[c]	150.80	63.30	0.21	136.70	0.03
Availability of ambulance suppliers (unique ambulance suppliers with RSNAT services per 100,000 beneficiaries)[c]	9.40	4.40	1.07	7.60	0.37
Proportion of beneficiaries using RSNAT services	0.20	0.08	1.38	0.17	0.37
Proportion frequently using RSNAT services	0.09	0.03	1.35	0.08	0.40
Percentage with ESRD	1.50	1.30	0.65	1.50	0.09
Mean beneficiary age	71.00	70.90	0.12	70.50	0.49
Percentage living in nursing homes	2.50	2.60	-0.08	2.80	-0.27
Percentage rural[d]	24.70	35.80	-0.58	32.10	-0.38
Year 1 model states					
RSNAT utilization (trips per 100,000 Medicare beneficiaries)[c]	3,385.05	525.00	2.34	2,565.06	0.67
Change in RSNAT utilization since 2012 (trips per 100,000 Medicare beneficiaries)[c]	2.33	63.30	-4.37	8.84	-0.47
Availability of ambulance suppliers (unique ambulance suppliers with RSNAT services per 100,000 beneficiaries)[c]	15.0	4.40	3.08	9.25	1.67
Proportion of beneficiaries using RSNAT services	0.31	0.08	2.07	0.27	0.36
Proportion frequently using RSNAT services	0.18	0.03	2.16	0.15	0.43
Percentage with ESRD	1.38	1.30	0.50	1.78	-2.51
Mean beneficiary age	71.29	70.90	0.34	69.96	1.15
Percentage living in nursing homes	2.67	2.60	0.22	2.63	0.12
Percentage rural[d]	16.73	35.80	-1.22	37.65	-1.33

Measure	Model states	Comparison, pre-matching, all states other than the model states and those listed below[a]		Comparison post-matching (the final matched-state list provided above)[b]	
	Mean	Mean	Standardized difference	Mean	Standardized difference
Year 2 model states					
RSNAT utilization (trips per 100,000 Medicare beneficiaries)[c]	881.51	525.00	0.59	752.14	0.21
Change in RSNAT utilization since 2012 (trips per 100,000 Medicare beneficiaries)[c]	225.04	63.30	0.29	200.56	0.04
Availability of ambulance suppliers (unique ambulance suppliers with RSNAT services per 100,000 beneficiaries)[c]	6.58	4.40	0.34	6.83	-0.04
Proportion of beneficiaries using RSNAT services	0.15	0.08	0.96	0.12	0.41
Proportion frequently using RSNAT services	0.06	0.03	0.83	0.04	0.55
Percentage with ESRD	1.56	1.30	0.43	1.31	0.41
Mean beneficiary age	70.80	70.90	-0.12	70.77	0.04
Percent living in nursing homes	2.47	2.60	-0.54	2.88	-1.69
Percentage rural[d]	28.64	35.80	-0.34	29.30	-0.03

[a]Excludes Alaska, Hawaii, North Dakota, Rhode Island, South Dakota, Vermont, and Wyoming.

[b]Weighted by using the state-level matching weights.

[c]Included in the matching specifications.

[d]Medicare improper payment rates are omitted from Table A.4. The post-matching difference was less than one percentage point.

ESRD = end-stage renal disease.

When all model states are considered, balance was increased on all of the key measures included in the matching as well as on several other measures. The initial differences in RSNAT utilization, supplier availability, and percentage of rural beneficiaries all decreased as a result of the matching. In addition, the differences in the proportion using RSNAT and the proportion frequently using RSNAT both decreased, even though we did not match on these measures, providing more evidence that the selected comparison states had higher-than-average baseline RSNAT utilization. High RSNAT utilization was a key factor in the assignment of states to the prior authorization model, so selecting comparison states with higher-than-average RSNAT utilization increases the validity of the selected comparison states as a counterfactual to the model states. Even though differences in mean age and the proportion of a state's residents living in nursing homes increased as a result of the matching, the differences were minor both before and after matching and are therefore not concerning. The weights generated as part of the matching process were incorporated into the analysis weights for beneficiaries and for suppliers. Appendices C and D address design effects due to weighting.

Matching generally improved balance for the Year 1 and Year 2 model states separately, though not on all measures. However, differences on these characteristics were still too small to prompt concern. For example, in the Year 1 model states, the difference in the mean beneficiary age increased from less than one half of a year to slightly over one year. Despite the increased difference, the difference is still only one year on average. So although the differences increased for a limited number of characteristics, they remained small.

This page has been left blank for double-sided copying.

Appendix B

Quantitative methods

This page has been left blank for double-sided copying.

A. Study sample

1. Beneficiaries

For the beneficiary analysis, we restricted our analysis group to beneficiaries who were in fee-for-service (FFS) for at least part of a given quarter, and were living in one of the included states (Year 1, Year 2, or comparison states) and were identified as having end-stage renal disease (ESRD) and/or stage 3 or 4 pressure ulcers (hereafter referred to as "pressure ulcers").[5] In any given quarter (for example, April–June 2015), the eligible sample included individuals designated as having ESRD and/or pressure ulcers based on their claims from the current calendar year (in this case, 2015). Beneficiaries with these conditions were believed to be most likely to be affected by the prior authorization model in both their ambulance use and possibility of any adverse events. Repeated non-emergency ambulance transportation is relatively uncommon among Medicare beneficiaries; however, over 85 percent of beneficiaries who used repetitive, scheduled, non-emergent ambulance transport (RSNAT) claims in our model and comparison states from January 2012 through December 2018 had ESRD and/or pressure ulcers. Thus, we could capture a substantial portion of the effect of the model while improving our likelihood of detecting an impact. In the next subsection, we describe our method of selecting the included chronic conditions.

We selected beneficiaries with ESRD and/or pressure ulcers as our population of interest through a multipart process. First, we identified all ambulance trips that met the definition of RSNAT for purposes of the model (coded as A0426 or A0428 and occurring with the requisite frequency[6]). Then, for individuals identified as having taken RSNAT ambulance trips, we examined all carrier and outpatient claims that occurred on the same day as an RSNAT trip. We reasoned that services received on the same day as an ambulance trip were likely the services that necessitated the trip. Using the Agency for Healthcare Research and Quality's Clinical Classifications Software, we then grouped the primary diagnosis codes and all the procedure codes from the claims, identifying the most common diagnosis and procedure categories to select a group of individuals who could be considered likely users of RSNAT services based on their health conditions. We consulted with a medical expert to ensure that our selection was reasonable and that we had indeed identified a group of individuals who were at elevated risk of using regular, nonemergency ambulance transportation.

At first, we identified three chronic condition groups as at high risk of RSNAT use: beneficiaries with ESRD, cancer, and skin ulcers (including pressure and non-pressure ulcers). These three groups together accounted for 97 percent of all RSNAT trips in our model and comparison states. We ultimately excluded beneficiaries with cancer only and beneficiaries with non-pressure ulcers because they had lower RSNAT utilization rates and lower likelihood of meeting RSNAT

[5] Pressure ulcers, also called decubitus ulcers or bedsores, are localized damage to the skin (and possibly the underlying tissue) that usually occur over a bony prominence as a result of pressure or a combination of pressure and friction. We included the most severe forms—pressure ulcer of skin with full thickness skin loss (stage 3) and pressure ulcer of the skin with necrosis through to muscle, tendon, or bone (stage 4).

[6] Three or more round trips in a 10-day period or at least one per week for three weeks or more.

medical necessity criteria, yielding a population that comprised 85 percent of RSNAT users in our states of interest. Although these exclusions yielded a study population with relatively high RSNAT utilization rates, if the model affects RSNAT users with cancer and/or non-pressure skin ulcers differently from beneficiaries with ESRD and/or pressure ulcers, our results could be slightly biased for the full population of RSNAT users.

We used the Hierarchical Condition Category (HCC) software to identify individuals with ESRD (HCC134 and HCC136) and/or stage 3 or 4 pressure ulcers (HCC157 and HCC158). The software analyzed a full year of claims for an individual, applying an algorithm that identified diagnosis or procedure codes associated with the selected condition (Centers for Medicare & Medicaid Services [CMS] 2017). To identify additional beneficiaries with ESRD, we used two variables from the Medicare denominator file: "original reason for entitlement" and "current reason for entitlement." If either variable indicated ESRD, we classified that beneficiary as having ESRD for that year.

The HCC software also produced a set of overall HCC scores for each beneficiary; it estimated the degree to which expected Medicare expenditures for a beneficiary would differ from the average in the next year. We used the HCC scores as proxy for overall health status—a measure of the beneficiary's relative risk of needing hospitalizations and other expensive Medicare services.

Our study sample is a repeated cross-section of the Medicare population with ESRD and/or pressure ulcers. Beneficiaries are in the sample in quarters during which they are FFS Medicare beneficiaries and are designated as having ESRD and/or pressure ulcers based on their claim history in the current year. For example, a beneficiary who had ESRD according to the 2013 claim history is included in the sample in all 2013 quarters in which he or she was enrolled in FFS Medicare. Applying these restrictions, our study group consisted of a total of 540,392 beneficiaries who resided exclusively in model states and 1,012,173 beneficiaries who resided exclusively in comparison states. We excluded beneficiaries who moved between model and comparison states during the study period, which resulted in dropping 2.9 percent of beneficiaries from the sample. The length of time that each beneficiary was part of our sample ranged from 1 to 28 quarters, with a mean duration of 7.6 quarters for model-only beneficiaries and 7.7 quarters for comparison-only beneficiaries, for a total of 11,883,591 beneficiary-quarters. Only 25,169 model-only beneficiaries (4.7 percent) and 45,557 comparison-only beneficiaries (4.5 percent) were included in all 28 quarters.

2. Suppliers

We identified suppliers from carrier claims based on National Provider Identifier (NPI) and provider state codes.[7] Our study population consists of all nonhospital-based ambulance

[7] If the provider state code indicated any of the model or comparison states, we matched the corresponding NPI with the National Plan and Provider Enumeration System (NPPES) file to verify the location of the supplier. We excluded three suppliers whose NPI numbers were invalid or who, when matched to the NPPES file, we determined were not garaged in a model or comparison state.

suppliers garaged in any of the model or comparison states that billed Medicare for ambulance services in any quarter of our study period. The population included 3,093 model state suppliers and 5,146 comparison state suppliers.

B. Outcome measures

1. Beneficiaries

We examined several quarterly outcomes related to Objective 1 (utilization and expenditures), Objective 2 (quality of care and access to care), and Objective 4 (denied claims) (Table B.1). We generated utilization outcomes by identifying claim lines that met the specified criteria and then aggregating each beneficiary's claim lines up to the quarter level. Outcomes related to utilization and expenditures fall into the following categories:

- RSNAT service utilization: Claim lines for nonemergency ambulance transportation using codes A0426 or A0428 and occurring as part of a sequence of trips that meets the frequency requirements specified in the model. RSNAT expenditures also included mileage claims using code A0425.
- Any Medicare-covered utilization of a ground ambulance: Claim lines for any ground ambulance transportation, with RSNAT claim lines (above) as a subset of these claim lines. Ambulance expenditures also included mileage claims.
- Total Medicare health care expenditures: The sum of claim line payment amounts across all claim types, including carrier, outpatient, inpatient, skilled nursing facility, home health, hospice, and durable medical equipment.

To assess whether there was evidence that the model resulted in unintended consequences, we examined impacts on quality of care and access to care (Objective 2). Quality of care measures included:

- Emergency department visits: Claim lines with revenue center codes 0450-0459 and 0981. We excluded claim lines with those revenue center codes that were for laboratory or imaging services.
- Emergency ambulance trips: Claim lines for ambulance transportation using codes A0427, A0429, A0433, and A0434.
- Unplanned inpatient admissions: Defined according to the specifications set forth by the Yale New Haven Health Services Corporation and the Center for Outcomes Research & Evaluation in their 2016 All-Cause Hospital-Wide Measure Updates and Specifications Report (CMS 2016a). This includes hospital admissions for acute conditions or for procedures that are not typically scheduled in advance. Admissions for planned procedures with no accompanying acute diagnosis are not included in the measure.
- Death: Identified using the date of death and the valid date of death variables in the Master Beneficiary Summary File.

To more closely study access to care among beneficiaries with ESRD, we examined the utilization of dialysis, as well as outcomes related to failure to receive timely dialysis—emergency dialysis and hospitalization for complications of ESRD. Access to care measures included:

- Scheduled dialysis: Claim lines with the codes listed in Table B.1.
- Emergency dialysis: Claims lines with the code listed in Table B.1.
- Hospitalization for complications of ESRD: Hospital admissions where the primary diagnosis is represented by one of the codes in Table B.2. We developed the list in collaboration with clinician researchers who have experience treating patients with ESRD.

Table B.1. Procedure codes included in dialysis measures

HCPCS code	Short description
Scheduled dialysis	
90935	Hemodialysis one evaluation
90937	Hemodialysis repeated evaluation
90999	Dialysis procedure
90945	Dialysis one evaluation
90947	Dialysis repeated evaluation
90997	Hemoperfusion
Emergency dialysis	
G0257	Unscheduled dialysis ESRD pt hos

Table B.2. Diagnosis codes indicating conditions related to exacerbation of untreated ESRD

ICD-9 code	ICD-10 code	Short description
275.2	E83.40, E83.41, E83.42, E83.49	Disorders of magnesium metabolism
275.3	E83.30, E83.31, E83.32, E83.39	Disorders of phosphorus metabolism
275.4		Disorders of calcium metabolism
275.40	E83.50	Unspecified disorder of calcium metabolism
275.42	E83.52	Hypercalcemia
276.1	E87.1	Hyposmolality and/or hyponatremia
276.2	E87.2	Acidosis
276.6	E87.70, E87.79	Fluid overload disorder
276.7	E87.5	Hyperpotassemia
276.9	E87.8	Electrolyte and fluid disorders not elsewhere classified
428.0	I50.9	Congestive heart failure, unspecified
428.1	I50.1	Left heart failure

ICD-9 code	ICD-10 code	Short description
428.2		Systolic heart failure
428.20	I50.20	Systolic heart failure, unspecified
428.23	I50.23	Acute on chronic systolic heart failure
428.9	I50.9	Heart failure, unspecified
586	N19	Renal failure, unspecified
782.3	R60.0, R60.1, R60.9	Edema
786.05	R06.02	Shortness of breath
780.97	R41.82	Altered mental status

To measure denied claims (Objective 3), we created the following measure:

- Denied non-emergency ambulance claims: Claims lines with codes A0426 or A0428 that were denied.

For most utilization measures, we considered both the likelihood of receiving any services in the quarter and the number of services received. The likelihood of receiving any services was represented as a binary variable equal to one if the beneficiary received at least one service in the category during the quarter. This approach allowed us to explore the degree to which the model influences the number of individuals who receive services, the average number of services received by individuals, or both. For ambulance trips and emergency department visits, the total number of trips or visits was of interest. For dialysis, we counted the number of days in the quarter in which an individual received the service. Given that beneficiaries required dialysis on a regularly scheduled basis, we also measured the average number of days between dialysis services. The recommended delivery schedule for dialysis usually does not vary for a given patient, thereby suggesting that an increase in the average number of days between treatments could possibly indicate a delay in receiving needed care. For denied claims, only the number of such claims was relevant.

The set of outcome measures is given in Table B.3.

Table B.3. Beneficiary-quarterly outcome measures

Research question	Quarterly measures
How does prior authorization affect total ambulance service use?	• Probability of RSNAT ambulance service utilization • Number of RSNAT ambulance trips • Probability of any Medicare ambulance utilization • Total number of Medicare ambulance trips
How does prior authorization affect Medicare expenditures for beneficiaries?	• RSNAT service expenditures • All Medicare ambulance expenditures • Total Medicare FFS expenditures

Research question	Quarterly measures
How does prior authorization affect the volume of services expected to be affected by access to RSNAT services? How does it impact quality and adverse outcomes?	• Probability of emergency department utilization • Number of emergency department visits • Probability of emergency ambulance utilization • Number of emergency ambulance trips • Probability of unplanned inpatient admission • Probability of death
How does prior authorization affect beneficiaries' use of dialysis services (beneficiaries with ESRD only)? Did prior authorization impact beneficiaries' frequency of ESRD complications (beneficiaries with ESRD only)?	• Probability of dialysis use • Number of days of dialysis use • Average number of days between dialysis services • Probability of emergency dialysis • Number of emergency dialysis treatments • Probability of hospitalizations for ESRD complications • Number of hospitalizations for ESRD complications
Does prior authorization impact claims denial rates?	• Number of denied non-emergency ambulance claims

2. Suppliers

The supplier analysis addresses research questions concerning the impact of prior authorization on supplier exit from the Medicare ambulance market. Our only outcome measure was whether or not the supplier billed Medicare for ambulance services in a given year. For example, we assumed that a supplier that billed Medicare in year *t* but did not bill in year *t+1* had exited the market.

C. Analysis

1. Beneficiary

Beneficiaries in the model and comparison states exhibited some small differences in their demographic and health characteristics that might be associated with our outcomes of interest. To improve the comparability of the two groups on demographic and health characteristics, we generated propensity score weights for each beneficiary based on his or her age, sex, race, and whether the beneficiary lived in a rural area. To generate the weights, we used logistic regression analysis, predicting whether a beneficiary lived in a model or comparison state by using the set of characteristics above. This regression produced estimated propensity scores, which we used to calculate weights to balance the model and comparison beneficiaries. In Appendix C, we provide more information on the beneficiary weights.

We generated weighted summary statistics of demographics and health for the model and comparison groups as well as their baseline levels of the outcome measures. We then used generalized difference-in-differences models to estimate the impact of prior authorization on each outcome. For binary variables, we used logistic regression; for count variables, we used negative binomial regression; for continuous variables, we used ordinary least squares. We weighted observations and adjusted standard errors to account for the effects of weighting and the non-independence of observations on the same individual in several quarters. We estimated the following regression equation:

(Equation 1) $$E\left[Y_{ist}\right] = F\left(\alpha + \sum_{s \in S_{-1}} \rho_s IS + \sum_{t \in T_{-1}} \gamma_t IT + \beta * Post_{st} + \delta X_{ist}\right),$$

where Y_{ist} is the outcome for beneficiary *i* in state *s* in quarter *t*. IS and IT are state and quarter fixed effects, respectively (omitting one indicator from each group). $Post_{st}$ takes value 1 in states and quarters when the model was in effect, and 0 otherwise. X_{ist} is a set of beneficiary-quarter–level control variables. Controls include age; age squared; the HCC score based on claims in the concurrent calendar year; length of time the beneficiary's county has been subject to a moratorium on new Medicare suppliers;[8] and indicators for race (white, black, or other), sex, rural residence, dual eligibility for Medicare and Medicaid, having a claim for a hospital bed for home use before or in quarter t, and residing in a county with a moratorium on new Medicare suppliers. The coefficient of interest is β, which gives the estimated per beneficiary per quarter impact of residing in a model state after prior authorization was implemented. *F(x)* is the cumulative logistic distribution for binary outcomes and the identity function for continuous outcomes. For ease of interpretation, we converted logistic regression coefficients into average marginal effects.

We estimated each regression on the full set of beneficiaries with ESRD and/or pressure ulcers, and then stratified by chronic condition, rural residence, and dual eligibility for Medicare and Medicaid, as these characteristics are likely to affect need for, access to, and utilization of ambulance services, and the impact of the control variables may differ among these groups as well. We also analyzed Year 1 and Year 2 states separately, which allowed us to see if results differed between the initially targeted states and the expansion states.[9] Finally, we estimated the regressions on the subsample of beneficiaries who had a claim for a hospital bed for home use, which we used as a proxy for mobility issues (Table B.4).

