New Medical Technology and Clinical Practice: Evidence from AI Stroke Triage

Miaoqing Jia, Lucy Xiaolu Wang *

Abstract

AI stroke-triage software aims to accelerate recognition and care coordination in ischemic stroke. Mechanical thrombectomy is highly effective, but its use depends on rapidly identifying eligible patients. Using national multi-payer claims covering 420,294 ischemic stroke admissions at 2,828 facilities from 2017 through 2023, we estimate the effect of facility-level adoption on the thrombectomy rate with a staggered-adoption difference-in-differences design. Adoption raises the thrombectomy rate by 1.08 percentage points, or 23 percent relative to the 4.73 percent pre-period baseline. The increase is immediate, stable, and concentrated at facilities with existing but not high baseline use. It is not driven by more scanning, as evidenced by imaging utilization rate patterns around adoption. We find no changes in thrombolysis, transfer, or hemorrhagic-stroke treatment rates. Results are robust to modern causal inference estimators, alternative exposure definitions, and different control groups. The pattern is consistent with faster recognition within established stroke programs rather than increased scanning or treatment substitution.

Keywords: artificial intelligence, technology adoption, stroke, mechanical thrombectomy, payment policy

JEL Codes: I11, I18, O33, L86

*Jia (mij4018@med.cornell.edu): Weill Cornell Medicine, Department of Population Health Sciences; Wang (xiaoluwang@umass.edu): University of Massachusetts Amherst, Max Planck Institute for Innovation and Competition, and Canadian Centre for Health Economics. The authors share equal contributions and are ordered alphabetically. All errors are ours.

1 Introduction

Mechanical thrombectomy is among the most effective treatments for ischemic stroke. In pooled trials of endovascular therapy for large-vessel-occlusion stroke, 46 percent of treated patients achieved functional independence against 27 percent of controls (Goyal et al., 2016), and extended-window trials with the 2018 guideline update widened eligibility to 24 hours after onset (Nogueira et al., 2018; Albers et al., 2018; Powers et al., 2018). Many eligible patients nevertheless go untreated (Kamel et al., 2021). The binding constraint is not procedural capacity but recognition: the occlusion must be identified on imaging quickly, and the patient connected to an interventional team while treatment remains possible.

Since 2018 the Food and Drug Administration has authorized a class of software aimed at exactly this step (U.S. Food and Drug Administration, 2018). The software analyzes computed tomographic angiography (CTA) as it is acquired and, when it suspects a large-vessel occlusion, notifies the stroke team directly, typically within minutes. It does not diagnose or direct treatment. Hospitals report use of this software with an inpatient procedure code, ICD-10-PCS 4A03X5D, which Medicare designated in October 2020 as the code to report for the New Technology Add-on Payment (NTAP) accompanying the first authorized device. This created the first national measure of facility adoption of this class of artificial intelligence (AI) software. The NTAP provided up to $1,040 per fee-for-service Medicare inpatient discharge through September 2022 (Hassan, 2021; Centers for Medicare and Medicaid Services, 2020).

Whether adoption of AI stroke-triage software increased the use of mechanical thrombectomy is unknown. Estimating its effect is challenging because adoption is nonrandom, the code became reportable only in October 2020, and billing declined after the NTAP expired. We addressed these challenges by comparing thrombectomy use within facilities before and after adoption with concurrent changes at facilities not yet adopting, and by complementing this with a co-primary analysis limited to the period when billing of the AI code was in continuous use. We estimated the effect of facility-level adoption on the share of ischemic stroke admissions treated with mechanical thrombectomy.

Adoption increases mechanical thrombectomy use by 1.08 percentage points, a 23 percent rise over the pre-code baseline. The increase appears in the month of adoption, persists through at least two years, and is concentrated at facilities with existing but not high baseline use, while thrombolysis, transfers, and a hemorrhagic-stroke negative control are unchanged. The increase is

not driven by more scanning: imaging volumes rise, but they rise before adoption and rise equally for imaging the software cannot affect, and imaging utilization rates show no increase at adoption. This pattern is consistent with faster recognition of already-eligible patients rather than more scanning or treatment substitution, and it is robust across several modern causal-inference estimators, alternative exposure definitions, and comparison groups.