[8] As of July 30, 2013, the Texas counties of Brazoria, Chambers, Fort Bend, Galveston, Harris, Liberty, Montgomery, and Waller were subject to a moratorium on enrollment of ambulance suppliers. As of January 30, 2014, several counties around Philadelphia, Pennsylvania, were added to the moratorium. The counties include Bucks, Delaware, Montgomery, and Philadelphia counties in Pennsylvania, and Burlington, Camden, and Gloucester counties in New Jersey. The Texas moratorium ended September 1, 2017.

[9] Our comparison group strategy optimized for balance between the full model and comparison groups. Our balance was therefore slightly worse when removing one of the cohorts to run these separate Year 1 and Year 2 analyses. However, for both groups, balance was still very good, and well within bounds generally considered suitable for this type of analysis (less than 0.05 standard deviations).

Table B.4. Model variants used in quantitative analysis of beneficiary outcomes

Model variant	Subgroups	Sample size (beneficiary-quarters)	Included beneficiaries
Full sample	None	11,566,321	All FFS beneficiaries with ESRD and/or pressure ulcers
Chronic condition stratification	ESRD only	8,109,475	All FFS beneficiaries with ESRD only
	Pressure ulcers only	3,118,359	All FFS beneficiaries with pressure ulcers only
	ESRD and pressure ulcers	338,487	All FFS beneficiaries with both ESRD and pressure ulcers
Cohort stratification	Year 1 cohort	9,480,975	All FFS beneficiaries with ESRD and/or pressure ulcers in the Year 1 states and comparison states[a]
	Year 2 cohort	9,628,136	All FFS beneficiaries with ESRD and/or pressure ulcers in the Year 2 states and comparison states[a]
Rural stratification	Rural	2,479,711	All FFS beneficiaries with ESRD and/or pressure ulcers residing in a zip code not included in a metropolitan statistical area
	Not rural	9,086,610	All FFS beneficiaries with ESRD and/or pressure ulcers residing in a zip code included in a metropolitan statistical area
Dually eligible stratification	Dually eligible	4,669,287	All FFS beneficiaries with ESRD and/or pressure ulcers dually eligible for Medicare and Medicaid
	Not dually eligible	6,897,034	All FFS beneficiaries with ESRD and/or pressure ulcers eligible for Medicare only
Hospital bed subsample	Hospital bed claim	928,464	All FFS beneficiaries with ESRD and/or pressure ulcers who had a claim for a hospital bed

[a] For the analyses stratified by cohort, the full set of comparison states were included as the comparison group for both the Year 1 and Year 2 model states.

We analyzed some alternative regression model specifications. For continuous and count outcomes (such as payments or number of trips), we conducted analyses on just the subset of beneficiaries with non-zero outcomes to assess the effects on these outcomes among users.

For the denied claims outcomes, we estimated a variant of the regression model that included separate indicator variables for each quarter after model implementation in the model states. That version of the regression analysis enabled us to assess whether the impact of the model differed over time. Specifically, we were interested in determining whether claim denial increases immediately following implementation but reverts to baseline levels as patients and suppliers acclimate to prior authorization.[10] The equation for this model variant is:

[10] For an earlier interim report, we ran a similar version of the model for all other outcomes, finding no evidence of a lag in impact on utilization, expenditures, quality, or access. With no evidence of lag or increasing/decreasing effects over time, we limit attention in this report to the main model for utilization and expenditures analyses.

(Equation 2)

$$E\left[Y_{ist}\right] = F\left(\alpha + \sum_{s \in S_{-1}} \rho_s IS + \sum_{s \in T_{-1}} \gamma_t IT + \sum_{\tau=1}^{6} \beta_\tau I\tau + \delta X_{ist}\right)$$

where all terms are as in Equation 1, and τ indexes quarters post-implementation in the model states (six quarters for Year 1 states and two quarters for Year 2 states). The coefficients of interest here are the 16 β terms, which give the impact of the model in each post-implementation quarter.

2. Suppliers

Because the demand for supplier services varies with the demographic and health characteristics of the beneficiaries served by suppliers, we identified a catchment area for each supplier. The catchment area consisted of the set of zip codes in the supplier's state and bordering states from which at least one beneficiary received a service from the supplier. We reasoned that, if at least one beneficiary in a zip code received a service from the supplier, then other beneficiaries in that zip code could also hire the supplier if they needed or wanted ambulance transportation. Note that beneficiaries could be counted in more than one supplier's catchment area. We then calculated the average characteristics of all beneficiaries residing in the catchment area to create aggregate measures of the demographic and health characteristics of the supplier's customer base for use in constructing supplier weights and as controls in regression analysis. The characteristics included average age, percentage female, percentage white, percentage black, percentage other race, percentage residing in a rural area, percentage dually eligible for Medicare and Medicaid, average HCC score, and percentage with each of three chronic conditions: ESRD, active cancer, and pressure and non-pressure skin ulcers. More information about the HCC score appears in the beneficiary analysis section.

Suppliers in the model and comparison states differed in the demographic and health composition of their customer bases. To improve the comparability of the two groups on demographic and health characteristics, we used a statistical approach (see p. A.5) called optimal matching (Hansen and Klopfer 2006) to form matched sets of suppliers from the model and comparison states and then generated weights for comparison suppliers to create balance on important characteristics within each matched set. The goal was to minimize the difference between the aggregate characteristics of beneficiaries served by model and comparison suppliers, particularly the percentage of beneficiaries living in rural areas. We used the weights to conduct descriptive analysis based on whether or not suppliers exited the market. In Appendix D, we provide more information on the construction of the supplier weights.

Based on supplier exit decisions, we examined weighted differences between suppliers on a number of measures of service provision and payment receipt.

For each supplier, we identified all claims for ambulance services rendered by the supplier and classified them by type of service (RSNAT or non-RSNAT). We did not limit to claims for beneficiaries with ESRD and/or severe pressure ulcers; instead we included all services provided to Medicare beneficiaries. We then aggregated services provided and payments to the supplier-quarter level. We included the same ambulance service categories as in the beneficiary analyses but added another category of payments to suppliers:

- Total Medicare FFS payments: payments to suppliers for any services, including transportation, mileage, and ancillary services (such as supplemental oxygen or additional personnel)

Recognizing that suppliers may provide services to beneficiaries living in states other than the state in which their ambulances are garaged, we did not restrict the analysis to claims for services rendered to beneficiaries in these states. The supplier analysis therefore included in the calculation of outcome measures any services delivered to beneficiaries residing in states that border the model and comparison states (although we did exclude a small number of claims for beneficiaries residing in states that do not border model or comparison states).[11] In Table B.5, we present the full set of supplier descriptive measures.

Table B.5. Supplier quarterly descriptive measures

Research question	Quarterly measures
How does prior authorization affect suppliers' average number of ambulance services provided?	• Number of beneficiaries served (any Medicare ambulance) • Number of Medicare ambulance trips • Number of Medicare ambulance trips per beneficiary • Number of beneficiaries served (RSNAT) • Number of RSNAT trips provided • Number of RSNAT trips per beneficiary • Percentage of ambulance trips that are RSNAT
How does prior authorization affect average payments to ambulance suppliers?	• RSNAT payments received • Percentage of ambulance payments that are RSNAT • Percentage of all payments that are RSNAT • Total Medicare payments received

[11] Almost all supplier claims (95.1 percent) that met our definitions pertained to beneficiaries in the same state as the supplier, whereas 2.9 percent of claims were linked to beneficiaries in bordering states and 2.0 percent to beneficiaries in other states. We excluded the last group because they are unlikely to represent potential regular customers for the providers.

Appendix C

Beneficiary balancing approach and findings

This page has been left blank for double-sided copying.

This appendix describes the approach we used to balance the beneficiaries in the model and matched-comparison states, and the results of this approach. It is divided into three sections:

- Defining included beneficiaries and examining covariate balance
- Propensity score weighting approach
- Choice of adjustment approach

We discuss each subsection below.

A. Defining included beneficiaries and examining covariate balance

To be included in the beneficiary impact analysis, a beneficiary had to reside in one of the model states (Delaware; Maryland; North Carolina; New Jersey; Pennsylvania; South Carolina; Virginia; Washington, DC; or West Virginia) or in one of the matched-comparison states (Alabama, Florida, Georgia, Indiana, Kentucky, Louisiana, Massachusetts, Montana, Nebraska, Ohio, Tennessee, Texas, or Washington State). In addition, beneficiaries had to be in the Medicare fee-for-service program for at least one month in the year and have end-stage renal disease (ESRD), skin ulcers, or both. These conditions were identified through the use of hierarchical condition category (HCC; Version 22) codes HCC134, HCC136, ESRD, HCC157, or HCC158 equal to 1. Additionally, if the reason for Medicare entitlement indicated a beneficiary with ESRD, we included that person regardless of the HCC indicators.[12] Table C.1 shows the resulting number, using these inclusion criteria, of beneficiaries included in the intervention and matched-comparison states.

In this analysis, there are small differences in the counts of eligible beneficiaries in each year as compared with the analysis presented in prior reports, despite the consistent application of eligibility requirements. These differences are a result of the longer claims runout period we used in this analysis and updates the Centers for Medicare and Medicaid Services (CMS) made to enrollment information for some beneficiaries that can change their eligibility status. For example, new claims may be added that trigger a qualifying HCC code, making a beneficiary eligible who was not included in previous reports. As a result, beneficiary counts and balance statistics presented below will differ slightly from previous reports.

[12] Beneficiaries with these conditions account for around 85 percent of all RSNAT services in our study states. We did not include other beneficiaries in the analysis, because beneficiaries without these chronic conditions were less likely to require RSNAT services consistently; thus, any observed changes in RSNAT utilization would likely be the result of other causes aside from the model.

Table C.1. Counts of beneficiaries based on initial inclusion criteria

	Intervention states	Matched-comparison states
2012	156,220	300,909
2013	156,057	299,142
2014	157,788	296,840
2015	161,996	302,899
2016	169,069	316,883
2017	171,038	319,772
2018	171,806	319,428

Note: Counts between years do not represent independent observations, because many beneficiaries overlap from year to year. The model states included Delaware; Maryland; New Jersey; North Carolina; Pennsylvania; South Carolina; Virginia; Washington, DC; and West Virginia. The comparison states included Alabama, Florida, Georgia, Indiana, Kentucky, Louisiana, Massachusetts, Montana, Nebraska, Ohio, Tennessee, Texas, and Washington.

After obtaining the sample of beneficiaries, we examined the initial balance between the model and comparison beneficiaries within each year along a set of key characteristics: age, rural location, sex, and race. For this analysis, we weighted the comparison beneficiaries using the weights generated by the state-level matching procedure. In general, we found modest differences in the characteristics of beneficiaries in the intervention and comparison states. About 20 percent of beneficiaries in model states lived in rural areas, compared with about 22 percent in comparison states, whereas about 38 percent of beneficiaries in the model states were black, compared with 32 percent in the comparison states. We found very small differences for the percentages for female. We present all of these differences, separately by year, in Figures C.1 through C.7.

B. Propensity score weighting approach

To adjust for the cited differences above, we used an inverse propensity score weighting approach, which involves two steps.[13] First, for each year of data, we estimated a weighted logistic regression, predicting study status (living in a model or a comparison state) based on the following set of characteristics: beneficiary age; gender; race (separate indicators for white or black, with all other races grouped together as the reference category); whether the person lived in a rural area; and indicators for whether the beneficiary had active cancer, ESRD, or skin ulcers.[14] We included beneficiaries from the Year 1 and Year 2 model states and their matched comparison states in this analysis. These regressions provided predicted probabilities for each

[13] We initially tried a calibration approach that involved calculating the weights by using discrete strata, defined by a set of key characteristics. Although this approach eliminated imbalance on the set of characteristics used in the weighting, the resulting design effects (which increase the variance of the outcomes) were unacceptably large. In addition, statistical matching was infeasible with the large number of beneficiaries observed in each year.

[14] When the study began, we considered including beneficiaries with active cancer, ESRD, or skin ulcers, as they were responsible for 97 percent of all RSNAT use. As a result, we included indicators for each of these three medical conditions in the regression model. However, after further examination of RSNAT usage, we decided to limit to only the population with ESRD, skin ulcers, or both, which has a much higher likelihood of utilization and are therefore of more interest in estimating the effects of prior authorization.

beneficiary that represent the likelihood each beneficiary lives in a state with prior authorization (Rosenbaum and Rubin 1983). The second stage of this process was calculating weights ω, for each beneficiary, defined as

(Equation 3) $$\omega(W,\chi) = W + (1-W)\frac{\hat{e}(x)}{1-\hat{e}(x)},$$

where W = 1 if a beneficiary lived in a model state, W = 0 if a beneficiary lived in a comparison state, x represents the set of characteristics included in the propensity score model, and $\hat{e}(\chi)$ represents the estimated propensity score (Guo and Fraser 2009). These propensity score weights reduce to 1 for beneficiaries living in model states and $\frac{\hat{e}(x)}{1-\hat{e}(x)}$ for beneficiaries living in comparison states. We then combined these weights with the state-level matching weights to form the beneficiary analysis weights.[15] As the following figures show, the analysis weights were highly effective in reducing imbalance on the key characteristics. In addition, the design effects of these weights were about 1.28 in each year—a negligible increase in variance over the state-level matching weights, which had design effects of 1.19.

In Figures C.1 through C.7, we show for each year (2012 through 2018) the standardized differences between the beneficiaries in the model states and those in the comparison states, weighted by the state-level matching weight (dark blue) and the propensity score adjusted analysis weight (red). The vertical bars demonstrate the size of the differences before and after the calibrations. Each figure also includes a data table with the numeric values for the standardized differences.[16] Table C.2 summarizes balance on the characteristics included in the calculation of the propensity score weights within each year.

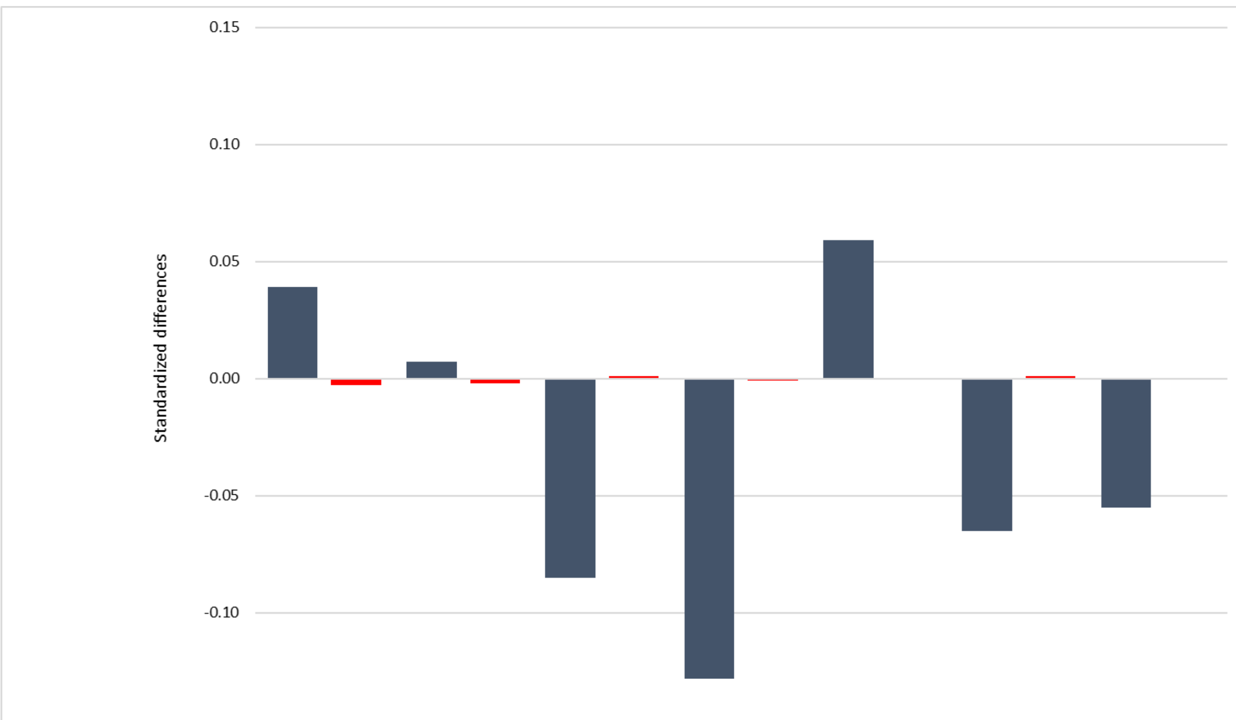
[15] To develop the weights, we used beneficiaries' state of residence as of December 31 or date of death for each study year. For example, a beneficiary designated as residing in a comparison state on December 31, 2012, would receive a comparison weight for each quarter of 2012 where they met the eligibility requirements.

[16] As discussed above, balance may differ slightly with data reported in earlier reports, due to small changes in the eligible beneficiary population resulting from longer claims runout.

Figure C.1. Beneficiary balance before and after propensity score adjustments, 2012

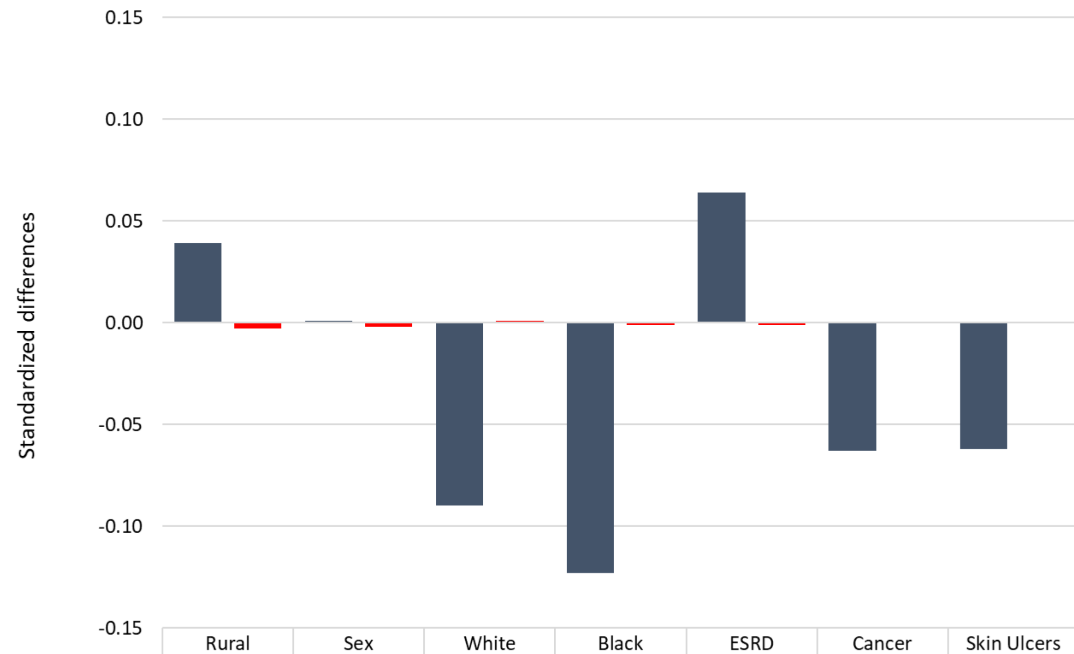
	Rural	Sex	White	Black	ESRD	Cancer	Skin Ulcers
State-level matching weights	0.04	0.002	-0.08	-0.134	0.055	-0.059	-0.051
Analysis weights	-0.003	-0.002	0	0	0	0	-0.001

Figure C.2. Beneficiary balance before and after propensity score adjustments, 2013



	Rural	Sex	White	Black	ESRD	Cancer	Skin Ulcers
State-level matching weights	0.039	0.007	-0.085	-0.128	0.059	-0.065	-0.055
Analysis weights	-0.003	-0.002	0.001	-0.001	0	0.001	0

Figure C.3. Beneficiary balance before and after propensity score adjustments, 2014



	Rural	Sex	White	Black	ESRD	Cancer	Skin Ulcers
■ State-level matching weights	0.039	0.001	-0.09	-0.123	0.064	-0.063	-0.062
■ Analysis weights	-0.003	-0.002	0.001	-0.001	-0.001	0	0

Figure C.4. Beneficiary balance before and after propensity score adjustments, 2015

	Rural	Sex	White	Black	ESRD	Cancer	Skin Ulcers
■ State-level matching weights	0.042	-0.001	-0.084	-0.123	0.087	-0.062	-0.086
■ Analysis weights	-0.003	-0.002	0.001	-0.001	-0.001	0.001	0.001

Figure C.5. Beneficiary balance before and after propensity score adjustments, 2016

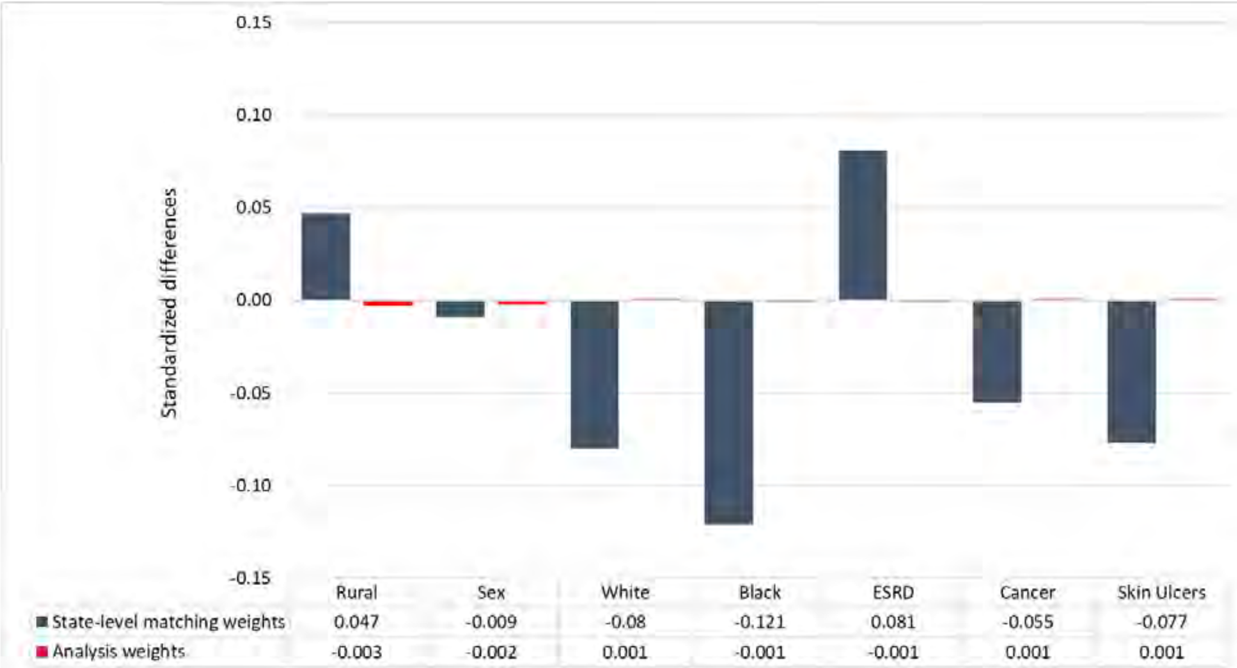


Figure C.6. Beneficiary balance before and after propensity score adjustments, 2017

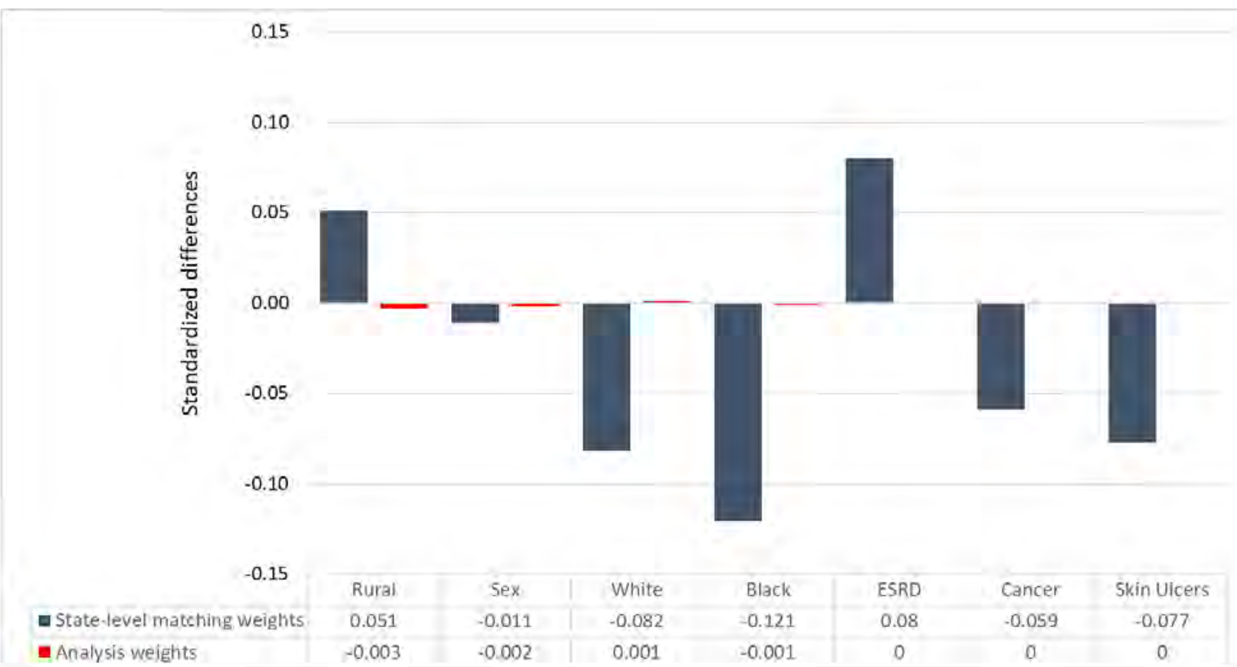
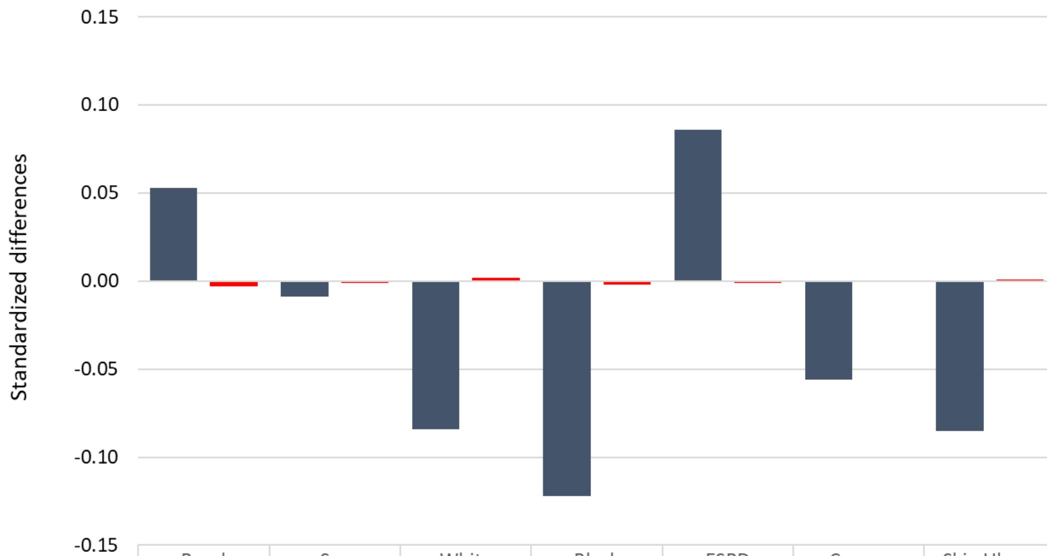


Figure C.7. Beneficiary balance before and after propensity score adjustments, 2018



	Rural	Sex	White	Black	ESRD	Cancer	Skin Ulcers
State-level matching weights	0.053	-0.009	-0.084	-0.122	0.086	-0.056	-0.085
Analysis weights	-0.003	-0.001	0.002	-0.002	-0.001	0	0.001

Table C.2. Beneficiary characteristics before and after propensity score weighting

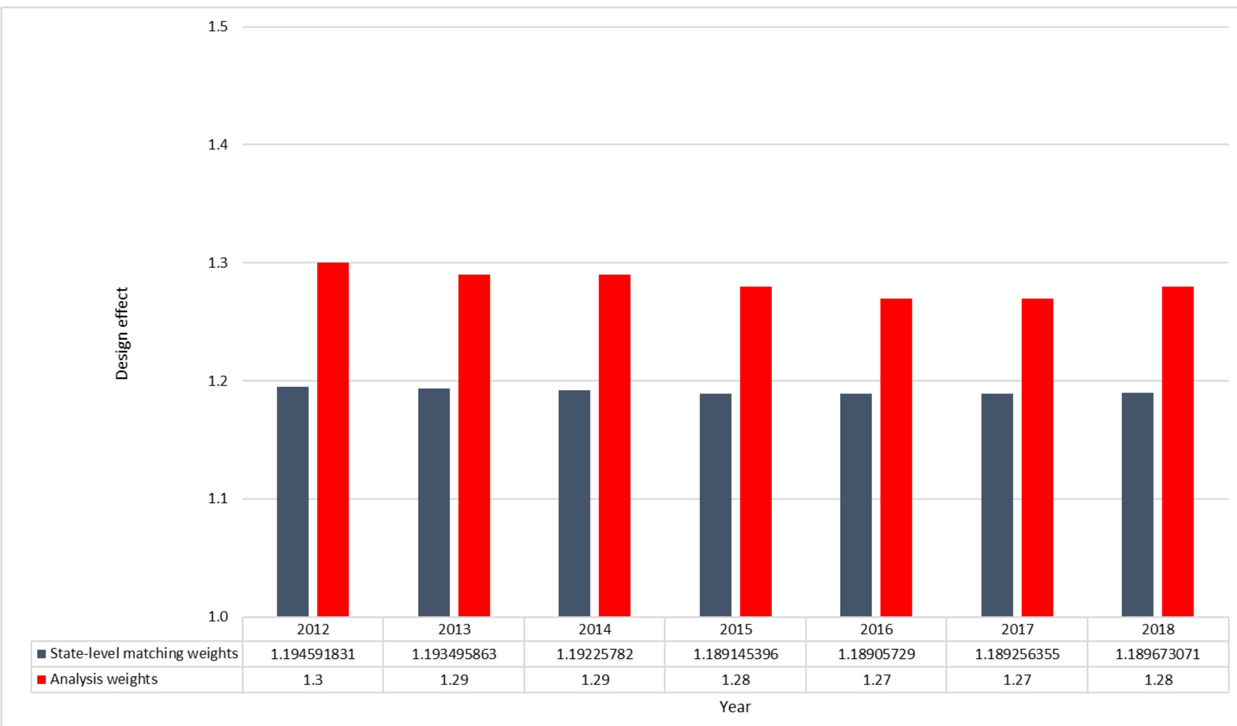
	2012			2013			2014			2015			2016			2017			2018		
	Model	Comparison		Model	Comparison		Model	Comparison		Model	Comparison		Model	Comparison		Model	Comparison		Model	Comparison	
Characteristic		Before	After		Before	After		Before	After		Before	After		Before	After		Before	After		Before	After
Rural	20.4	22.0	20.3	20.3	21.9	20.2	20.2	21.8	20.1	20.4	22.1	20.3	20.5	22.4	20.4	20.4	22.5	20.3	20.5	22.6	20.4
Female	48.0	48.1	47.9	47.5	47.8	47.4	47.4	47.5	47.3	47.5	47.4	47.4	47.9	47.5	47.8	47.6	47.1	47.6	47.1	46.7	47.1
Race																					
White	55.1	51.1	55.1	55.0	50.8	55.0	54.9	50.4	54.9	55.3	51.2	55.4	56.2	52.2	56.2	56.2	52.1	56.2	56.1	51.9	56.2
Black	38.4	31.9	38.4	38.2	32.0	38.2	38.0	32.1	38.0	37.3	31.4	37.3	36.3	30.5	36.3	36.0	30.3	36.0	35.9	30.0	35.8
With active cancer	14.0	11.9	14.0	14.1	11.9	14.1	14.2	12.0	14.2	14.5	12.3	14.5	14.3	12.4	14.4	14.7	12.6	14.7	14.8	12.8	14.8
With end-stage renal disease	77.1	79.4	77.1	76.0	78.5	76.0	75.0	77.8	74.9	71.9	75.8	71.8	68.0	71.8	68.0	67.7	71.5	67.7	67.7	71.7	67.6
With skin ulcers	25.3	23.1	25.3	26.6	24.2	26.6	27.7	25.0	27.7	31.1	27.1	31.1	35.4	31.7	35.4	35.8	32.1	35.8	36.0	31.9	36.0

Note: The model states included Delaware; Maryland; New Jersey; North Carolina; Pennsylvania; South Carolina; Virginia; Washington, DC; and West Virginia. The comparison states included Alabama, Florida, Georgia, Indiana, Kentucky, Louisiana, Massachusetts, Montana, Nebraska, Ohio, Tennessee, Texas, and Washington. In May 2017, the CMS Virtual Research Data Center updated the source of enrollment and eligibility data from the Enrollment Database to the Common Medicare Environment. One result of this change was more beneficiaries classified as Hispanic, rather than white, based on an algorithm CMS used that reassigns beneficiaries using first and last names. As a result, the percentage of white beneficiaries decreased by less than 2 percentage points in the model states and about 7 percentage points in the comparison states. This change does not threaten the validity of the analysis contained in this report, as the analysis weights were still able to remove any imbalance on the percentage of white beneficiaries between the model and comparison states. Specific values may differ from previous reports due to small changes in the eligible populations of beneficiaries due to longer claims runout.

C. Choice of adjustment approach

A common approach to reduce or eliminate imbalance between an intervention and comparison sample calls for selecting a subset of the possible comparison units that are the closest matches to treated units, often through using a propensity score. We initially considered such an approach but rejected it in favor of the propensity score weighting approach. The main reason we decided against the use of a matching approach relates to computational efficiency. Regardless of the exact details of a matching task, it is essential to construct a distance matrix, which contains the calculated distance between each treated and untreated unit. A computer algorithm then searches the matrix to select the matches based on the chosen matching criteria. With so many beneficiaries observed in the model and comparison states in each year, reliance on such a matrix would be unwieldy. Initial tests with basic distance measures suggest that implementing this approach would be time-consuming. The propensity score weighting approach, on the other hand, took considerably less time to design and implement, and presented no computational difficulties, as it did not rely on iteratively searching for the best set of matches for all individual beneficiaries. However, we had to examine the design effect caused by differential weights to ensure that the weights did not substantially increase the standard errors of our estimates. We show the design effects due to weighting before and after propensity score adjustment in Figure C.8.

Figure C.8. Design effects, by weight and year



The design effect due to weighting is a measure of the increase in the variance of an outcome measure induced by weights. It is scaled to a simple random sample (SRS), which has no differential weights and therefore no design effect. Therefore, a design effect of 1.0 indicates no increase in variance due to weighting, whereas a design effect of 1.5 indicates a 50 percent increase in variance attributable to weighting. The state-level matching weights had design effects of about 1.19 across all years, making the variance of the outcomes about 20 percent higher than a SRS using just the state-level matching weights. The propensity score adjusted weights increased the design effects to about 1.28. Such an increase in the variance of the outcomes is marginal and will increase the size of confidence intervals of estimates by only fractions of a percent. Therefore, we find that the propensity score adjustment approach does not significantly reduce the statistical power of our analyses.

We believe that the choice of propensity score weighting provides impact estimates on RSNAT utilization and access as equally unbiased as those potentially resulting from a matching approach (Wang et al. 2014; Posner and Ash n.d.). The goal of matching is to eliminate imbalance on important measured characteristics; the figures above clearly show that the propensity score weighting approach eliminated those imbalances. In this project, the weights slightly increased the variance of the state-level matching weights. However, an increase in the variance of the weights would almost certainly also result from a matching approach. The propensity score weighting approach also included all comparison beneficiaries, resulting in a larger sample size than would be expected with the use of a matching approach, which improves statistical precision.

D. Summary of balance on beneficiary characteristics

Before running the impact analysis, we confirmed that beneficiaries in the comparison states were similar on average to those in the model states, across the full range of data, rather than within each year as in Table C.2 above. After applying calibration weights, we compared the means between the two groups for baseline demographic and health characteristics. All but one group difference was about 2 percent or less of the model group mean. The exception is a higher proportion of comparison beneficiaries who are dually eligible for Medicare and Medicaid. Table C.3 contains weighted summary statistics for fee-for-service (FFS) beneficiaries with ESRD, pressure ulcers, or both, in model and comparison states, averaged across all years.

Table C.3. Beneficiary summary statistics at baseline (weighted)

	Model mean (SD)	Comparison mean (SD)	Difference	Percentage difference (%)
Age (years)	71.5 (15.0)	71.5 (15.0)	-0.0	-0.1
Female (percentage)	50.8 (50.0)	50.2 (50.0)	0.6***	1.2
Race (percentage)				
White	64.1 (47.96)	64.7 (47.80)	-0.5***	-0.8
Black	30.4 (46.00)	29.8 (45.76)	0.6***	1.8

	Model mean (SD)	Comparison mean (SD)	Difference	Percentage difference (%)
Other	5.5 (22.72)	5.5 (22.80)	-0.0	-0.7
Rural (percentage)	20.2 (40.12)	20.4 (40.27)	-0.2*	-1.0
Dual (percentage)	33.9 (47.34)	36.8 (48.24)	-2.9***	-8.6
HCC score	4.3 (2.70)	4.3 (2.72)	-0.0*	-0.2
Chronic condition (percentage)				
ESRD only	56.2 (49.61)	56.4 (49.59)	-0.1	-0.3
Pressure ulcers only	41.7 (49.30)	41.6 (49.28)	0.1	0.2
ESRD and pressure ulcers	2.11 (14.38)	2.1 (14.24)	0.0	1.9
Number of beneficiaries	398,088	746,892		

Note: This table presents weighted means (and standard deviations) of beneficiary characteristics for beneficiaries with ESRD, pressure ulcers, or both. Comparison group individuals are propensity-score weighted to resemble model state individuals on baseline demographic and health characteristics. The model states were Delaware, Maryland, New Jersey, North Carolina, Pennsylvania, South Carolina, Virginia, West Virginia, and the District of Columbia. The comparison states were Alabama, Florida, Georgia, Indiana, Kentucky, Louisiana, Massachusetts, Montana, Nebraska, Ohio, Tennessee, Texas, and Washington.

*p < 0.05, **p < 0.01, ***p < 0.001.

ESRD = end-stage renal disease; HCC = hierarchical condition category; SD = standard deviation.

These statistics suggest that the comparison states provide a suitable basis of comparison to the model states. (Appendix A describes how the comparison states were selected.)

This page has been left blank for double-sided copying.

Appendix D

Supplier matching approach and findings

This page has been left blank for double-sided copying.