These results add to a small but growing literature on artificial intelligence in acute stroke care. A four-hospital randomized trial of this class of software found faster treatment but was not powered to assess utilization (Martinez-Gutierrez et al., 2023); a study of 26 English hospitals found thrombectomy roughly doubled within a single-payer system (Nagaratnam et al., 2025); and prior U.S. work using the same billing code characterized which hospitals adopt (Pelzl et al., 2026), while a national claims study deemed treatment effects from this application unmeasurable in its data (Wu et al., 2024). We provide national, multi-payer evidence that adoption changed treatment use, and connect it to the temporary payment that accompanied the technology's arrival.

This paper proceeds as follows. Section 2 describes the setting and data. Section 3 details the empirical strategy. Section 4 presents the main results, heterogeneity, and robustness. Section 5 discusses mechanisms and policy implications and concludes.

2 Setting and Data

2.1 Data and Study Population

We used the Inovalon closed-claims database, a national multi-payer database of medical claims for beneficiaries in Medicare Advantage, managed Medicaid, and commercial insurance across all 50 states (Inovalon, 2024). Traditional fee-for-service Medicare is not included. Because the NTAP accompanying the AI billing code applied only to fee-for-service Medicare inpatient discharges, none of the admissions in these data received the NTAP. However, facilities in our sample may also have treated fee-for-service Medicare patients outside the database and therefore may have been eligible to receive the NTAP for those admissions.

We identified inpatient admissions with a principal diagnosis of acute ischemic stroke (International Classification of Diseases, Tenth Revision, Clinical Modification (ICD-10-CM) code I63.x) from January 2017 through December 2023. We restricted the analysis to facilities that regularly treated ischemic stroke, defined as reporting at least one ischemic stroke admission in 12 or more

of the 84 study months, and excluded facilities that billed the AI code on fewer than five admissions. The estimation sample comprised 2,828 facilities and 117,664 facility-months, of which 643 facilities billed the AI code. Admissions at the same facilities with a principal diagnosis of hemorrhagic stroke (I60.x, I61.x, I62.x), for which mechanical thrombectomy is not indicated, served as a negative-control population. To hold the facility set constant, we included only hemorrhagic stroke admissions from facilities in the estimation sample. Figure 1 shows sample construction.