A. Supplier characteristics and balance

For this analysis, we defined suppliers using a combination of National Provider Identifier (NPI)/provider ID number, state, and zip code. This approach addresses the complication of single suppliers operating in multiple states. Using this definition, there were, on average, about 65 percent more providers in the matched-comparison states than the model states in each year (Table D.1).

Table D.1. Counts of RSNAT service suppliers, by year

	2012	2013	2014	2015	2016	2017	2018
Model states	2,451	2,452	2,376	2,263	2,181	2,156	2,080
Matched-comparison states	4,005	3,869	3,801	3,779	3,814	3,845	3,793

Note: Counts are unweighted. The model states included Delaware; Maryland; New Jersey; North Carolina; Pennsylvania; South Carolina; Virginia; Washington, DC; and West Virginia. The comparison states included Alabama, Florida, Georgia, Indiana, Kentucky, Louisiana, Massachusetts, Montana, Nebraska, Ohio, Tennessee, Texas, and Washington.

To assess the similarity of the suppliers in the matched-comparison states with those in the model states, we used the aggregated characteristics of the potential service users living within the catchment areas of each supplier. We chose this approach for two reasons. First, very limited data were available at the supplier level—primarily whether the supplier was located in a rural or urban area and the ownership type for the supplier. Ownership type was not very informative, because it was available only for the institutional ambulance providers, and we considered rural/urban location potentially suspect because this designation was based solely on the zip code of the supplier's location, not the areas where services are provided. A supplier may choose to locate its garage in a rural area due to lower property taxes, rent, or other reasons, yet the majority of the beneficiaries served would receive services in an urban location. If prior authorization has a differential impact on the services available to rural versus urban beneficiaries, using the supplier zip codes could suppress the discovery of such an effect.

Potential service users were defined as beneficiaries diagnosed with end-stage renal disease (ESRD), skin ulcers, or both, and who lived within a supplier catchment area, the latter defined as the set of zip codes of previous RSNAT-covered claims for the supplier (see the methods section for more information). The data available for the beneficiaries within the supplier catchment areas included age, race, urban/rural locality, and chronic conditions. We aggregated these characteristics to the supplier level, weighted by the state-level matching weight in the case of the suppliers in matched-comparison states. Table D.2 shows the average of these characteristics for each year.

As Table D.2 shows, the suppliers were similar on the average age, percentage female, percentage rural, and percentage with ESRD, active cancer, or skin ulcers. There are differences for the average percentages rural, white, and black, but they are mostly moderate differences, on the order of about 2 to 4 percentage points or less.

Table D.2. Summary statistics for aggregated potential customer base characteristics in catchment areas, by year

	2012		2013		2014		2015		2016		2017		2018	
	Model	Comparison	Model	Comparison	Model	Comparison	Model	Comparison	Model	Comparison	Model	Comparison	Model	Comparison
Average age	71.1	70.9	71.0	70.8	71.0	70.8	71.1	70.9	71.2	70.9	71.3	71.1	71.4	71.4
Percentage female	55.7	55.3	55.6	55.1	55.4	55.0	55.3	55.0	55.1	54.8	55.1	54.8	55.0	54.9
Percentage rural	23.8	24.6	23.6	24.8	23.5	24.6	24.1	25.4	24.7	25.5	25.2	24.4	25.4	23.3
Percentage white	79.0	82.4	77.7	81.4	77.1	80.6	78.2	80.1	78.6	78.9	78.5	80.0	78.4	80.5
Percentage black	15.0	11.3	15.7	11.8	15.6	12.1	14.5	12.5	14.1	13.2	13.9	11.7	13.8	11.1
Percentage with end-stage renal disease	1.5	1.4	1.5	1.4	1.5	1.4	1.4	1.3	1.4	1.4	1.4	1.3	1.3	1.2
Percentage with active cancer	8.2	7.8	8.1	7.6	8.0	7.4	8.0	7.3	7.8	7.1	7.6	7.0	7.6	7.4
Percentage with skin ulcers	2.2	1.9	2.2	1.9	2.2	1.8	1.8	1.5	1.9	1.7	1.9	1.7	1.9	1.8

Note: All rows labeled as percentages of a category represent averages of within-catchment area beneficiary characteristics, aggregated to the provider level. The model states included Delaware; Maryland; New Jersey; North Carolina; Pennsylvania; South Carolina; Virginia; Washington, DC; and West Virginia. The comparison states included Alabama, Florida, Georgia, Indiana, Kentucky, Louisiana, Massachusetts, Montana, Nebraska, Ohio, Tennessee, Texas, and Washington.

B. Matching approach

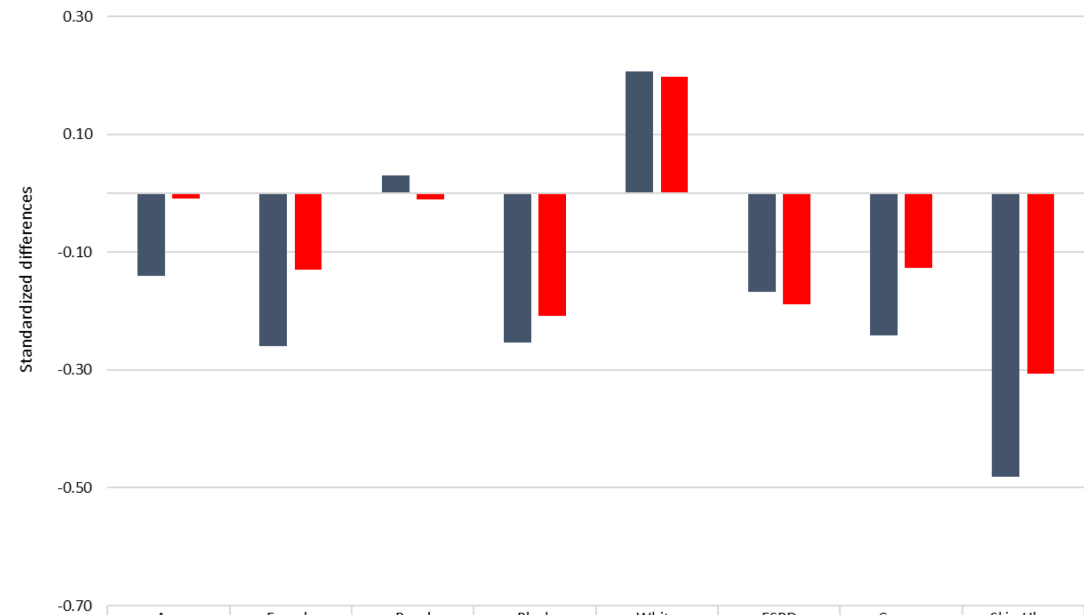
We used the statistical technique of optimal matching to select a sample of suppliers from the matched-comparison states similar to the suppliers in the model states on aggregated characteristics of beneficiaries within their service areas.[17] The priority for matching was to maintain the imbalance on the percentage of rural potential service users to less than 0.10 standard deviations (SD), a common threshold in evaluations for the Centers for Medicare and Medicaid Innovation (CMMI); we considered any larger imbalance on rural location a threat to the validity of the impact estimates. Our secondary goals were to minimize the differences on the remaining characteristics of interest (for example, percentage with ESRD or skin ulcers, race, gender) and keep the design effect due to weighting in the 1.7–1.8 range, disallowing any above 2.0.[18] After achieving a minimal difference on the percentage of rural service users (< 0.10 SD), we used these two goals to strike a balance between reducing differences on other characteristics and minimizing the effect of the weights on the variance of the outcomes. Separately for each year, we tested a series of matching specifications, with variations on a caliper on the percentage of rural beneficiaries, the propensity-score based distance for the combination of all variables, and the minimum and maximum matching ratio within the matched sets. We chose final matching specifications based primarily on the reduction in the differences on percentage of rural potential service users, as well as balance on the remaining characteristics, the resulting design effect of the matching weights, and changes in the weights for suppliers across years.

Figures D.1 through D.7 present the standardized differences on the key characteristics used in matching before and after that matching. In general, the matching improved balance on all characteristics, with the post-match balance on percentage of rural beneficiaries meeting the target of less than 0.10 SD in all years. Although the matching did not consistently achieve that level of balance on other characteristics, these mean differences were nevertheless quite small in raw terms. Table D.3 presents the weighted post-match means. After the matching procedure, the difference in the average percentage of potential service users living in rural areas was less than or equal to 2 percentage points in any year. Differences for all other characteristics marginally improved compared with their prior-to-matching balance.

[17] We initially used a propensity-score weighting approach to balance the provider characteristics, similar to the approach deployed for the beneficiary matching analysis. However, we found that complex modeling was required to achieve balance similar to what we could achieve using the matching. In addition, the propensity score-adjusted weights had design effects that were much higher than the matching-based weights. As a result, we decided to use the matching analysis instead.

[18] Higher design effects would have an adverse impact on the precision of impact estimates.

Figure D.1. Supplier balance on aggregated beneficiary characteristics before and after matching, 2012

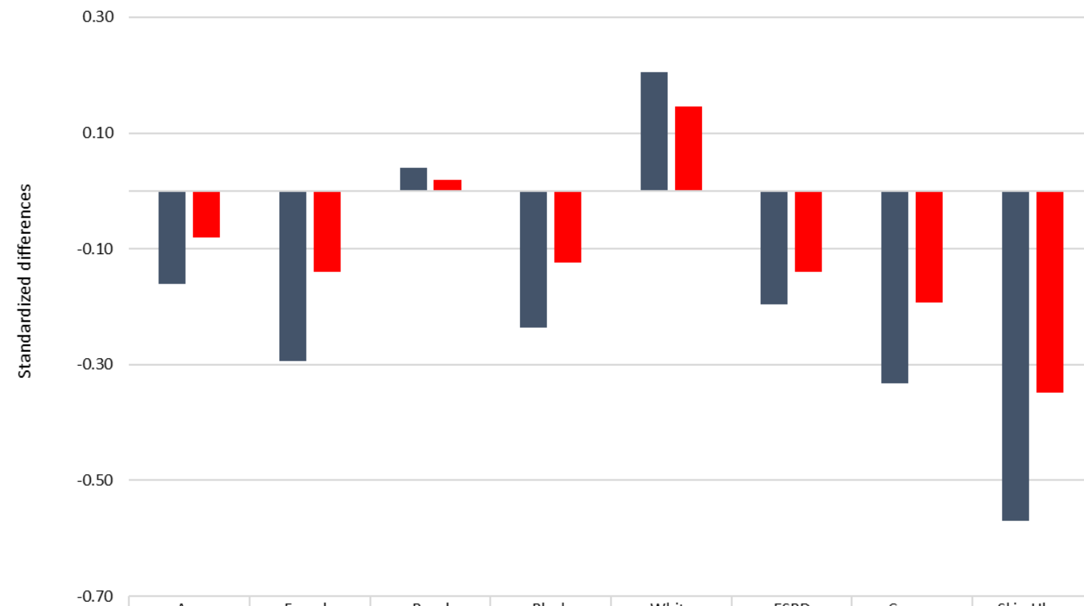


	Age	Female	Rural	Black	White	ESRD	Cancer	Skin Ulcer
State-level matching weights	-0.14	-0.26	0.03	-0.25	0.21	-0.17	-0.24	-0.48
Analysis weights	-0.01	-0.13	-0.01	-0.21	0.20	-0.19	-0.13	-0.31

Figure D.2. Supplier balance on aggregated beneficiary characteristics before and after matching, 2013

	Age	Female	Rural	Black	White	ESRD	Cancer	Skin Ulcer
State-level matching weights	-0.16	-0.29	0.04	-0.26	0.22	-0.17	-0.32	-0.55
Analysis weights	-0.05	-0.16	0.00	-0.18	0.17	-0.19	-0.16	-0.34

Figure D.3. Supplier balance on aggregated beneficiary characteristics before and after matching, 2014



	Age	Female	Rural	Black	White	ESRD	Cancer	Skin Ulcer
State-level matching weights	-0.16	-0.29	0.04	-0.24	0.20	-0.20	-0.33	-0.57
Analysis weights	-0.08	-0.14	0.02	-0.12	0.15	-0.14	-0.19	-0.35

Figure D.4. Supplier balance on aggregated beneficiary characteristics before and after matching, 2015

	Age	Female	Rural	Black	White	ESRD	Cancer	Skin Ulcer
State-level matching weights	-0.24	-0.24	0.04	-0.15	0.12	-0.09	-0.39	-0.58
Analysis weights	-0.10	-0.10	0.00	-0.07	0.08	-0.10	-0.23	-0.35

Figure D.5. Supplier balance on aggregated beneficiary characteristics before and after matching, 2016

	Age	Female	Rural	Black	White	ESRD	Cancer	Skin Ulcer
State-level matching weights	-0.27	-0.22	0.03	-0.08	0.02	0.01	-0.41	-0.60
Analysis weights	-0.17	-0.08	-0.01	0.00	0.00	-0.02	-0.20	-0.37

Figure D.6. Supplier balance on aggregated beneficiary characteristics before and after matching, 2017

	Age	Female	Rural	Black	White	ESRD	Cancer	Skin Ulcer
State-level matching weights	-0.22	-0.23	-0.03	-0.19	0.11	-0.12	-0.39	-0.51
Analysis weights	-0.08	-0.08	-0.07	-0.13	0.09	-0.15	-0.15	-0.26

Figure D.7. Supplier balance on aggregated beneficiary characteristics before and after matching, 2018

Table D.3. Summary statistics for aggregated beneficiary characteristics in catchment areas, by year, post-matching

	2012		2013		2014		2015		2016		2017		2018	
	Model	Comparison	Model	Comparison	Model	Comparison	Model	Comparison	Model	Comparison	Model	Comparison	Model	Comparison
Average age	71.1	71.1	71.0	71.0	71.0	70.9	71.1	71.0	71.2	71.0	71.3	71.2	71.4	71.4
Percentage female	55.7	55.5	55.6	55.3	55.4	55.2	55.3	55.1	55.1	55.0	55.1	55.0	55.0	54.9
Percentage rural	23.8	23.5	23.6	23.6	23.5	24.0	24.1	24.1	24.7	24.5	25.2	23.1	25.4	23.3
Percentage white	79.0	82.3	77.7	80.7	77.1	79.6	78.2	79.4	78.6	78.6	78.5	79.7	78.4	80.5
Percentage black	15.0	12.0	15.7	13.0	15.6	13.8	14.5	13.6	14.1	14.2	13.9	12.4	13.8	11.1
Percentage with end-stage renal disease	1.5	1.4	1.5	1.4	1.5	1.4	1.4	1.3	1.4	1.4	1.4	1.3	1.3	1.2
Percentage with active cancer	8.2	8.0	8.1	7.8	8.0	7.6	8.0	7.6	7.8	7.5	7.6	7.4	7.6	7.4
Percentage with skin ulcers	2.2	2.1	2.2	2.0	2.2	2.0	1.8	1.6	1.9	1.8	1.9	1.8	1.9	1.8

Note: The model states included Delaware; Maryland; New Jersey; North Carolina; Pennsylvania; South Carolina; Virginia; Washington, DC; and West Virginia. The comparison states included Alabama, Florida, Georgia, Indiana, Kentucky, Louisiana, Massachusetts, Montana, Nebraska, Ohio, Tennessee, Texas, and Washington.

The final matching specifications, particularly the maximum of three model suppliers matched to a single comparison supplier (or vice versa), was informed by the year-to-year change in weights for each supplier, as well as the imbalance on the percentage of rural service users. Initial matches allowed for more flexible ratios of model to comparison suppliers, as high as 10-to-one and vice versa, to reduce the imbalances; however, this approach resulted in large changes in the weights for individual suppliers across years, creating an artificial seasonal cycle in the outcome measures. As a result, we restricted the matching ratios in the final matching specifications to minimize changes in weights across years, which also reduced the design effect of the weights. See Appendices F and G for a discussion of the statistical precision of the supplier-level analyses.

This page has been left blank for double-sided copying.

Appendix E

Quantitative results

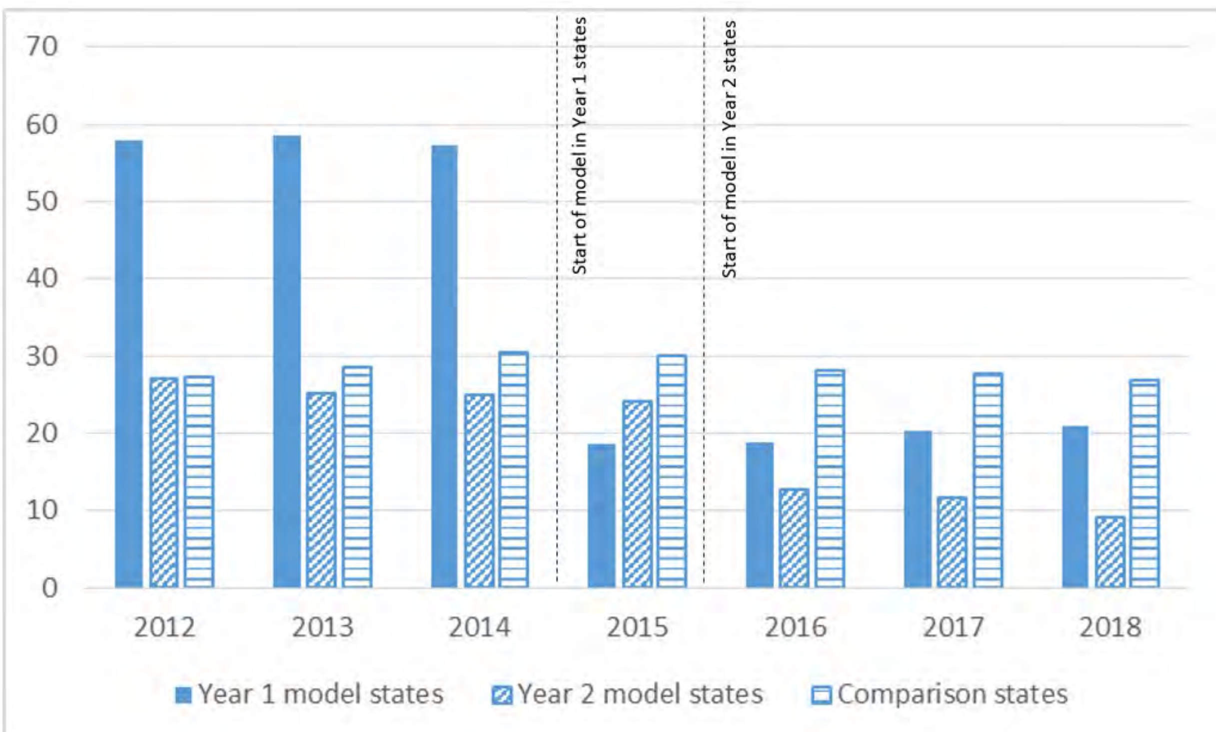
This page has been left blank for double-sided copying.

In this appendix, we present the additional quantitative analysis results we referred to in the quantitative results section of the report. We have divided this appendix into subsections on descriptive results, beneficiary summary statistics, utilization and expenditures, quality of and access to care, and supplier experience.

A. Descriptive results

In the quantitative results section, we presented unadjusted beneficiary-level utilization and expenditures as well as overall expenditure figures to provide an estimate of the Medicare savings on repetitive, scheduled, non-emergent ambulance transport (RSNAT) services attributable to the model. Here we provide results of descriptive analyses of Medicare utilization and expenditures per 100,000 FFS beneficiaries. Figure E.1 shows that the number of RSNAT trips per 100,000 Medicare fee-for-service (FFS) months declined in Year 1 model states following implementation in December 2014, and in Year 2 expansion states following implementation in January 2016. The same pattern holds for expenditures on RSNAT services normalized to Medicare population size, as seen in Figure E.2.

Figure E.1. RSNAT trips per 100,000 Medicare FFS months

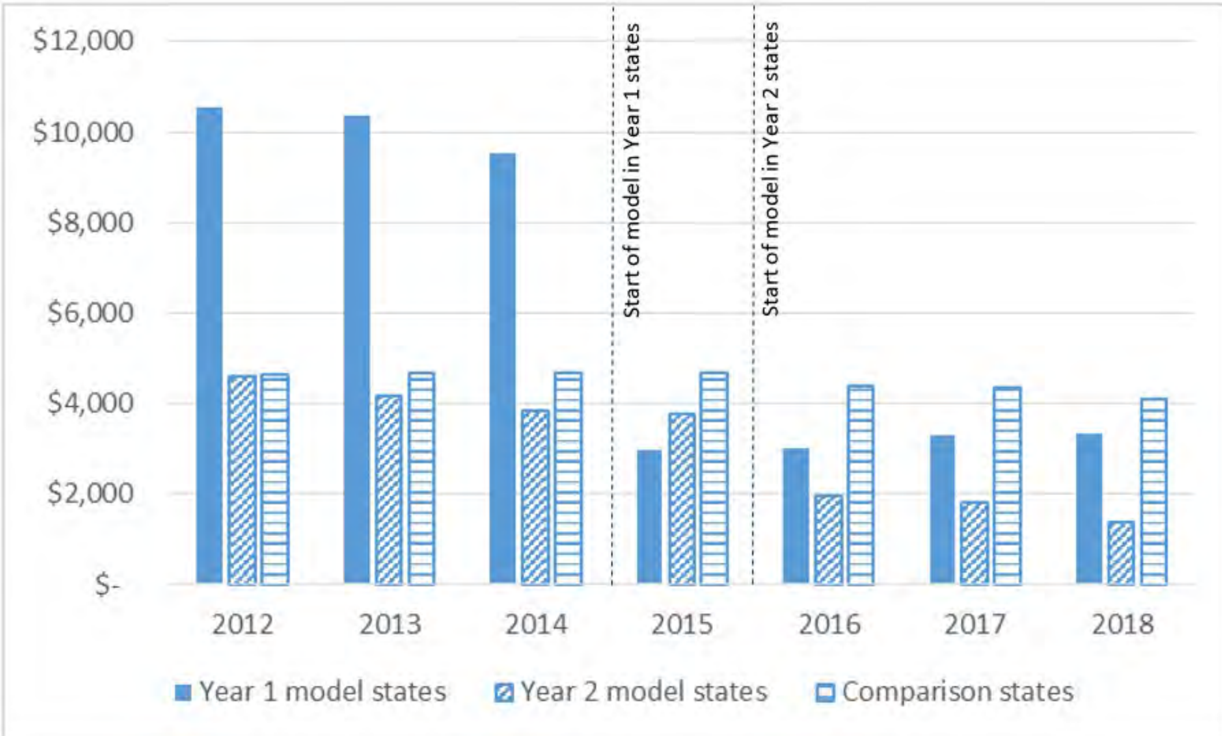


Source: Medicare FFS claims, January 2012 through December 2018.

Note: Year 1 model states included: New Jersey, Pennsylvania, and South Carolina. Year 2 expansion states included: Delaware; Maryland; North Carolina; Virginia; Washington, DC; and West Virginia. Comparison states included: Alabama, Florida, Georgia, Indiana, Kentucky, Louisiana, Massachusetts, Montana, Nebraska, Ohio, Tennessee, Texas, and Washington.

FFS = fee-for-service; RSNAT = repetitive, scheduled, non-emergent ambulance transport.

Figure E.2. RSNAT payments per 100,000 Medicare FFS months



Source: Medicare FFS claims, January 2012 through December 2018.

Note: Year 1 model states included: New Jersey, Pennsylvania, and South Carolina. Year 2 expansion states included: Delaware; Maryland; North Carolina; Virginia; Washington, DC; and West Virginia. Comparison states included: Alabama, Florida, Georgia, Indiana, Kentucky, Louisiana, Massachusetts, Montana, Nebraska, Ohio, Tennessee, Texas, and Washington.

FFS = fee-for-service; RSNAT = repetitive, scheduled, non-emergent ambulance transport.

B. Utilization and expenditures

1. Descriptive characteristics

Before the model was implemented, beneficiaries with ESRD and/or pressure ulcers in model states had higher quarterly utilization of and expenditures for ambulance services, with RSNAT utilization over 30 percent higher and expenditures nearly 40 percent higher. Table E.1 shows the baseline levels of utilization and expenditures in the model and comparison states. This difference is by design—CMS selected the states with the highest rates of RSNAT use to be subject to the prior authorization (Year 1 states). Other than expenditures related to RSNAT, the composition of baseline total expenditures was comparable in the two groups of states with one exception: model states had much lower home health expenditures than comparison states.[19]

[19] The difference arose primarily from a greater proportion of comparison state beneficiaries having any home health expenditures (18 percent versus 13 percent in model states). Among beneficiaries with any expenditures in a quarter, expenditure amounts were similar between model ($3400) and comparison states ($3600).

Table E.1. Baseline quarterly utilization and expenditures per beneficiary

	Quarterly model mean (SD)	Quarterly comparison mean (SD)	Difference	Percentage difference (%)
Utilization				
Probability of RSNAT ambulance utilization (%)	4.1 (19.8)	2.8 (16.5)	1.3***	31.4
Number of RSNAT ambulance trips	2.2 (11.9)	1.4 (9.4)	0.8***	36.4
Probability of any Medicare ambulance utilization (%)	24.8 (43.2)	24.4 (42.9)	0.4***	1.7
Total number of Medicare ambulance trips	2.7 (12.2)	2.0 (9.7)	0.8***	29.2
Expenditures				
RSNAT expenditures ($)	370 (2,049)	223 (1,533)	146***	39.6
All Medicare ambulance expenditures ($)	642 (2,712)	471 (2,160)	171***	26.6
Total Medicare FFS expenditures ($)	15,904 (29,411)	15,629 (27,259)	276***	1.7
Major expenditures categories				
Outpatient expenditures ($)	3,998 (5,203)	3,833 (4,916)	165***	4.1
Professional services[a] expenditures ($)	2,785 (4,596)	2,6232 (4,199)	162***	5.8
Inpatient expenditures ($)	6,519 (23,368)	6,502 (21,975)	17	0.3
Skilled nursing facility expenditures ($)	1,296 (4,790)	1,263 (4,791)	33***	2.6
Home health expenditures	440 (1,403)	652 (1,740)	-211***	-48.0
Number of beneficiary-quarter observations	1,950,170	3,998,231		

Note: The table presents baseline weighted means (and standard deviations) of quarterly beneficiary utilization and expenditures outcomes for beneficiaries with ESRD and/or pressure ulcers in the model and comparison state groups. Baseline period for Year 1 states is 2012–2014. Baseline period for Year 2 states is 2012–2015. Comparison group individuals are weighted to resemble model state individuals on baseline demographic and health characteristics. The model states were Delaware, Maryland, New Jersey, North Carolina, Pennsylvania, South Carolina, Virginia, West Virginia, and the District of Columbia. The comparison states were Alabama, Florida, Georgia, Indiana, Kentucky, Louisiana, Massachusetts, Montana, Nebraska, Ohio, Tennessee, Texas, and Washington.

[a] Professional services providers include physicians, physician assistants, clinical social workers, and nurse practitioners, as well as some organizational providers such as independent clinical laboratories, ambulance providers, free-standing ambulatory surgical centers, and free-standing radiology centers.

*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

ESRD = end-stage renal disease; FFS = fee-for-service; RSNAT = repetitive, scheduled, non-emergent ambulance transport; SD = standard deviation.

Baseline rates of quality of care measures were similar for model and comparison states (Table E.2). Although many of the baseline differences between model and comparison states are significantly different from zero statistically, these differences were quite small in magnitude.

Table E.2. Baseline measures of quality of care per beneficiary per quarter

	Model mean (SD)	Comparison mean (SD)	Difference	Percentage difference (%)
Probability of emergency department utilization (%)	35.9 (48.0)	35.4 (47.8)	0.5***	1.3
Number of emergency department visits	0.7 (1.3)	0.6 (1.3)	0.0***	1.5
Probability of emergency ambulance utilization (%)	18.8 (39.1)	19.9 (39.9)	-1.1***	-5.8
Number of emergency ambulance trips	0.3 (0.9)	0.3 (0.9)	-0.0***	-6.5
Probability of unplanned hospital admission (%)	26.3 (44.0)	26.1 (43.9)	-0.2**	0.7
Probability of death (%)	5.4 (22.7)	5.6 (23.0)	-0.1***	-2.6
Number of observations	1,950,170	3,998,231		

Note: This table presents baseline weighted means (and standard deviations) of quarterly beneficiary quality of care and access to treatment outcomes for beneficiaries with ESRD and/or pressure ulcers. Baseline period for Year 1 states is 2012–2014. Baseline period for Year 2 states is 2012–2015.Comparison group individuals are weighted to resemble model state individuals on baseline demographic and health characteristics. The model states were Delaware, Maryland, New Jersey, North Carolina, Pennsylvania, South Carolina, Virginia, West Virginia, and the District of Columbia. The comparison states were Alabama, Florida, Georgia, Indiana, Kentucky, Louisiana, Massachusetts, Montana, Nebraska, Ohio, Tennessee, Texas, and Washington.

*p < 0.05, **p < 0.01, ***p < 0.001.

ESRD = end-stage renal disease; SD = standard deviation.

Table E.3 presents access to care outcomes for beneficiaries with ESRD. Emergency dialysis treatments were rarer among model than comparison group beneficiaries, although baseline utilization was low for both groups. On other measures, the two groups were similar.

Table E.3. Baseline measures of access to care per beneficiary per quarter, beneficiaries with ESRD

	Model mean (SD)	Comparison mean (SD)	Difference	Percentage difference (%)
Probability of any dialysis use (%)	54.3 (49.8)	55.1 (49.7)	-0.8***	-1.5
Number of dialysis treatments	21.4 (22.9)	21.6 (22.9)	-0.2*	-0.8
Number of days between dialysis treatments	2.4 (1.3)	2.5 (1.4)	-0.1***	-2.1
Probability of emergency dialysis (%)	2.3 (15.1)	3.2 (17.6)	-0.9***	-38.2

	Model mean (SD)	Comparison mean (SD)	Difference	Percentage difference (%)
Number of emergency dialysis treatments	0.0 (0.5)	0.1 (0.4)	-0.0***	-25.0
Probability of hospitalization for ESRD-related conditions (%)	1.9 (13.6)	1.8 (13.3)	0.1***	5.3
Number of hospitalizations for ESRD-related conditions	0.0 (0.2)	0.0 (0.2)	0.0***	0.0
Number of observations	1,491,021	3,063,906		

Note: This table presents baseline weighted means (and standard deviations) of quarterly beneficiary quality of care and access to treatment outcomes for beneficiaries with ESRD. Baseline period for Year 1 states is 2012–2014. Baseline period for Year 2 states is 2012–2015. Comparison group individuals are weighted to resemble model state individuals on baseline demographic and health characteristics. The model states were Delaware, Maryland, New Jersey, North Carolina, Pennsylvania, South Carolina, Virginia, West Virginia, and the District of Columbia. The comparison states were Alabama, Florida, Georgia, Indiana, Kentucky, Louisiana, Massachusetts, Montana, Nebraska, Ohio, Tennessee, Texas, and Washington.

*p < 0.05, **p < 0.01, ***p < 0.001.

ESRD = end-stage renal disease; SD = standard deviation.

2. Full results from multivariate regressions

In this section, we present more detailed regression results for the outcomes than we presented in the main report.

Utilization and expenditures

Tables E.4 and E.5 present detailed results from our analyses of utilization and expenditures for the full sample and by chronic condition (Objective 1).

Table E.4. Impact of RSNAT-PA on utilization and expenditures per beneficiary per quarter, by chronic condition

	Probability of RSNAT ambulance service utilization (percentage points) (I)	Number of RSNAT ambulance trips (II)	Probability of any Medicare ambulance utilization (percentage points) (III)	Total number of Medicare ambulance trips (IV)	RSNAT service expenditures ($) (V)	All Medicare ambulance expenditures ($) (VI)	Total Medicare FFS expenditures ($) (VII)
ESRD and/or pressure ulcers (11,566,321 beneficiary-quarters)							
Average marginal effect (AME)	-2.6***	-1.6***	-1.8***	-1.6***	-265***	-333***	-316***
(standard error)	(0.05)	(0.03)	(0.07)	(0.03)	(5)	(7)	(35)
Baseline mean	4.1	2.2	24.8	2.7	370	642	15,904
AME as percentage of baseline	-63.2	-71.8	-7.2	-58.2	-71.7	-51.9	-2.0
ESRD only (8,109,475 beneficiary-quarters)							
Average marginal effect	-3.20***	-1.92***	-1.97***	-1.94***	-326***	-402***	-443***
(standard error)	(0.08)	(0.04)	(0.08)	(0.04)	(6)	(9)	(37)
Baseline mean	4.4	2.5	18.3	2.8	423	640	13,705
AME as percentage of baseline	-72.8	-77.3	-10.8	-68.8	-76.9	-62.8	-3.2
Pressure ulcers only (3,118,359 beneficiary-quarters)							
Average marginal effect	-0.3***	-0.1***	-1.1***	-0.2***	-13***	-41***	245***
(standard error)	(0.03)	(0.01)	(0.13)	(0.01)	(1)	(3)	(70)
Baseline mean	1.1	0.2	41.6	1.4	28	399	19,891
AME as percentage of baseline	-29.9	-47.0	-2.7	-11.7	-48.6	-10.2	1.2
ESRD and pressure ulcers (338,487 beneficiary-quarters)							
Average marginal effect	-11.2***	-6.1***	-4.3***	-6.3***	-1039***	-1325***	-1060**
(standard error)	(0.53)	(0.32)	(0.46)	(0.32)	(53)	(72)	(355)
Baseline mean	23.8	11.9	62.2	13.7	2,007	3,092	47,219
AME as percentage of baseline	-47.0	-51.5	-6.9	-45.5	-51.8	-42.9	-2.2

Note: The table presents average marginal effects and (standard errors) from weighted logistic (columns I and III) and ordinary least squares columns (II, IV, V, VI, VII) regression analyses. Control variables include age, age squared, sex, race, rural residence, dual eligibility for Medicare and Medicaid, hospital bed claim, an indicator for residing in a county with a moratorium on new Medicare suppliers, Hierarchical Condition Category (HCC) score, and length of time since the county moratorium took effect. Standard errors are adjusted to account for correlation between observations on the same individual. Coefficients from logistic regressions have been transformed into average marginal effects. The model states were Delaware, Maryland, New Jersey, North Carolina, Pennsylvania, South Carolina, Virginia, West Virginia, and the District of Columbia. The comparison states were Alabama, Florida, Georgia, Indiana, Kentucky, Louisiana, Massachusetts, Montana, Nebraska, Ohio, Tennessee, Texas, and Washington.

*p < 0.05, **p < 0.01, ***p < 0.001.

AME = average marginal effect; ESRD = end-stage renal disease; FFS = fee-for-service; RSNAT-PA = Prior Authorization Model for Repetitive, Scheduled, Non-emergent Ambulance Transport.

Table E.5. Impact of RSNAT-PA on Medicare FFS expenditures ($) per beneficiary per quarter, by chronic condition

	Outpatient[a] expenditures (I)	Professional services[b] expenditures (II)	Inpatient expenditures (III)	SNF expenditures (IV)	Home health expenditures (V)	Total expenditures (VI)
ESRD and/or pressure ulcers (11,566,321 beneficiary-quarters)						
Average marginal effect (AME)	0	-341***	8	12†	22***	-316***
Baseline mean	3,998	2,785	6,519	1,296	440	15,904
AME as percentage of baseline	0.0	-12.3	0.1	0.9	5.0	-2.0
ESRD only (8,109,475 beneficiary-quarters)						
Average marginal effect	-14	-412***	-30	-6	36***	-443***
Baseline mean	4,830	2,640	4,806	552	256	13,705
AME as percentage of baseline	-0.3	-15.6	-0.6	-1.1	14.1	-3.2
Pressure ulcers only (3,118,359 beneficiary-quarters)						
Average marginal effect	41***	-8	163**	42*	7	245***
Baseline mean	1,045	2,675	10,234	3,414	978	19,891
AME as percentage of baseline	4.0	-0.3	1.6	1.2	0.8	1.2
ESRD and pressure ulcers (338,487 beneficiary-quarters)						
Average marginal effect	119†	-1,280***	275	-88	13	-1,060**
Baseline mean	6,335	8,529	24,887	4,353	1,055	47,219
AME as percentage of baseline	1.9	-15.0	1.10	-2.0	1.2	-2.2

Note: The table presents average marginal effects from ordinary least squares regression analyses. Control variables include age, age squared, sex, race, rural residence, median income, missing indicator for median income, dual eligibility for Medicare and Medicaid, hospital bed, cohort, an indicator for residing in a county with a moratorium on new Medicare suppliers, Hierarchical Condition Category (HCC) score, and length of time since the county moratorium went into effect. Standard errors are adjusted to account for correlation between observations on the same individual. The model states were Delaware, Maryland, New Jersey, North Carolina, Pennsylvania, South Carolina, Virginia, West Virginia, and the District of Columbia. The comparison states were Alabama, Florida, Georgia, Indiana, Kentucky, Louisiana, Massachusetts, Montana, Nebraska, Ohio, Tennessee, Texas, and Washington.

[a] Outpatient providers include hospital outpatient departments, rural health clinics, renal dialysis facilities, outpatient rehabilitation facilities, comprehensive outpatient rehabilitation facilities, federally qualified health centers, and community mental health centers.

[b] Professional services providers include physicians, physician assistants, clinical social workers, and nurse practitioners, as well as some organizational providers such as independent clinical laboratories, ambulance providers, free-standing ambulatory surgical centers, and free-standing radiology centers.

†p < 0.10, *p < 0.05, **p < 0.01, ***p < 0.001

AME = average marginal effect; ESRD = end-stage renal disease; FFS = fee-for-service; RSNAT-PA = Prior Authorization Model for Repetitive, Scheduled, Non-emergent Ambulance Transport. SNF = skilled nursing facility.

Tables E.6 and E.7 present detailed results from our analyses of utilization and expenditures by model cohort.