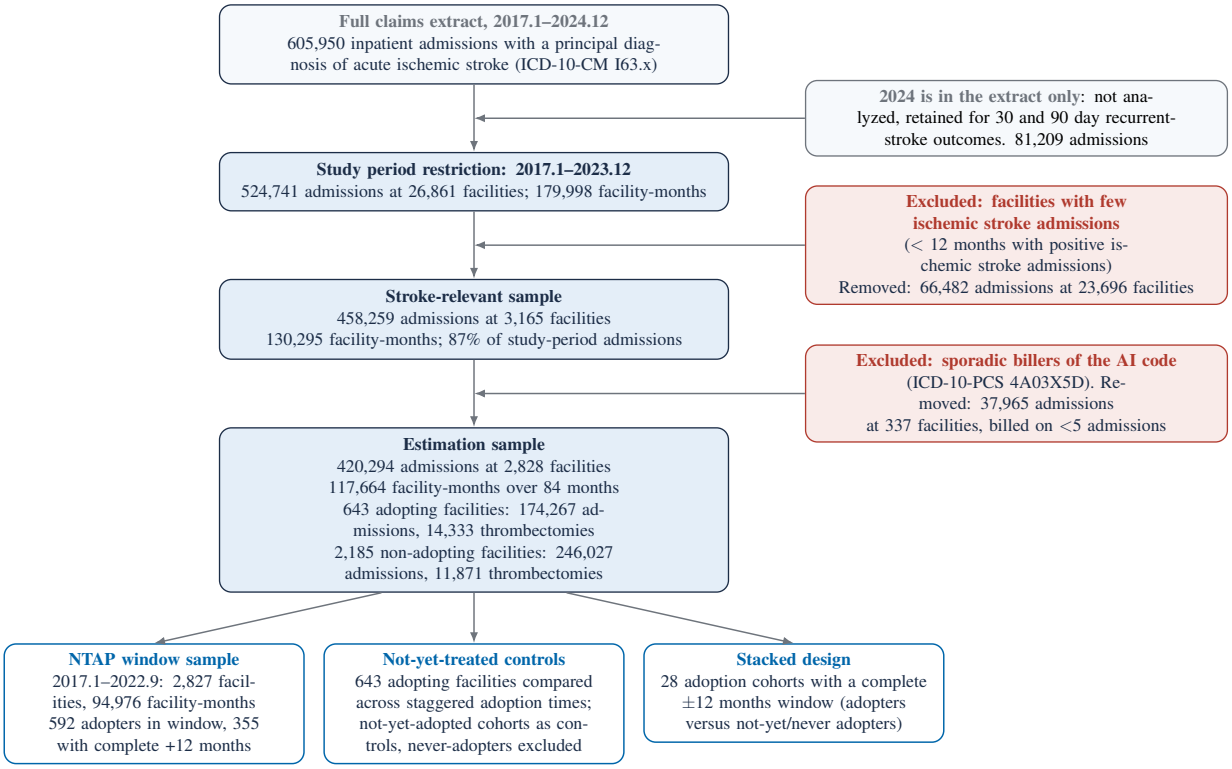


Figure 1: Study Population Flow

Notes: Facilities are billing National Provider Identifiers; a hospital may bill under more than one, and one may cover more than one site. Admissions in 2024 are used to complete recurrent-stroke follow-up and to classify billing-based exposure; outcome analyses end in December 2023. Facilities regularly treating ischemic stroke are those with an admission in at least 12 of the 84 study months. Adopting facilities are those billing the AI code on at least five ischemic stroke admissions across the full extract; facilities billing on fewer are excluded rather than counted as non-adopting. The code became reportable in October 2020, the same month the NTAP began, and the NTAP ran through September 2022. Of the 643 adopting facilities, 3 first billed in September 2020, 592 during the NTAP window, 42 after it expired and within the study period, and 6 in 2024. Because procedure codes apply by discharge date, an admission in late September 2020 discharged after October 1 falls in the fiscal year in which the code became reportable. The negative-control panel comprises 113,311 hemorrhagic stroke admissions and 44,143 facility-months at 2,122 of the same facilities. Abbreviations: NTAP (New Technology Add-on Payment).

2.2 Measuring Adoption

We defined facility-level adoption as the admission month of the first ischemic stroke admission at a billing National Provider Identifier (NPI) carrying ICD-10-PCS code 4A03X5D, among NPIs billing the code on at least five such admissions across the full extract, January 2017 through December 2024. Because hospitals may bill under multiple NPIs and a single NPI may cover multiple sites, adoption was measured at the billing-NPI level. The February 2018 De Novo authorization established the device classification under which subsequent products were cleared (U.S. Food and Drug Administration, 2018), so the code identifies a class of software rather than a specific manufacturer.

Billing-based measures of adoption are subject to three sources of potential misclassification. The code became reportable only in October 2020, billing reflects coding practice rather than software deployment, and billing declined after the NTAP expired. Each source would tend to identify adoption later than it occurred, or fail to identify some adoption altogether, biasing estimates toward the null. We prespecified a co-primary analysis truncated at September 2022, when the NTAP expired and billing of the AI code began to decline, restricting adopters to facilities whose first billing occurred during that period. We also assessed sensitivity to the five-admission requirement by defining adoption as any billing of the code, as billing in at least three distinct months, and as billing on at least eight admissions. We separately re-estimated the model without the 12-month stroke-volume screen.

2.3 Outcomes

The primary outcome was the proportion of ischemic stroke admissions at each facility in each month that were treated with mechanical thrombectomy. The software is intended to accelerate recognition for thrombectomy, not alter thrombolysis eligibility. Thrombolysis was prespecified as a substitution test of whether increased thrombectomy use reduced thrombolysis. Secondary outcomes were length of stay and recurrent ischemic stroke admission within 30 and 90 days. Recurrent admissions were identified by following each admission forward to the patient's next ischemic stroke admission using claims through December 2024. In-hospital mortality could not be measured. Discharge disposition was examined but not carried forward as an outcome, because adoption was associated with a change in how often it was recorded (Figure 5). Table A1 lists all

diagnosis and procedure codes.

3 Empirical Strategy

The unit of analysis was the facility-month, and the primary outcome was the proportion of ischemic stroke admissions treated with mechanical thrombectomy. We compared changes in thrombectomy use before and after adoption with concurrent changes at facilities that had not yet adopted or never adopted. Rate outcomes were weighted by monthly ischemic stroke admissions; count outcomes, including imaging volume and admissions, were unweighted. Standard errors were clustered by facility throughout.

Following published guidance for staggered-adoption difference-in-differences (DiD) designs, we account for time-invariant facility differences and common calendar-time shocks (Wing et al., 2024b). The TWFE specification includes facility and calendar-month fixed effects, comparing within-facility changes after adoption with concurrent changes at facilities not yet or never adopting. Because a conventional two-way fixed-effects (TWFE) estimator can use already-treated facilities as controls and place negative weight on some comparisons when effects vary across adoption cohorts, we treat TWFE as the parsimonious main specification but complement it with three estimators that accommodate cohort-specific effects, set out below; broad concordance across them indicates that TWFE's weighting is not driving the results. The panel truncated at NTAP expiration is prespecified as a co-primary analysis.