Table E.6. Impact of RSNAT-PA on utilization and expenditures per beneficiary per quarter, by cohort

	Probability of RSNAT ambulance service utilization (percentage points) (I)	Number of RSNAT ambulance trips (II)	Probability of any Medicare ambulance utilization (percentage points) (III)	Total number of Medicare ambulance trips (IV)	RSNAT service expenditures ($) (V)	All Medicare ambulance expenditures ($) (VI)	Total Medicare FFS expenditures ($) (VII)
Year 1 cohort (9,480,975 beneficiary-quarters)							
Average marginal effect (AME)	-3.9***	-2.5***	-3.2***	-2.6***	-440***	-544***	-211***
(standard error)	(0.1)	(0.1)	(0.1)	(0.1)	(9)	(11)	(49)
Baseline mean	6.8	3.7	26.8	4.3	653	971	16,864
AME as percentage of baseline	-57.2	-66.9	-11.9	-59.1	-67.4	-56.0	-1.3
Year 2 cohort (9,628,136 beneficiary-quarters)							
Average marginal effect	-1.3***	-0.8***	-0.8***	-0.8***	-121***	-165***	-388***
(standard error)	(0.0)	(0.0)	(0.1)	(0.0)	(5)	(7)	(44)
Baseline mean	2.3	1.1	23.4	1.6	175	416	15,245
AME as percentage of baseline	-56.7	-72.6	-3.3	-49.5	-69.1	-39.7	-2.5

Note: The table presents average marginal effects and (standard errors) from weighted logistic (columns I and III) and ordinary least squares (columns II, IV, V, VI, VII) regression analyses using dates of service January 2012 through December 2018. Control variables include age, age squared, sex, race, rural residence, dual eligibility for Medicare and Medicaid, hospital bed, an indicator for residing in a county with a moratorium on new Medicare suppliers, Hierarchical Condition Category score, and length of time since the county moratorium went into effect. Standard errors are adjusted to account for correlation between observations on the same individual. Coefficients from logistic regressions have been transformed into average marginal effects. The Year 1 model states included New Jersey, Pennsylvania, and South Carolina. The Year 2 model states were Delaware, Maryland, North Carolina, Virginia, West Virginia, and the District of Columbia. The comparison states were Alabama, Florida, Georgia, Indiana, Kentucky, Louisiana, Massachusetts, Montana, Nebraska, Ohio, Tennessee, Texas, and Washington.
*p < 0.05, **p < 0.01, ***p < 0.001.

AME = average marginal effect; FFS = fee-for-service; RSNAT-PA = Prior Authorization Model for Repetitive, Scheduled, Non-emergent Ambulance Transport.

Table E.7. Impact of RSNAT-PA on utilization and expenditures per beneficiary per quarter, by cohort

	Outpatient[a] expenditures (I)	Professional services[b] expenditures (II)	Inpatient expenditures (III)	SNF expenditures (IV)	Home health expenditures (V)	Total expenditures (VI)
Year 1 cohort (9,480,975 beneficiary-quarters)						
Average marginal effect (AME)	90***	-514***	193***	32**	24***	-211***
(standard error)	(15)	(14)	(35)	(10)	(4)	(49)
Baseline mean	3,812	3,296	6,781	1,534	459	16,864
AME as percentage of baseline	2.4	-15.6	2.9	2.1	5.3	-1.3
Year 2 cohort (9.628,136 beneficiary-quarters)						
Average marginal effect	-43**	-180***	-134***	-14	23***	-388***
(standard error)	(15)	(11)	(33)	(8)	(3)	(44)
Baseline mean	4,133	2,433	6,338	1,133	427	15,245
AME as percentage of baseline	-1.1	-7.4	-2.1	-1.2	5.3	-2.5

Note: The table presents average marginal effects from ordinary least squares regression analyses using dates of service January 2012 through December 2018. Control variables include age, age squared, sex, race, rural residence, dual eligibility for Medicare and Medicaid, hospital bed, an indicator for residing in a county with a moratorium on new Medicare suppliers, Hierarchical Condition Category score, and length of time since the county moratorium went into effect. Standard errors are adjusted to account for correlation between observations on the same individual. Coefficients from logistic regressions have been transformed into average marginal effects. The Year 1 model states were New Jersey, Pennsylvania, and South Carolina. The Year 2 model states were Delaware, Maryland, North Carolina, Virginia, West Virginia, and the District of Columbia. The comparison states were Alabama, Florida, Georgia, Indiana, Kentucky, Louisiana, Massachusetts, Montana, Nebraska, Ohio, Tennessee, Texas, and Washington.

[a] Outpatient providers include hospital outpatient departments, rural health clinics, renal dialysis facilities, outpatient rehabilitation facilities, comprehensive outpatient rehabilitation facilities, federally qualified health centers, and community mental health centers.

[b] Professional services providers include physicians, physician assistants, clinical social workers, and nurse practitioners, as well as some organizational providers such as independent clinical laboratories, ambulance providers, and free-standing ambulatory surgical and radiology centers.

*p < 0.05, **p < 0.01, ***p < 0.001.

AME = average marginal effect; RSNAT-PA = Prior Authorization Model for Repetitive, Scheduled, Non-emergent Ambulance Transport; SNF = skilled nursing facility.

Access to care and quality of care

Tables E.8 and E.9 provide greater detail on our results related to quality of care. Table E.10 provides more detail on our results related to access to care for beneficiaries with ESRD.

Table E.8. Impact of RSNAT-PA on quality of care per beneficiary per quarter, full sample and by chronic condition

	Probability of emergency department utilization (percentage points) (I)	Number of emergency department visits (II)	Probability of emergency ambulance utilization (percentage points) (III)	Number of emergency ambulance trips (IV)	Probability of hospital admission (percentage points) (V)	Probability of death (percentage points) (VI)
ESRD and/or pressure ulcers (11,566,321 beneficiary-quarters)						
Average marginal effect (AME)	-0.77***	-0.02***	0.08	-0.00*	-0.65***	-0.04
(standard error)	(0.07)	(0.00)	(0.06)	(0.00)	(0.05)	(0.03)
Baseline mean	35.90	0.65	18.78	0.31	27.03	5.44
AME as percentage of baseline	-2.15	-3.65	0.43	-1.25	-2.42	-0.66
ESRD only (8,109,475 beneficiary-quarters)						
Average marginal effect	-0.71***	-0.02***	0.05	-0.00	-0.62***	-0.05
(standard error)	(0.08)	(0.00)	(0.07)	(0.00)	(0.06)	(0.03)
Baseline mean	31.87	0.58	14.04	0.23	22.11	3.28
AME as percentage of baseline	-2.24	-3.50	0.35	-1.14	-2.80	-1.48
Pressure ulcers only (3,118,359 beneficiary-quarters)						
Average marginal effect	-0.97***	-0.03***	0.17	-0.00	-0.71***	0.01
(standard error)	(0.13)	(0.00)	(0.12)	(0.00)	(0.11)	(0.07)
Baseline mean	45.73	0.80	31.29	0.51	39.34	11.30
AME as percentage of baseline	-2.12	-4.12	0.56	-0.43	-1.81	0.09
ESRD and pressure ulcers (338,487 beneficiary-quarters)						
Average marginal effect	-1.62***	-0.07***	-0.44	-0.02	-1.46***	-0.54*
(standard error)	(0.40)	(0.02)	(0.45)	(0.01)	(0.39)	(0.25)
Baseline mean	65.12	1.41	45.08	0.86	59.26	14.36
AME as percentage of baseline	-2.48	-4.75	-0.99	-1.76	-2.46	-3.75

Note: This table presents average marginal effects and (standard errors) from weighted logistic (columns I, III, V, and VI) and ordinary least squares (columns II and IV) regression analyses using 11,566,321 beneficiary-quarters from dates of service from January 2012 through December 2018 for beneficiaries with ESRD only (8,109,475 beneficiary-quarters), pressure ulcers only (3,118,359 beneficiary-quarters), or both (338,487 beneficiary quarters). Control variables include age, age squared, sex, race, rural residence, dual eligibility for Medicare and Medicaid, hospital bed, an indicator for residing in a county with a moratorium on new Medicare suppliers, Hierarchical Condition Category (HCC) score, and length of time since the county moratorium went into effect. Standard errors are adjusted to account for correlation between observations on the same individual. Coefficients from logistic regressions have been transformed into average marginal effects. The model states were Delaware, Maryland, New Jersey, North Carolina, Pennsylvania, South Carolina, Virginia, West Virginia, and the District of Columbia. The comparison states were Alabama, Florida, Georgia, Indiana, Kentucky, Louisiana, Massachusetts, Montana, Nebraska, Ohio, Tennessee, Texas, and Washington.

*p < 0.05, **p < 0.01, ***p < 0.001.

AME = average marginal effect; ESRD = end-stage renal disease; RSNAT-PA = Prior Authorization Model for Repetitive, Scheduled, Non-emergent Ambulance Transport.

Table E.9. Impact of RSNAT-PA on quality of care per beneficiary per quarter, by cohort

	Probability of emergency department utilization (percentage points) (I)	Number of emergency department visits (II)	Probability of emergency ambulance utilization (percentage points) (III)	Number of emergency ambulance trips (IV)	Probability of hospital admission (percentage points) (V)	Probability of death (percentage points) (VI)
Year 1 cohort (9,480,975 beneficiary-quarters)						
Average marginal effect (AME)	-0.34***	-0.02***	-0.41***	-0.01***	-0.27***	-0.04
(standard error)	(0.09)	(0.00)	(0.08)	(0.00)	(0.08)	(0.04)
Baseline mean	35.26	0.61	19.42	0.31	26.93	5.77
AME as percentage of baseline	-0.96	-2.69	-2.09	-4.47	-1.02	-0.76
Year 2 cohort (9,628,136 beneficiary-quarters)						
Average marginal effect	-1.18***	-0.03***	0.50***	0.00*	-1.44***	-0.02
(standard error)	(0.09)	(0.00)	(0.08)	(0.00)	(0.07)	(0.04)
Baseline mean	36.34	0.67	18.34	0.30	25.84	5.21
AME as percentage of baseline	-3.26	-4.82	2.70	1.40	-5.56	-0.31

Note: Table presents average marginal effects and (standard errors) from weighted logistic (columns I, III, V, and VI) and ordinary least squares (columns II and IV) regression analyses using 9,480,975 beneficiary-quarters (Year 1 model states) and 9,628,136 beneficiary-quarters (Year 2 states) from dates of service January 2012 through December 2018. Control variables include age, age squared, sex, race, rural residence, dual eligibility for Medicare and Medicaid, hospital bed, an indicator for residing in a county with a moratorium on new Medicare suppliers, Hierarchical Condition Category score, and length of time since the county moratorium went into effect. Standard errors are adjusted to account for correlation between observations on the same individual. Coefficients from logistic regressions have been transformed into average marginal effects. The Year 1 cohort of states included New Jersey, Pennsylvania, and South Carolina. The Year 2 cohort of states were Delaware, Maryland, North Carolina, Virginia, West Virginia, and the District of Columbia. The comparison states were Alabama, Florida, Georgia, Indiana, Kentucky, Louisiana, Massachusetts, Montana, Nebraska, Ohio, Tennessee, Texas, and Washington.

*p < 0.05, **p < 0.01, ***p < 0.001.

AME = average marginal effect; RSNAT-PA = Prior Authorization Model for Repetitive, Scheduled, Non-emergent Ambulance Transport.

Table E.10. Impact of RSNAT-PA on access to care per beneficiary per quarter, beneficiaries with ESRD

	Probability of dialysis use (percentage points) (I)	Number of days of dialysis use (II)	Average number of days between dialysis use (III)	Probability of emergency dialysis (percentage points) (IV)	Number of emergency dialysis treatments (V)	Probability of hospitalization due to ESRD complications (percentage points) (VI)	Number of hospitalizations due to ESRD complications (VII)
All ESRD beneficiaries (8,447,962 beneficiary-quarters)							
Average marginal effect (AME)	-0.66***	-0.21**	-0.01**	0.44***	0.01***	-0.26***	-0.00***
(standard error)	(0.14)	(0.07)	(0.00)	(0.03)	(0.00)	(0.02)	(0.00)
Baseline mean	54.29	21.37	2.42	2.34	0.04	1.89	0.02
AME as percentage of baseline	-1.21	-0.96	-0.41	18.73	13.88	-13.92	-15.67
ESRD only (8,109,475 beneficiary-quarters)							
Average marginal effect	-0.73***	-0.25***	-0.01*	0.41***	0.01***	-0.25***	-0.00***
(standard error)	(0.14)	(0.07)	(0.00)	(0.03)	(0.00)	(0.02)	(0.00)
Baseline mean	53.62	21.24	2.41	2.29	0.04	1.81	0.02
AME as percentage of baseline	-1.36	-1.17	-0.33	18.07	13.47	-13.85	-15.44
Year 1 states (6,891,712 beneficiary-quarters)							
Average marginal effect	0.26	0.12	-0.00	0.40***	0.00	-0.30***	-0.00***
(standard error)	(0.20)	(0.10)	(0.00)	(0.04)	(0.00)	(0.03)	(0.00)
Baseline mean	53.09	20.75	2.42	1.90	0.03	1.83	0.02
AME as percentage of baseline	0.49	0.56	-0.17	20.92	7.57	-16.60	-17.61
Year 2 states (7,096,367 beneficiary-quarters)							
Average marginal effect	-1.34***	-0.43***	-0.02***	0.55***	0.01***	-0.30***	-0.00***
(standard error)	(0.18)	(0.10)	(0.00)	(0.04)	(0.00)	(0.03)	(0.00)
Baseline mean	55.08	21.78	2.42	2.63	0.04	1.93	0.02
AME as percentage of baseline	-2.43	-1.97	-0.71	20.92	18.84	-15.67	-17.67

Note: The table presents average marginal effects and (standard errors) from weighted logistic (columns I, IV, and VI) and ordinary least squares (columns II, III, VII) regression analyses from dates of service from January 2012 through December 2018 for beneficiaries with ESRD. Control variables include age, age squared, sex, race, rural residence, dual eligibility for Medicare and Medicaid, hospital bed, an indicator for residing in a county with a moratorium on new Medicare suppliers, Hierarchical Condition Category score, and length of time since the county moratorium went into effect. Standard errors are adjusted to account for correlation between observations on the same individual. Coefficients from logistic regressions have been transformed into average marginal effects. The model states were Delaware, Maryland, New Jersey, North Carolina, Pennsylvania, South Carolina, Virginia, West Virginia, and the District of Columbia. The comparison states were Alabama, Florida, Georgia, Indiana, Kentucky, Louisiana, Massachusetts, Montana, Nebraska, Ohio, Tennessee, Texas, and Washington.

*p < 0.05, **p < 0.01, ***p < 0.001.

AME = average marginal effect; ESRD = end-stage renal disease; RSNAT-PA = Prior Authorization Model for Repetitive, Scheduled, Non-emergent Ambulance Transport.

3. Subgroup impacts from multivariate regression

We stratified our multivariate analysis by rural residence and dual eligibility for Medicare and Medicaid and we found that estimated impacts were larger in both magnitude and percentage terms for urban than for rural residents, with the exception of impacts on total Medicare expenditures, where impacts were larger for rural residents. Estimated impacts for dually eligible beneficiaries were larger in magnitude than for non-dually eligible beneficiaries, but similar in percentage terms (Table E.11).

Table E.11. Impact of RSNAT-PA on utilization and expenditures per beneficiary per quarter, by rural residence and dual eligibility

	Probability of RSNAT ambulance service utilization (percentage points) (I)	Number of RSNAT ambulance trips (II)	Probability of any Medicare ambulance utilization (percentage points) (III)	Total number of Medicare ambulance trips (IV)	RSNAT service expenditures ($) (V)	All Medicare ambulance expenditures ($) (VI)	Total Medicare FFS expenditures ($) (VII)
Rural status							
Rural							
Average marginal effect (AME)	-1.41***	-0.88***	-1.62***	-0.92***	-138***	-223***	-463***
(standard error)	(0.10)	(0.07)	(0.15)	(0.07)	(10)	(17)	(69)
Baseline mean	4.04	2.01	25.91	2.68	320	723	14,708
AME as percentage of baseline	-34.83	-43.79	-6.27	-34.31	-43.02	-31	-3.15
Not rural							
Average marginal effect	-3.52***	-1.92***	-2.18***	-1.97***	-333***	-401***	-322***
(standard error)	(0.07)	(0.04)	(0.08)	(0.04)	(6)	(7)	(39)
Baseline mean	4.13	2.21	24.51	2.76	383	620	16,235
AME as percentage of baseline	-85.36	-87.18	-8.90	-71.48	-86.74	-65	-1.98

	Probability of RSNAT ambulance service utilization (percentage points) (I)	Number of RSNAT ambulance trips (II)	Probability of any Medicare ambulance utilization (percentage points) (III)	Total number of Medicare ambulance trips (IV)	RSNAT service expenditures ($) (V)	All Medicare ambulance expenditures ($) (VI)	Total Medicare FFS expenditures ($) (VII)
Dual eligibility							
Dually eligible							
Average marginal effect	-3.59***	-2.26***	-2.42***	-2.34***	-382***	-489***	-419***
(standard error)	(0.10)	(0.06)	(0.12)	(0.06)	(10)	(13)	(58)
Baseline mean	6.09	3.32	32.62	4.17	562	970	18,336
AME as percentage of baseline	-58.91	-68.19	-7.41	-56.01	-68.01	-50.46	-2.29
Not dually eligible							
Average marginal effect	-2.03***	-1.13***	-1.46***	-1.15***	-195***	-240***	-290***
(standard error)	(0.06)	(0.03)	(0.08)	(0.03)	(5)	(7)	(43)
Baseline mean	2.93	1.48	20.18	1.90	256	449	14,467
AME as percentage of baseline	-69.29	-75.91	-7.22	-60.53	-76.08	-53.51	-2.01

Note: Table presents average marginal effects (and standard errors) from weighted logistic (columns I and III) and ordinary least squares (columns II, IV, V, VI, VII) regression analyses using 2,479,711 beneficiary quarters (rural); 9,086,610 beneficiary quarters (not rural); 4,669,287 beneficiary quarters (dually eligible); and 6,897,034 beneficiary quarters (not dually eligible) from dates of service January 2012 through December 2018 for beneficiaries with ESRD and/or pressure ulcers. Control variables included age, age squared, sex, race, rural residence, dual eligibility for Medicare and Medicaid, hospital bed, an indicator for residing in a county with a moratorium on new Medicare suppliers, HCC score, and length of time since the county moratorium went into effect. Errors are clustered at the individual level. Coefficients from logistic regressions have been transformed into average marginal effects. The model states included: Delaware; Maryland; New Jersey; North Carolina; Pennsylvania; South Carolina; Virginia; Washington, DC; and West Virginia. The comparison states included: Alabama, Florida, Georgia, Indiana, Kentucky, Louisiana, Massachusetts, Montana, Nebraska, Ohio, Tennessee, Texas, and Washington.

*p < 0.05, **p < 0.01, ***p < 0.001.

AME = average marginal effect; ESRD = end-stage renal disease; FFS = fee-for-service; RSNAT-PA = Prior Authorization Model for Repetitive, Scheduled, Non-emergent Ambulance Transport.

We also examined Objective 1 outcomes for the subgroup of beneficiaries who had a claim for a hospital bed. This subgroup had sizeable decreases in utilization and expenditures, larger in magnitude than for the full set of beneficiaries but similar in percentage terms. Table E.12 contains these results.

Table E.12. Impact of RSNAT-PA on utilization and expenditures per beneficiary per quarter, beneficiaries with a hospital bed claim

	Probability of RSNAT ambulance service utilization (percentage points) (I)	Number of RSNAT ambulance trips (II)	Probability of any Medicare ambulance utilization (percentage points) (III)	Total number of Medicare ambulance trips (IV)	RSNAT service expenditures ($) (V)	All Medicare ambulance expenditures ($) (VI)	Total Medicare FFS expenditures ($) (VII)
Average marginal effect (AME)	-5.47***	-3.47***	-3.32***	-3.57***	-584***	-746***	-478***
(standard error)	(0.25)	(0.17)	(0.29)	(0.17)	(27)	(37)	(142)
Baseline mean	11.75	6.21	51.93	7.69	1,038	1,777	26,459
AME as percentage of baseline	-46.58	-55.92	-6.40	-46.43	-56.26	-41.97	-1.81

Note: Table presents average marginal effects (and standard errors) from weighted logistic (columns I and III) and ordinary least squares (columns II, IV, V, VI, VII) regression analyses using 928,464 beneficiary quarters from dates of service from January 2012 through December 2018 for beneficiaries with ESRD and/or pressure ulcers who have a claim for a hospital bed. Control variables included age, age squared, sex, race, rural residence, dual eligibility for Medicare and Medicaid, an indicator for residing in a county with a moratorium on new Medicare suppliers, HCC score, and length of time since the county moratorium took effect. Errors are clustered at the individual level. Coefficients from logistic regressions have been transformed into average marginal effects. The model states included: Delaware; Maryland; New Jersey; North Carolina; Pennsylvania; South Carolina; Virginia; Washington, DC; and West Virginia. The comparison states included: Alabama, Florida, Georgia, Indiana, Kentucky, Louisiana, Massachusetts, Montana, Nebraska, Ohio, Tennessee, Texas, and Washington.