Formally, let y_{it} be the share of facility i's ischemic stroke admissions in month t treated with mechanical thrombectomy, G_i the adoption month, $D_{it} = \mathbf{1}\{t \geq G_i\}$, and n_{it} monthly admissions; rate outcomes are weighted by n_{it} and count outcomes are unweighted, with standard errors clustered by facility throughout. The main specification is

$$y_{it} = \beta D_{it} + \alpha_i + \lambda_t + \varepsilon_{it}, \tag{1}$$

where α_i are facility fixed effects and λ_t calendar-month fixed effects, so β is the change in the thrombectomy share after adoption relative to facilities not adopting in the same calendar month. To trace the dynamics and examine pre-trends, the event-study analog replaces the single post-

adoption indicator with relative-time indicators,

$$y_{it} = \sum_{\substack{k=-12 \\ k\neq -1}}^{12} \beta_k \mathbf{1}\{t - G_i = k\} + \beta^{\text{pre}} \mathbf{1}\{t - G_i < -12\} + \beta^{\text{post}} \mathbf{1}\{t - G_i > 12\} + \alpha_i + \lambda_t + \varepsilon_{it}, \quad (2)$$

with the month before adoption as the reference; periods more than 12 months before or after adoption are binned into single endpoints (Schmidheiny and Siegloch, 2023).

The Callaway and Sant'Anna estimator (Callaway and Sant'Anna, 2021) compares each adopting cohort only with facilities not yet adopting—future adopters, which resemble adopters more closely than never-adopters do—through the group-time average treatment effect

$$ATT(g,t) = \mathbb{E}\left[y_t - y_{g-1} \mid G = g\right] - \mathbb{E}\left[y_t - y_{g-1} \mid G > t\right]. \quad (3)$$

We report both its overall admission-weighted aggregation, $\theta^{\text{overall}} = \sum_g \sum_{t \geq g} w_{gt} ATT(g,t)$ (Table 2, Figure 2), and its event-time aggregation, $\theta(k) = \sum_g w_g ATT(g, g+k)$ (Figure 3); the months 0–12 average is the mean of $\theta(k)$ over $k \in [0, 12]$. The de Chaisemartin and D'Haultfœuille estimator (de Chaisemartin and D'Haultfœuille, 2020) identifies effects from facilities switching into adoption, with symmetric placebos for the pre-adoption months. The stacked design (Wing et al., 2024a) builds, for each cohort g with a complete window, a dataset from cohort-g adopters and comparison facilities over months $g-12$ to $g+12$ indexed by stack s, pooled as

$$y_{ist} = \beta D_{ist} + \alpha_{is} + \lambda_{ks} + \varepsilon_{ist}, \quad (4)$$

with facility-by-stack and relative-month-by-stack fixed effects so that no comparison crosses stacks. As a further robustness check, we add facility-specific linear trends $\gamma_i t$ to equation (1).

Pre-adoption coefficients were used to assess time-varying confounders. We evaluated the power of these tests to detect a linear violation of parallel trends (Roth, 2022). Separately, we re-estimated the main specification with facility-specific linear trends, which may partly absorb a persistent treatment effect when fitted over the full panel (Wolfers, 2006). Three additional analyses were prespecified. First, we estimated event studies for imaging volume and utilization rate across modalities, including computed tomographic angiography, non-contrast computed tomography, magnetic resonance imaging of the brain, and magnetic resonance angiography, with com-

puted tomographic angiography of the chest, abdomen, and pelvis as a negative control. Claims were counted across all settings, because the software operates at imaging in the emergency department rather than after admission. Second, we used randomization inference for facilities that adopted after the NTAP expired. Only 42 facilities did so, too few to rely on the asymptotic approximation behind cluster-robust standard errors. Third, we translated the estimated effect into additional thrombectomies and additional patients achieving functional independence. This used annual ischemic stroke volume at adopting facilities and the absolute difference in functional independence from the pooled thrombectomy trials (Goyal et al., 2016). The analysis used deidentified claims under a data-use agreement and did not constitute human subjects research. The estimating equations are set out above. Analysis code is available in a dedicated repository.

4 Results

4.1 Selection into Adoption

Of the 2,828 facilities in the estimation sample, 643 adopted. Adopting facilities were larger and more likely to have performed the procedure before the software existed: they averaged 3.92 ischemic stroke admissions per month against 2.36 at other facilities, and their thrombectomy rate over 2017 through 2019 was 4.85 percent against 2.13 percent (Table 1). Non-adopting facilities carried a larger share of managed Medicaid admissions than adopters, 41.2 against 35.1 percent. In a facility-level model, admission volume, any prior thrombectomy, and baseline thrombolysis use predicted adoption, whereas payer mix did not (Table A2).

4.2 Main Estimates and Timing

Thrombectomy use rose after adoption. The main specification gave 1.08 percentage points (95% CI, 0.54 to 1.62), and the NTAP-window sample 0.85 points (95% CI, 0.25 to 1.45). Against an admission-weighted rate of 4.73 percent across all facilities before October 2020, this is a relative increase of 23 percent. Estimators designed to be robust to heterogeneous treatment effects across adoption cohorts gave 1.48 points for CSDID and 1.79 for DCDH, both above the main specification, and 0.82 to 0.94 for the stacked designs, which restrict comparisons to a plus-or-minus-12-month window around 28 adoption cohorts. The three estimators trace a similar path in

Table 1: Characteristics of Facilities by Adoption Status

	Adopting	**Non-adopting**
Facilities	643	2,185
Mean monthly ischemic stroke admissions	3.92	2.36
Mean monthly ischemic stroke admissions, 2017–2019	2.93	1.95
Mechanical thrombectomy rate, 2017–2019, %	4.85	2.13
Intravenous thrombolysis rate, 2017–2019, %	10.05	6.84
Payer mix, % of ischemic stroke admissions		
Commercial	35.0	29.4
Medicare Advantage	29.9	29.4
Managed Medicaid	35.1	41.2
Event counts, 2017–2023		
Ischemic stroke admissions	174,267	246,027
Mechanical thrombectomy	14,333	11,871
Intravenous thrombolysis	18,648	21,313
Recurrent ischemic stroke, 30 day	3,397	6,873
Recurrent ischemic stroke, 90 day	5,698	11,795

Notes: Values describe the 2,828 facilities in the estimation sample. Adopting facilities are those billing the AI code on at least five ischemic stroke admissions. Values for 2017 through 2019 precede the introduction of the code in October 2020. Payer shares are each facility's share of ischemic stroke admissions in each payer group. Recurrent ischemic stroke is the next ischemic stroke admission for the same patient within the stated interval, not all-cause readmission.

the months after adoption (Figure A2). The four alternative sample and adoption definitions ranged from 0.98 to 1.13 (Figure 2).

Pre-adoption coefficients are flat and jointly indistinguishable from zero. The increase appears in the month of adoption and persists, with an average effect over months 0 to 12 of 1.44 points (95% CI, 0.02 to 2.85; Figure 3, Panel A). Truncating the panel at September 2022, when the code was in continuous use, gives the same shape (Figure 3, Panel B). Allowing facility-specific linear trends produced a smaller and less precise estimate (0.54 percentage points; 95% CI, −0.17 to 1.24), consistent with partial absorption of a persistent treatment effect by trends fitted over the full panel (Wolfers, 2006). In addition, a power calculation following Roth showed that the event-study pre-trend tests could not reliably detect the divergence of 0.068 percentage points per month needed to explain the full estimate (Roth, 2022). We therefore evaluated that threat directly: pre-code trends differed by only 0.003 points per month between adopting and non-adopting facilities, in a direction that would attenuate rather than inflate the estimate (Table A3).

4.3 Heterogeneity

The increases in thrombectomy rate vary by baseline capability (P<0.001 for equality across groups). Increases were concentrated among facilities in the lower two tertiles of the positive baseline distribution: 1.85 percentage points in Tertile 1 and 1.67 points in Tertile 2. By contrast, no increase was detected at either extreme—facilities that performed no thrombectomies during 2017–2019 (-0.43 points) or those in Tertile 3 (0.40 points; median baseline rate, 14.4 percent). On average, estimates for Tertiles 1 and 2 exceeded those for the two extreme groups by 1.77 points (95% CI, 0.76 to 2.78; Table 2 and Figure 4). Among facilities with any baseline thrombectomy use, mean monthly ischemic stroke admission volume was similar across tertiles (5.6 to 6.2 admissions), so volume differences do not explain the decline in effect from Tertile 1 to Tertile 3. Estimates differed by payer (joint P=0.048), with the largest increase among managed Medicaid admissions and the smallest increase among commercially insured admissions (Table 2).