*p < 0.05, **p < 0.01, ***p < 0.001.

AME = average marginal effect; ESRD = end-stage renal disease; FFS = fee-for-service; RSNAT-PA = Prior Authorization Model for Repetitive, Scheduled, Non-emergent Ambulance Transport.

C. Quality of care and access to care

Table E.13 contains results of the multivariate analyses of quality of care outcomes stratified by rural residence and dual eligibility for Medicare and Medicaid. Rural and urban beneficiaries experienced similar impacts in both magnitude and percentage terms, with two exceptions. For unplanned hospital admission, urban residents saw larger impacts than rural residents, although the probability of admission decreased for both groups. Results were similar for dual and non-dual beneficiaries.

We find no discernible impacts on emergency department or emergency ambulance utilization for the subgroup of beneficiaries with a hospital bed claim (Table E.14). We find a small decrease in the probability of unplanned hospital admission.

Table E.13. Impact of RSNAT-PA on quality of care per beneficiary per quarter, by rural residence and dual eligibility

	Probability of emergency department utilization (percentage points) (I)	Number of emergency department visits (II)	Probability of emergency ambulance utilization (percentage points) (III)	Number of emergency ambulance trips (IV)	Probability of unplanned admission (percentage points) (V)	Probability of death (percentage points) (VI)
Rural status						
Rural						
Average marginal effect (AME)	-0.01	-0.00*	-0.75***	-0.01***	-0.56***	-0.06
(standard error)	(0.09)	(0.00)	(0.14)	(0.00)	(0.12)	(0.06)
Baseline mean	6.05	0.11	20.46	0.35	26.14	5.32
AME as percentage of baseline	-0.22	-3.54	-3.68	-3.99	-2.16	-1.14
Not rural						
Average marginal effect	-0.01	-0.00***	0.16*	-0.00*	-0.96***	-0.04
(standard error)	(0.04)	(0.00)	(0.07)	(0.00)	(0.06)	(0.03)
Baseline mean	6.22	0.10	18.31	0.29	26.32	5.47
AME as percentage of baseline	-0.20	-3.87	0.86	-1.37	-3.66	-0.64

	Probability of emergency department utilization (percentage points) (I)	Number of emergency department visits (II)	Probability of emergency ambulance utilization (percentage points) (III)	Number of emergency ambulance trips (IV)	Probability of unplanned admission (percentage points) (V)	Probability of death (percentage points) (VI)
Dual eligibility						
Dually eligible						
Average marginal effect (AME)	0.01	-0.00**	0.06	-0.01*	-0.86***	-0.07
(standard error)	(0.06)	(0.00)	(0.11)	(0.00)	(0.09)	(0.05)
Baseline mean	6.32	0.11	24.60	0.43	30.20	5.79
AME as percentage of baseline	0.23	-2.98	0.23	-1.57	-2.85	-1.16
Not dually eligible						
Average marginal effect	0.03	-0.00**	0.06	-0.00*	-0.89***	0.01
(standard error)	(0.05)	(0.00)	(0.07)	(0.00)	(0.07)	(0.04)
Baseline mean	6.11	0.10	15.33	0.23	23.96	5.23
AME as percentage of baseline	0.55	-2.45	0.42	-1.31	-3.71	0.21

Note: Table presents average marginal effects (and standard errors) from weighted logistic (columns I, III, V, and VI) and ordinary least squares (columns II and IV) regression analyses using 2,479,707 beneficiary quarters (rural); 9,086,610 beneficiary quarters (not rural); 4,669,287 beneficiary quarters (dually eligible); and 6,897,034 beneficiary quarters (not dually eligible) from dates of service January 2012 through December 2018 for beneficiaries with ESRD and/or pressure ulcers. Control variables included age, age squared, sex, race, rural residence, dual eligibility for Medicare and Medicaid, hospital bed, an indicator for residing in a county with a moratorium on new Medicare suppliers, HCC score, and length of time since the county moratorium went into effect. Errors are clustered at the individual level. Coefficients from logistic regressions have been transformed into average marginal effects. The model states included: Delaware; Maryland; New Jersey; North Carolina; Pennsylvania; South Carolina; Virginia; Washington, DC; and West Virginia. The comparison states included: Alabama, Florida, Georgia, Indiana, Kentucky, Louisiana, Massachusetts, Montana, Nebraska, Ohio, Tennessee, Texas, and Washington.

*p < 0.05, **p < 0.01, ***p < 0.001.

AME = average marginal effect; ESRD = end-stage renal disease; RSNAT-PA = Prior Authorization Model for Repetitive, Scheduled, Non-emergent Ambulance Transport.

Table E.14. Impact of RSNAT-PA on quality of care per beneficiary per quarter, beneficiaries with a hospital bed claim

	Probability of emergency department utilization (percentage points) (I)	Number of emergency department visits (II)	Probability of emergency ambulance utilization (percentage points) (III)	Number of emergency ambulance trips (IV)	Probability of unplanned admission (percentage points) (V)	Probability of death (percentage points) (VI)
Average marginal effect (AME)	-0.09	-0.01*	0.11	-0.01	-0.94***	-0.01
(standard error)	(0.16)	(0.00)	(0.26)	(0.01)	(0.22)	(0.13)
Baseline mean	13.05	0.24	37.47	0.67	44.70	11.71
AME as percentage of baseline	-0.66	-2.87	0.30	-0.99	-2.09	-0.13

Note: Table presents average marginal effects (and standard errors) from weighted logistic (columns I, III, V, and VI) and ordinary least squares (columns II and IV) regression analyses using 928,464 beneficiary quarters from dates of service from January 2012 through June 2017 for beneficiaries with ESRD and/or pressure ulcers who have a claim for a hospital bed. Control variables included age, age squared, sex, race, rural residence, dual eligibility for Medicare and Medicaid, an indicator for residing in a county with a moratorium on new Medicare suppliers, Hierarchical Condition Category score, and length of time since the county moratorium took effect. Errors are clustered at the individual level. Coefficients from logistic regressions have been transformed into average marginal effects. The model states included: Delaware; Maryland; New Jersey; North Carolina; Pennsylvania; South Carolina; Virginia; Washington, DC; and West Virginia. The comparison states included: Alabama, Florida, Georgia, Indiana, Kentucky, Louisiana, Massachusetts, Montana, Nebraska, Ohio, Tennessee, Texas, and Washington.

*p < 0.05, **p < 0.01, ***p < 0.001.

AME = average marginal effect; ESRD = end-stage renal disease; RSNAT-PA = Prior Authorization Model for Repetitive, Scheduled, Non-emergent Ambulance Transport.

D. Suppliers

Table E.15 contains weighted descriptive statistics of beneficiaries in supplier catchment areas for the suppliers in model and comparison states. Although most of the differences are significantly different from zero statistically, the differences are proportionally small in magnitude (less than 10 percent of the comparison group mean) for all characteristics except the percentage of the population that is of non-white race and the percentage that has pressure ulcers.

Table E.15. Characteristics of beneficiaries in supplier catchment areas (weighted)

	Model mean (SD)	Comparison mean (SD)	Difference	Percentage difference (%)
Age of attributed beneficiaries (years)	71.04 (1.34)	70.95 (1.36)	0.09*	0.13
Sex of attributed beneficiaries (% female)	55.67 (1.74)	55.45 (1.64)	0.23***	0.41
Race of attributed beneficiaries (%)				
White	76.85 (18.63)	80.02 (16.80)	-3.18***	-3.97
Black	16.37 (16.06)	14.12 (16.24)	2.25***	15.93
Other	6.16 (6.79)	5.23 (5.61)	0.93***	17.78
Rural attributed beneficiaries (%)	23.09 (29.46)	24.28 (30.20)	-1.20	-4.94
Dually eligible attributed beneficiaries (%)	19.21 (6.31)	20.43 (7.48)	0.22	1.16
Average HCC score of attributed beneficiaries	1.22 (0.09)	1.20 (0.09)	0.02***	1.67
Chronic conditions of attributed beneficiaries (%)				
ESRD	1.57 (0.59)	1.49 (0.63)	0.08***	5.37
Pressure ulcers	2.23 (0.59)	2.01 (0.45)	0.22***	10.95
Number of suppliers	3,093	5,146		

Note: Table presents weighted means (and standard deviations) of supplier characteristics. Comparison state suppliers are weighted to resemble model state suppliers in the baseline demographic and health characteristics of their customer base. The model states included: Delaware; Maryland; New Jersey; North Carolina; Pennsylvania; South Carolina; Virginia; Washington, DC; and West Virginia. The comparison states included: Alabama, Florida, Georgia, Indiana, Kentucky, Louisiana, Massachusetts, Montana, Nebraska, Ohio, Tennessee, Texas, and Washington.

*p < 0.05, **p < 0.01, ***p < 0.001.

ESRD = end-stage renal disease; HCC = Hierarchical Condition Category; SD = standard deviation.

Table E.16 shows the baseline levels of the supplier outcome measures regarding utilization and payments. On average, suppliers in model states served fewer beneficiaries but made more trips per beneficiary served, as well as more trips that met the RSNAT definition. They received more in RSNAT payments but similar amounts in payments for all ambulance services and total payments.

Table E.16. Baseline supplier quarterly utilization and payments

	Model mean (SD)	Comparison mean (SD)	Difference	Percentage difference (%)
Number of beneficiaries served (any Medicare ambulance)	186.33 (476.56)	205.82 (565.99)	-19.50***	-9.47
Number of RSNAT trips	139.33 (470.40)	87.27 (503.60)	52.06***	59.65
Number of Medicare ambulance trips	388.23 (900.52)	367.51 (1,114.03)	20.73*	5.64
Number of Medicare ambulance trips per beneficiary	4.87 (11.64)	2.49 (6.80)	2.38***	95.58
RSNAT payments ($)	23,838.67 (81,326.90)	14,176.91 (81,483.12)	9,661.75***	68.15
Total Medicare ambulance payments ($)	89,293.63 (203,202.25)	89,834.89 (254,057.22)	-541.27	-0.60
Total Medicare FFS payments ($)	113,475.56 (248,254.30)	113,814.65 (311,352.41)	-339.09	-0.30
Number of observations	32,672	59,248		

Note: Table presents weighted means (and standard deviations) of quarterly supplier utilization and payment outcomes. Comparison state suppliers are weighted to resemble model state suppliers in the baseline demographic and health characteristics of their customer base. The model states included: Delaware; Maryland; New Jersey; North Carolina; Pennsylvania; South Carolina; Virginia; Washington, DC; and West Virginia. The comparison states included: Alabama, Florida, Georgia, Indiana, Kentucky, Louisiana, Massachusetts, Montana, Nebraska, Ohio, Tennessee, Texas, and Washington.

*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

FFS = fee-for-service; RSNAT = repetitive, scheduled, non-emergent ambulance transport; SD = standard deviation.

In the main report, we characterized service provision and payment receipt by suppliers based on whether or not they exited the Medicare ambulance market around the time of model implementation. In addition, we compared the pre-model characteristics of beneficiaries attributed to suppliers that left the program in the model states to those of beneficiaries attributed to suppliers that stayed, and also to suppliers that stayed and left in comparison states. We considered the subset of suppliers that were active before the model went into effect and divided them into three groups: stayers, which also billed Medicare for any ambulance service in two years after prior authorization went into effect; triers, which billed Medicare for any ambulance service in the first year after model implementation but not after that; and leavers, which did not bill Medicare for ambulance services at any point after prior authorization was implemented. Our results are presented in Table E.17. Stayers, triers, and leavers were similar in terms of their customer bases in Year 2 and comparison states. In contrast, stayers, triers, and leavers in Year 1 model states served significantly different types of Medicare beneficiaries. Suppliers that exited the market served customer bases that were less white, less rural, and more likely to be dually eligible. Table E.18 provides more detail on our descriptive analyses of stayers, triers, and leavers.

Table E.17. Characteristics of beneficiaries in supplier catchment areas in the year before RSNAT-PA, by supplier group (stayers, triers, and leavers)

	Year 1 model states			Year 2 model states			Comparison states		
	Stayers	Triers	Leavers	Stayers	Triers	Leavers	Stayers	Triers	Leavers
	Weighted mean (standard deviation)			Weighted mean (standard deviation)			Weighted mean (standard deviation)		
Age of attributed beneficiaries (years)	71 (1)	71 (1)	71* (2)	71 (1)	71 (1)	71 (1)	71 (2)	71 (1)	71 (2)
Sex of attributed beneficiaries (% female)	56 (1)	56* (1)	57*,† (1)	55 (1)	55 (2)	55 (2)	55 (2)	55 (1)	55 (2)
Race of attributed beneficiaries (%)									
White	81 (15)	64* (22)	54*,† (26)	79 (12)	74 (18)	75 (21)	80 (14)	73 (19)	74 (25)
Black	11 (11)	23* (18)	31*,† (24)	16 (11)	22 (17)	20 (19)	13 (14)	21* (18)	19 (23)
Other	7 (7)	11* (10)	13* (10)	4 (3)	4 (3)	4 (8)	6 (5)	5 (5)	6 (10)
Rural attributed beneficiaries (%)	17 (26)	9* (21)	6* (19)	35 (27)	30 (30)	45 (42)	24 (28)	21 (24)	35 (39)
Dually eligible attributed beneficiaries (%)	18 (5)	22* (8)	25*,† (11)	17 (4)	18 (5)	19 (6)	19 (5)	19 (5)	21 (9)
Average HCC score of attributed beneficiaries	1.25 (0.07)	1.28* (0.09)	1.31*,† (0.10)	1.21 (0.05)	1.23 (0.07)	1.23 (0.09)	1.24 (0.08)	1.25 (0.07)	1.24 (0.10)
Number of suppliers (percentage)	1,167 (79)	136 (9)	171 (12)	818 (92)	37 (4)	36 (4)	3,446 (92)	148 (4)	162 (4)

Note: The table presents weighted means and (standard deviations) of supplier characteristics from before model implementation. Stayers are suppliers that were active both before and in at least two years after implementation; Triers are suppliers that were active before and in the first year of implementation; Leavers are suppliers that were active before, but not after, implementation. Comparison state suppliers are weighted to resemble model state suppliers in the demographic and health characteristics of their customer base. The model states included Delaware; Maryland; New Jersey; North Carolina; Pennsylvania; South Carolina; Virginia; Washington, DC; and West Virginia. The comparison states included Alabama, Florida, Georgia, Indiana, Kentucky, Louisiana, Massachusetts, Montana, Nebraska, Ohio, Tennessee, Texas, and Washington.
*Statistically significantly different from Stayer value at 0.05 level.
† Statistically significantly different from Trier value at 0.05 level.

HCC = Hierarchical Condition Category; RSNAT-PA = Prior Authorization Model for Repetitive, Scheduled, Non-emergent Ambulance Transport.

Table E.18. Quarterly services provided and payments received by stayer, trier, and leaver suppliers in the year before RSNAT-PA

	Year 1 model states			Year 2 model states			Comparison states		
	Stayers	Triers	Leavers	Stayers	Triers	Leavers	Stayers	Triers	Leavers
	Weighted mean (standard deviation)			Weighted mean (standard deviation)			Weighted mean (standard deviation)		
RSNAT services provided									
Number of beneficiaries served (any Medicare ambulance)	163 (420)	85* (176)	50* (193)	268 (587)	248 (421)	61*,† (157)	207 (526)	121* (297)	51*,† (139)
Number of Medicare ambulance trips	347 (819)	572* (867)	362† (597)	439 (924)	702 (1,141)	218*,† (653)	384 (1,096)	439 (778)	151*,† (327)
Number of Medicare ambulance trips per beneficiary	4 (10)	17* (19)	24*,† (20)	2 (3)	5 (8)	4 (9)	2 (6)	10* (15)	7* (13)
Number of beneficiaries served (RSNAT)	3 (10)	9* (15)	7* (10)	2 (7)	9* (18)	4 (12)	2 (11)	6* (12)	2† (5)
Number of RSNAT trips	133 (476)	471* (798)	301*,† (522)	79 (277)	360* (701)	146 (558)	101 (543)	268* (552)	82† (201)
Number of RSNAT trips per beneficiary	40 (17)	45* (15)	39† (15)	35 (18)	37 (12)	26 (17)	37 (18)	40 (16)	31 (15)
Percentage of ambulance trips that are RSNAT	14 (30)	58* (42)	73*,† (38)	9 (22)	26* (35)	14 (30)	8 (22)	32* (39)	27* (41)
Payments received									
RSNAT payments ($)	21,691 (76,463)	80,716* (140,910)	51,592*,† (91,885)	12,339 (43,719)	56,589* (109,933)	22,880 (87,923)	15,677 (82,790)	40,950* (84,080)	12,285† (29,800)
Percentage of ambulance payments from RSNAT (%)	13 (29)	57* (42)	72*,† (38)	8 (20)	25* (34)	13 (30)	7 (21)	31* (39)	27* (41)

	Year 1 model states			Year 2 model states			Comparison states		
	Stayers	Triers	Leavers	Stayers	Triers	Leavers	Stayers	Triers	Leavers
	Weighted mean (standard deviation)			Weighted mean (standard deviation)			Weighted mean (standard deviation)		
Percentage of all payments from RSNAT (%)	11 (24)	49* (37)	64*,† (35)	5 (14)	19* (28)	11 (26)	6 (16)	23* (29)	20* (32)
Total Medicare FFS payments ($)	96,936 (222,180)	122,662 (178,379)	77,566† (131,740)	142,856 (276,822)	176,080 (273,953)	52,838*,† (137,760)	117,183 (297,076)	114,625 (211,976)	46,585*,† (80,324)
Number of suppliers (percent)	1,167 (79)	136 (9)	171 (12)	818 (92)	37 (4)	36 (4)	3,446 (92)	148 (4)	162 (4)

Note: The table presents weighted means and (standard deviations) of supplier characteristics from before model implementation. Stayers are suppliers that were active both before and in at least two years after implementation; triers are suppliers that were active before and in the first year of implementation; leavers are suppliers that were active before, but not after, implementation. Comparison state suppliers are weighted to resemble model state suppliers in the demographic and health characteristics of their customer base. The model states were Delaware, Maryland, New Jersey, North Carolina, Pennsylvania, South Carolina, Virginia, West Virginia, and the District of Columbia. The comparison states were Alabama, Florida, Georgia, Indiana, Kentucky, Louisiana, Massachusetts, Montana, Nebraska, Ohio, Tennessee, Texas, and Washington.

*Statistically significantly different from stayer value at 0.05 level.

† Statistically significantly different from trier value at 0.05 level.

FFS = fee-for-service; RSNAT-PA = Prior Authorization Model for Repetitive, Scheduled, Non-emergent Ambulance Transport.

Appendix F

Statistical precision of estimated model effects on RSNAT utilization

This page has been left blank for double-sided copying.