4.4 Alternative Explanations and Secondary Outcomes

We found no evidence for three alternative explanations that would also produce more thrombectomy. First, adopting facilities may have scanned more patients. Imaging utilization rates show no

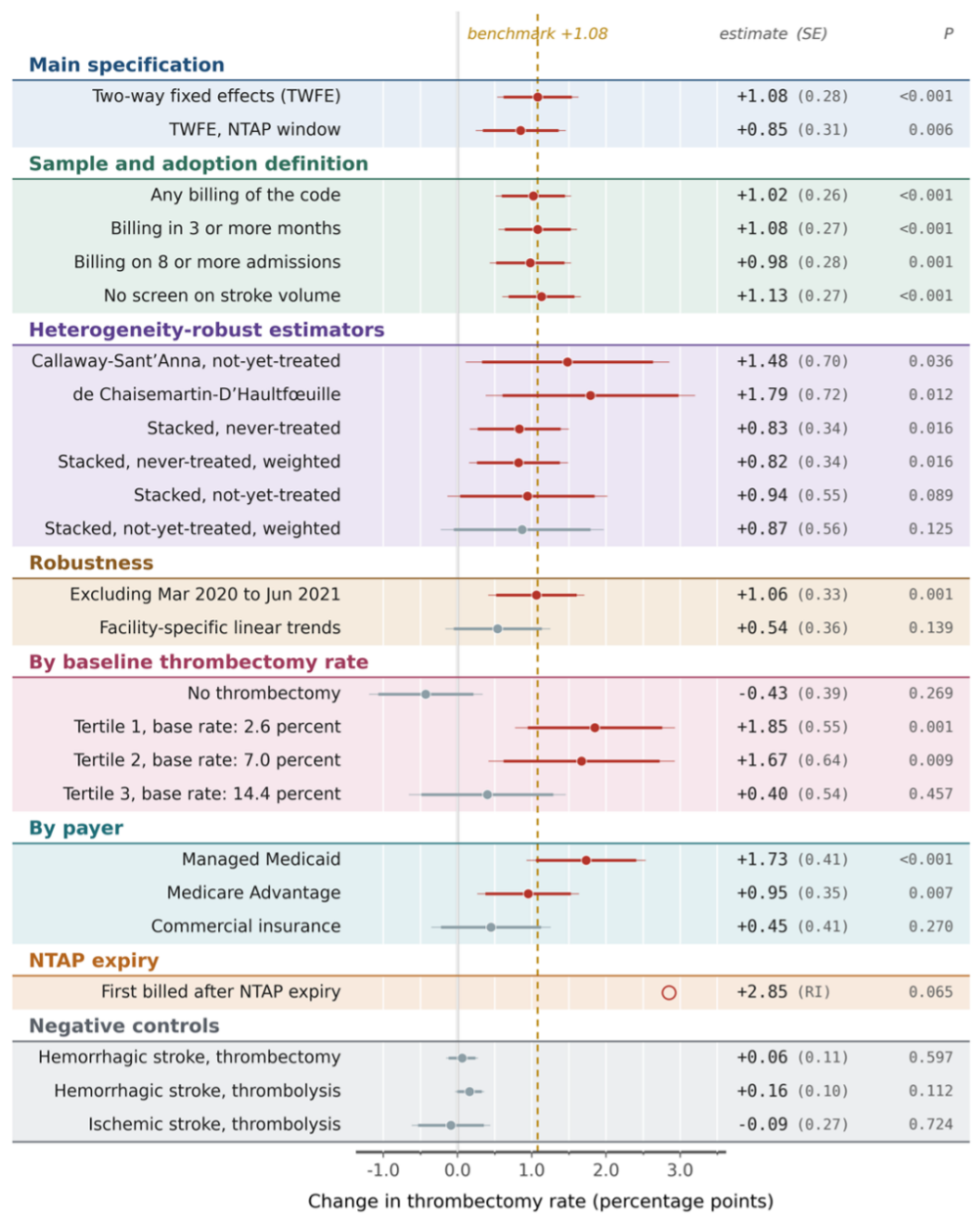


Figure 2: Change in Mechanical Thrombectomy Use Across Specifications

Notes: Each row estimates the percentage-point change in thrombectomy use. Unless otherwise noted, specifications use the main panel of 2,828 facilities and 117,664 facility-months. Bars show 90% (thick) and 95% (thin) confidence intervals. Estimates significant at the 10 percent level are red and the remainder grey. Dashed gold and solid grey lines mark the 1.08-point main estimate and zero, respectively. The main TWFE specification uses the full panel; the next row truncates it at NTAP expiration in September 2022. CSDID and DCDH allow cohort-specific effects. Stacked designs use 28 cohorts with complete $\pm$12-month windows; weighted rows apply optimal weights, and variants differ only in comparison groups (718,608 vs. 49,843 observations). Facility-specific linear trends are fitted over the full panel, which partly absorb the treatment effect and attenuate the estimate. Estimates by baseline thrombectomy rate and by payer each come from a single model with interaction terms; tertiles divide facilities with a positive baseline rate into equal thirds. The NTAP-expiry row rests on 42 adopting facilities compared with 2,185 that never billed the code, with a P value from randomization inference over 1,000 permutations; cluster-robust inference gives 2.85 (SE 1.70), P=0.094. Hemorrhagic stroke admissions at the same facilities serve as negative controls. Abbreviations: CSDID (Callaway and Sant'Anna estimator), DCDH (de Chaisemartin and D'Haultfœuille estimator), NTAP (New Technology Add-on Payment), SE (standard error), TWFE (two-way fixed effects).

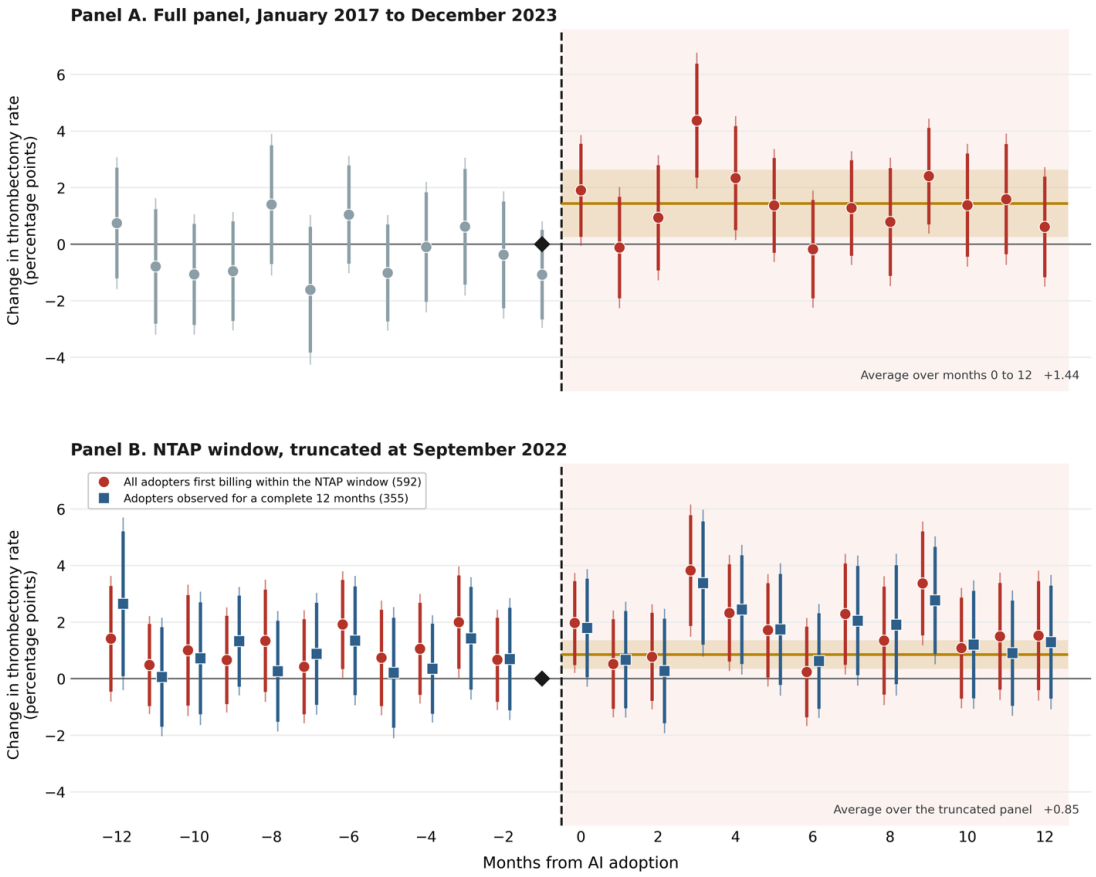


Figure 3: Change in Mechanical Thrombectomy Use by Month from AI Software Adoption