In this appendix, we provide an analysis of the available statistical precision to detect effects of the prior authorization model on different outcomes related to RSNAT utilization. Statistical power analysis uses the variance of the impact estimates to demonstrate the precision with which the impacts are measured. We present precision estimates for utilization outcomes among beneficiaries. Some of these estimated effects are quite small and perhaps not substantively important. Outcomes at the beneficiary level include:

- any ambulance utilization
- number of ambulance trips
- any ambulance utilization specifically for RSNAT
- the number of ambulance trips for those purposes
- several measures of adverse outcomes (emergency department visits, unplanned hospital admissions, and death)

The primary factor that determines statistical power is the sample size used to generate estimates. All other factors being equal, larger sample sizes result in more precise estimates and therefore increase statistical power. Other factors that affect statistical power include the observed variance of the outcomes, design effects due to weighting, and design effects due to clustering. Each of these factors has an inverse relationship with statistical precision: as they increase, statistical precision decreases. The design effect due to weighting is the increase in the variance of an estimate due to unequal weighting. For analyses using beneficiaries, we used this weighting to match model states with comparison states and balance beneficiaries in the comparison states with those in model states. For analyses using suppliers, the weights reflect matching at the state and supplier levels within states.

Using a difference-in-differences (DD) analytic approach, statistical power is also affected by the extent to which the baseline observations predict the post-intervention outcomes, as determined by two factors: (1) the correlation of the outcome measure between the pre- and post-intervention samples, which is highly dependent on the sample overlap of these two time periods, and (2) the correlation between the covariates used in the model with the outcome measure. As this correlation increases, the variance in the predicted outcome decreases, thereby increasing statistical precision. This effect is amplified the more the pre- and post-intervention samples overlap. Analyses with strong predictor variables and high proportions of overlap in the samples will have greater statistical precision than those with less of either or both factors.

Table F.1 presents 90 percent half-confidence intervals for the various beneficiary-level utilization outcomes. These intervals are based on the observed variance of the DD impact estimates, which encapsulates all of the factors described above. The 90 percent confidence interval estimates account for a Type I error rate of 10 percent; in other words, if the data were resampled 100 times, 90 of the estimated confidence intervals would cover the true difference in means. The confidence interval for outcome j CI_j is calculated using the following formula:

(Equation 4) $CI_j = \hat{\beta}_j \pm \left(Z_{1-\alpha} * S_j\right),$

where $\hat{\beta}_j$ is the impact estimate for outcome j, $Z_{1-\alpha}$ is the critical value of the normal distribution, with $\alpha = 0.10$ for the 90 percent confidence intervals, and S_j is the standard error for the impact estimate. Separate results are provided for beneficiaries living in rural areas, with hospital bed claims, and with dual eligibility for Medicare and Medicaid. These three subgroups are defined using unrelated characteristics and are not mutually exclusive.

Table F.1. Statistical precision for RSNAT utilization outcomes

Outcome and sample	Beneficiary quarterly observations in model states	Beneficiary quarterly observations in comparison states	90 percent half-confidence interval	Observed effect
Any ambulance utilization	**4,023,531**	**7,542,790**	**0.1127**	**-1.80*** **
Beneficiaries with ESRD	2,791,479	5,317,996	0.1301	-1.97***
Beneficiaries with skin ulcers	1,115,686	2,002,673	0.2133	-1.11***
Beneficiaries with multiple conditions	116,366	222,121	0.7522	-4.31***
Beneficiaries in rural areas	832,442	1,647,269	0.2503	-1.62***
Beneficiaries with hospital bed claims	351,644	576,820	0.4744	-3.32***
Beneficiaries with dual eligibility	1,493,793	3,175,494	0.1989	-2.42***
Number of ambulance trips	**4,023,531**	**7,542,790**	**0.0503**	**-1.60*** **
Beneficiaries with ESRD	2,791,479	5,317,996	0.0651	-1.94***
Beneficiaries with skin ulcers	1,115,686	2,002,673	0.0169	-0.16***
Beneficiaries with multiple conditions	116,366	222,121	0.5342	-6.25***
Beneficiaries in rural areas	832,442	1,647,269	0.1107	-0.92***
Beneficiaries with hospital bed claims	351,644	576,820	0.2765	-3.57***
Beneficiaries with dual eligibility	1,493,793	3,175,494	0.0997	-2.34***
Any ambulance utilization (RSNAT)	**4,023,531**	**7,542,790**	**0.0896**	**-2.60*** **
Beneficiaries with ESRD	2,791,479	5,317,996	0.1250	-3.20***
Beneficiaries with skin ulcers	1,115,686	2,002,673	0.0490	-0.33***
Beneficiaries with multiple conditions	116,366	222,121	0.8754	-11.19***
Beneficiaries in rural areas	832,442	1,647,269	0.1709	-1.41***
Beneficiaries with hospital bed claims	351,644	576,820	0.4066	-5.47***
Beneficiaries with dual eligibility	1,493,793	3,175,494	0.1674	-3.59***

Outcome and sample	Beneficiary quarterly observations in model states	Beneficiary quarterly observations in comparison states	90 percent half-confidence interval	Observed effect
Number of ambulance trips (RSNAT)	**4,023,531**	**7,542,790**	**0.0495**	**-1.55*** **
Beneficiaries with ESRD	2,791,479	5,317,996	0.0643	-1.92***
Beneficiaries with skin ulcers	1,115,686	2,002,673	0.0117	-0.07***
Beneficiaries with multiple conditions	116,366	222,121	0.5265	-6.11***
Beneficiaries in rural areas	832,442	1,647,269	0.1092	-0.88***
Beneficiaries with hospital bed claims	351,644	576,820	0.2734	-3.47***
Beneficiaries with dual eligibility	1,493,793	3,175,494	0.0984	-2.26***
Probability of ED utilization	**4,023,531**	**7,542,790**	**0.1151**	**0.01[a]**
Beneficiaries with ESRD	2,791,479	5,317,996	0.1316	0.02[a]
Beneficiaries with skin ulcers	1,115,686	2,002,673	0.2138	-0.09[a]
Beneficiaries with multiple conditions	116,366	222,121	0.6579	-0.39[a]
Beneficiaries in rural areas	832,442	1,647,269	0.1441	-0.01[a]
Beneficiaries with hospital bed claims	351,644	576,820	0.2601	-0.09[a]
Beneficiaries with dual eligibility	1,493,793	3,175,494	0.1066	0.01[a]
Number of emergency ED	**4,023,531**	**7,542,790**	**0.0000**	**-0.00*** **
Beneficiaries with ESRD	2,791,479	5,317,996	0.0000	-0.00***
Beneficiaries with skin ulcers	1,115,686	2,002,673	0.0000	-0.01***
Beneficiaries with multiple conditions	116,366	222,121	0.0329	-0.02*[a]
Beneficiaries in rural areas	832,442	1,647,269	0.0026	-0.00*[a]
Beneficiaries with hospital bed claims	351,644	576,820	0.0058	-0.07*[a]
Beneficiaries with dual eligibility	1,493,793	3,175,494	0.0020	-0.00**[a]

Outcome and sample	Beneficiary quarterly observations in model states	Beneficiary quarterly observations in comparison states	90 percent half-confidence interval	Observed effect
Probability of emergency ambulance utilization	**4,023,531**	**7,542,790**	**0.0987**	**0.08[a]**
Beneficiaries with ESRD	2,791,479	5,317,996	0.1079	0.05[a]
Beneficiaries with skin ulcers	1,115,686	2,002,673	0.2048	0.17[a]
Beneficiaries with multiple conditions	116,366	222,121	0.7367	-0.44[a]
Beneficiaries in rural areas	832,442	1,647,269	0.2267	-0.75***
Beneficiaries with hospital bed claims	351,644	576,820	0.4208	0.11[a]
Beneficiaries with dual eligibility	1,493,793	3,175,494	0.1752	0.06[a]
Number of emergency ambulance trips	**4,023,531**	**7,542,790**	**0.0026**	**-0.00*[a]**
Beneficiaries with ESRD	2,791,479	5,317,996	0.0031	-0.00[a]
Beneficiaries with skin ulcers	1,115,686	2,002,673	0.0048	-0.00[a]
Beneficiaries with multiple conditions	116,366	222,121	0.0235	-0.02[a]
Beneficiaries in rural areas	832,442	1,647,269	0.0056	-0.01***
Beneficiaries with hospital bed claims	351,644	576,820	0.0117	-0.07[a]
Beneficiaries with dual eligibility	1,493,793	3,175,494	0.0056	-0.07*[a]
Probability of unplanned admission	**4,023,531**	**7,542,790**	**0.0916**	**-0.84*** **
Beneficiaries with ESRD	2,791,479	5,317,996	0.1043	-0.74***
Beneficiaries with skin ulcers	1,115,686	2,002,673	0.1821	-1.06***
Beneficiaries with multiple conditions	116,366	222,121	0.6499	-1.77***
Beneficiaries in rural areas	832,442	1,647,269	0.1975	-0.56***
Beneficiaries with hospital bed claims	351,644	576,820	0.3599	-0.94***
Beneficiaries with dual eligibility	1,493,793	3,175,494	0.1546	-0.86***

Outcome and sample	Beneficiary quarterly observations in model states	Beneficiary quarterly observations in comparison states	90 percent half-confidence interval	Observed effect
Number of unplanned admissions	**4,023,531**	**7,542,790**	**0.0018**	**-0.02*****
Beneficiaries with ESRD	2,791,479	5,317,996	0.0020	-0.01***
Beneficiaries with skin ulcers	1,115,686	2,002,673	0.0035	-0.03***
Beneficiaries with multiple conditions	116,366	222,121	0.0158	-0.06***
Beneficiaries in rural areas	832,442	1,647,269	0.0038	-0.01***
Beneficiaries with hospital bed claims	351,644	576,820	0.0077	-0.03***
Beneficiaries with dual eligibility	1,493,793	3,175,494	0.0033	-0.02***
Probability of death (percentage)	**4,023,531**	**7,542,790**	**0.0498**	**-0.04[a]**
Beneficiaries with ESRD	2,791,479	5,317,996	0.0469	-0.05†[a]
Beneficiaries with skin ulcers	1,115,686	2,002,673	0.1188	0.01[a]
Beneficiaries with multiple conditions	116,366	222,121	0.4158	-0.54*[a]
Beneficiaries in rural areas	832,442	1,647,269	0.1059	-0.06[a]
Beneficiaries with hospital bed claims	351,644	576,820	0.2217	-0.01[a]
Beneficiaries with dual eligibility	1,493,793	3,175,494	0.0817	-0.07[a]

Note: Observed effects are calculated as regression-based average marginal effects and are rounded to the nearest one-hundredth of a percent. As a result, some observed effects, including those that are statistically significant, round to zero. The analysis is based on observations in each quarter, so the observed effects and intervals should be interpreted as the average observed change in each quarter of the post-implementation years and the area around that observed difference where the true difference likely is found. For example, the observed impact for "any ambulance utilization" was -1.80, which indicates the average probability of a beneficiary in the model states using an ambulance after prior authorization began was 1.8 percent lower than for beneficiaries in the comparison states. The 90 percent half-confidence interval for this estimate was 0.11, so the true difference in probabilities is likely to be -1.91 to -1.69 percent, using a 10 percent Type I error rate. As this interval does not include zero, we reject the null hypothesis of no difference. The model states included Delaware; Maryland; New Jersey; North Carolina; Pennsylvania; South Carolina; Virginia; Washington, DC; and West Virginia. The comparison states included Alabama, Florida, Georgia, Indiana, Kentucky, Louisiana, Massachusetts, Montana, Nebraska, Ohio, Tennessee, Texas, and Washington.

[a] Test of a difference different from 0 has less than 80 percent power (1 minus Type II error).

† $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

ED = emergency department; ESRD = end-stage renal disease.

In general, our analyses had high statistical precision. We estimated nearly all impacts precisely enough to reject the null hypothesis of no difference. Thus, we recommend reviewing both the statistical significance and the magnitude of the impact estimates, as some of them are very precise yet small in magnitude.

This page has been left blank for double-sided copying.

Appendix G

Statistical precision of estimated model effects on costs

This page has been left blank for double-sided copying.

In this appendix, we provide an analysis of the statistical precision for detecting effects of the RSNAT prior authorization model on different outcomes related to RSNAT service costs, assuming such effects exist. Outcomes related to these costs include expenditures for RSNAT services and all ambulatory services (measured at the beneficiary level) and payments for them (measured at the supplier level). Table G.1 presents precision estimates for the expenditure outcomes. They include separate results for beneficiary-level outcomes for beneficiaries with end-stage renal disease (ESRD) only, pressure ulcers only, or multiple conditions; beneficiaries in rural areas; beneficiaries with hospital bed claims; and dual-eligible beneficiaries. See Appendix F for more information on how precision estimates are calculated.

As with the utilization estimates, we generally measured the impact estimates for cost outcomes with high precision. Again, we caution that some estimates are very precisely estimated, resulting in statistical significance even when the impact estimates are quite small.

Table G.1. Statistical precision for RSNAT expenditure outcomes

Outcome and sample	Beneficiary or supplier quarterly observations in model states	Beneficiary or supplier quarterly observations in matched-comparison states	90 percent half-confidence interval	Observed effect
Expenditures, any ambulance	**4,023,531**	**7,542,790**	**10.9**	**-333*** **
Beneficiaries with ESRD	2,791,479	5,317,996	14.1	-402***
Beneficiaries with skin ulcers	1,115,686	2,002,673	4.7	-41***
Beneficiaries with multiple conditions	116,366	222,121	118.0	-1325***
Beneficiaries in rural areas	832,442	1,647,269	27.4	-223***
Beneficiaries with hospital bed claims	351,644	576,820	60.9	-746***
Beneficiaries with dual eligibility	1,493,793	3,175,494	21.8	-489***
Expenditures, RSNAT services	**4,023,531**	**7,542,790**	**8.3**	**-265*** **
Beneficiaries with ESRD	2,791,479	5,317,996	10.7	-326***
Beneficiaries with skin ulcers	1,115,686	2,002,673	2.0	-13***
Beneficiaries with multiple conditions	116,366	222,121	86.8	-1039***
Beneficiaries in rural areas	832,442	1,647,269	17.1	-138***
Beneficiaries with hospital bed claims	351,644	576,820	44.9	-584***
Beneficiaries with dual eligibility	1,493,793	3,175,494	16.3	-382***
Total expenditures	**4,023,531**	**7,542,790**	**56.9**	**-316*** **
Beneficiaries with ESRD	2,791,479	5,317,996	61.1	-443***
Beneficiaries with skin ulcers	1,115,686	2,002,673	115.9	245***
Beneficiaries with multiple conditions	116,366	222,121	583.9	-1060**
Beneficiaries in rural areas	832,442	1,647,269	112.9	-463***
Beneficiaries with hospital bed claims	351,644	576,820	233.3	-478***
Beneficiaries with dual eligibility	1,493,793	3,175,494	94.8	-419***

Outcome and sample	Beneficiary or supplier quarterly observations in model states	Beneficiary or supplier quarterly observations in matched-comparison states	90 percent half-confidence interval	Observed effect
Number of denied non-emergency ambulance claims	**4,023,531**	**7,542,790**		
Q1 average marginal effect			0.01	0.05***
Q2 average marginal effect			0.01	0.03***
Q3 average marginal effect			0.01	0.03***
Q4 average marginal effect			0.01	0.03***
Q5 average marginal effect			0.01	0.03***
Q6 average marginal effect			0.01	0.02***
Q7 average marginal effect			0.01	0.02***
Q8 average marginal effect			0.01	0.01*a
Q9 average marginal effect			0.01	0.00
Q10 average marginal effect			0.01	0.00
Q11 average marginal effect			0.01	0.00
Q12 average marginal effect			0.01	0.01*a
Q13 average marginal effect			0.01	0.01*a
Q14 average marginal effect			0.01	0.00
Q15 average marginal effect			0.01	0.00a
Q16 average marginal effect			0.01	0.00a
Proportion of non-emergency ambulance claims denied	**462,227**	**782,325**		
Q1 average marginal effect			0.00	0.03***
Q2 average marginal effect			0.00	0.02***
Q3 average marginal effect			0.00	0.02***
Q4 average marginal effect			0.00	0.02***
Q5 average marginal effect			0.00	0.02***
Q6 average marginal effect			0.00	0.01***
Q7 average marginal effect			0.00	0.02***
Q8 average marginal effect			0.00	0.01***

Outcome and sample	Beneficiary or supplier quarterly observations in model states	Beneficiary or supplier quarterly observations in matched-comparison states	90 percent half-confidence interval	Observed effect
Q9 average marginal effect			0.00	0.01***
Q10 average marginal effect			0.00	0.01***
Q11 average marginal effect			0.00	0.01***
Q12 average marginal effect			0.00	0.01***
Q13 average marginal effect			0.00	0.01***
Q14 average marginal effect			0.00	0.01***
Q15 average marginal effect			0.00	0.01***
Q16 average marginal effect			0.00	0.01***
Number of denied Medicare ambulance claims	**4,023,531**	**7,542,790**		
Q1 average marginal effect			0.01	0.05***
Q2 average marginal effect			0.01	0.03***
Q3 average marginal effect			0.01	0.03***
Q4 average marginal effect			0.01	0.03***
Q5 average marginal effect			0.01	0.03***
Q6 average marginal effect			0.01	0.03***
Q7 average marginal effect			0.01	0.03***
Q8 average marginal effect			0.01	0.02**
Q9 average marginal effect			0.01	0.01[a]
Q10 average marginal effect			0.01	0.00[a]
Q11 average marginal effect			0.01	0.01[a]
Q12 average marginal effect			0.01	0.02***
Q13 average marginal effect			0.01	0.01**
Q14 average marginal effect			0.01	0.00[a]
Q15 average marginal effect			0.01	0.01[a]
Q16 average marginal effect			0.01	0.01**

Outcome and sample	Beneficiary or supplier quarterly observations in model states	Beneficiary or supplier quarterly observations in matched-comparison states	90 percent half-confidence interval	Observed effect
Proportion of Medicare ambulance claims denied	**770,471**	**1,397,207**		
Q1 average marginal effect			0.00	0.01***
Q2 average marginal effect			0.00	0.01***
Q3 average marginal effect			0.00	0.01***
Q4 average marginal effect			0.00	0.01***
Q5 average marginal effect			0.00	0.01***
Q6 average marginal effect			0.00	0.01***
Q7 average marginal effect			0.00	0.01***
Q8 average marginal effect			0.00	0.01***
Q9 average marginal effect			0.00	0.01***
Q10 average marginal effect			0.00	0.01***
Q11 average marginal effect			0.00	0.01***
Q12 average marginal effect			0.00	0.01***
Q13 average marginal effect			0.00	0.01***
Q14 average marginal effect			0.00	0.01***
Q15 average marginal effect			0.00	0.01***
Q16 average marginal effect			0.00	0.01***

Note: Observed effects are calculated as regression-based average marginal effects and are rounded to the nearest one-hundredth of a percent. As a result, some observed effects, including those that are statistically significant, round to zero. The model states included Delaware; Maryland; New Jersey; North Carolina; Pennsylvania; South Carolina; Virginia; Washington, DC; and West Virginia. The comparison states included Alabama, Florida, Georgia, Indiana, Kentucky, Louisiana, Massachusetts, Montana, Nebraska, Ohio, Tennessee, Texas, and Washington.

[a] Test of a difference different from 0 has less than 80 percent power (1 minus Type II error).

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

ESRD = end-stage renal disease.

Mathematica

Princeton, NJ • Ann Arbor, MI • Cambridge, MA
Chicago, IL • Oakland, CA • Seattle, WA
Tucson, AZ • Woodlawn, MD • Washington, DC

EDI Global, a Mathematica Company

Bukoba, Tanzania • High Wycombe, United Kingdom

Mathematica
Progress Together

Mathematica.org