Notes: Points show the estimated change in the thrombectomy rate in each month relative to the month before adoption, marked by the black diamond. Thick bars show 90 percent confidence intervals and thin lines 95 percent; shading marks months after adoption. Estimates are from CSDID with not-yet-adopting facilities as the comparison group, weighted by monthly admissions, with standard errors clustered by facility. Panel A uses 2,828 facilities and 117,664 facility-months; its gold band is the average effect over months 0 to 12, 1.44 percentage points. Panel B truncates the panel at September 2022, so no month after the NTAP expired contributes, and shows all 592 facilities first billing within the NTAP window and the 355 observed for a complete 12 months. Pre-adoption coefficients are jointly consistent with zero in both, P=0.77 and P=0.79. The gold band in Panel B is a TWFE estimate over the truncated panel, 0.85 percentage points, and is not the average of the points above it; the two bands are not directly comparable. Abbreviations: CSDID (Callaway and Sant'Anna estimator), NTAP (New Technology Add-on Payment), TWFE (two-way fixed effects).

Table 2: Change in Mechanical Thrombectomy Use After Adoption of AI Stroke Triage

	Change, pp	95% CI	P	Pre-period rate, %
Main specification				
Two-way fixed effects	1.08	0.54 to 1.62	<0.001	4.73
NTAP-window sample	0.85	0.25 to 1.45	0.006	
Heterogeneity-robust estimators				
Callaway-Sant'Anna, not-yet-treated	1.48	0.10 to 2.86	0.04	
de Chaisemartin-D'Haultfœuille	1.79	0.39 to 3.19	0.01	
Stacked, never-treated	0.83	0.15 to 1.50	0.02	
Stacked, never-treated, weighted	0.82	0.15 to 1.49	0.02	
Stacked, not-yet-treated	0.94	-0.14 to 2.02	0.09	
Stacked, not-yet-treated, weighted	0.87	-0.24 to 1.97	0.13	
Robustness				
Excluding March 2020 through June 2021	1.06	0.41 to 1.71	0.001	
Facility-specific linear trends	0.54	-0.17 to 1.24	0.14	
By months from adoption				
Months 0 to 12	1.44	0.02 to 2.85	0.05	
Months 13 to 24	1.39	-0.17 to 2.94	0.08	
By baseline thrombectomy rate, 2017–2019				
No thrombectomy	-0.43	-1.20 to 0.33	0.27	0.0
Tertile 1, baseline rate: 2.6 percent	1.85	0.76 to 2.93	0.001	2.6
Tertile 2, baseline rate: 7.0 percent	1.67	0.42 to 2.92	0.009	7.0
Tertile 3, baseline rate: 14.4 percent	0.40	-0.66 to 1.47	0.46	14.4
By payer				
Managed Medicaid	1.73	0.93 to 2.52	<0.001	3.95
Medicare Advantage	0.95	0.25 to 1.64	0.007	4.42
Commercial insurance	0.45	-0.35 to 1.25	0.27	5.53
Negative controls, hemorrhagic stroke				
Mechanical thrombectomy	0.06	-0.16 to 0.28	0.60	0.29
Intravenous thrombolysis	0.16	-0.04 to 0.36	0.11	0.44
Secondary outcomes				
Intravenous thrombolysis	-0.09	-0.62 to 0.43	0.72	9.37
Recurrent ischemic stroke, 30 day	-0.10	-0.31 to 0.10	0.33	2.76
Recurrent ischemic stroke, 90 day	0.02	-0.27 to 0.31	0.88	4.72
Length of stay, days	0.03	-0.12 to 0.19	0.68	5.20

Notes: Values are changes in the share of ischemic stroke admissions treated with mechanical thrombectomy, in percentage points; length of stay is in days. The TWFE specifications include facility and calendar-month fixed effects; all models in this table are weighted by monthly ischemic stroke admissions and cluster standard errors by facility. Pre-period rates are admission-weighted shares in months before October 2020; for rows by baseline thrombectomy rate, this column reports the 2017 through 2019 rate defining each group. The estimating equations are set out above. Estimates by baseline thrombectomy rate and by payer each come from a single model with interaction terms, not separate regressions on subgroups. Baseline groups are fixed on 2017 through 2019 rates among the 1,179 facilities with at least 30 ischemic stroke admissions in that period; tertiles divide those with a positive rate into equal thirds. Effects differ across baseline groups (P<0.001) and across payers (P=0.048). Facility-specific linear trends are fitted over the full panel, which partly absorb the treatment effect and attenuate the estimate. Negative controls are hemorrhagic stroke admissions at the same facilities, for which thrombectomy is not indicated. Abbreviations: NTAP (New Technology Add-on Payment).

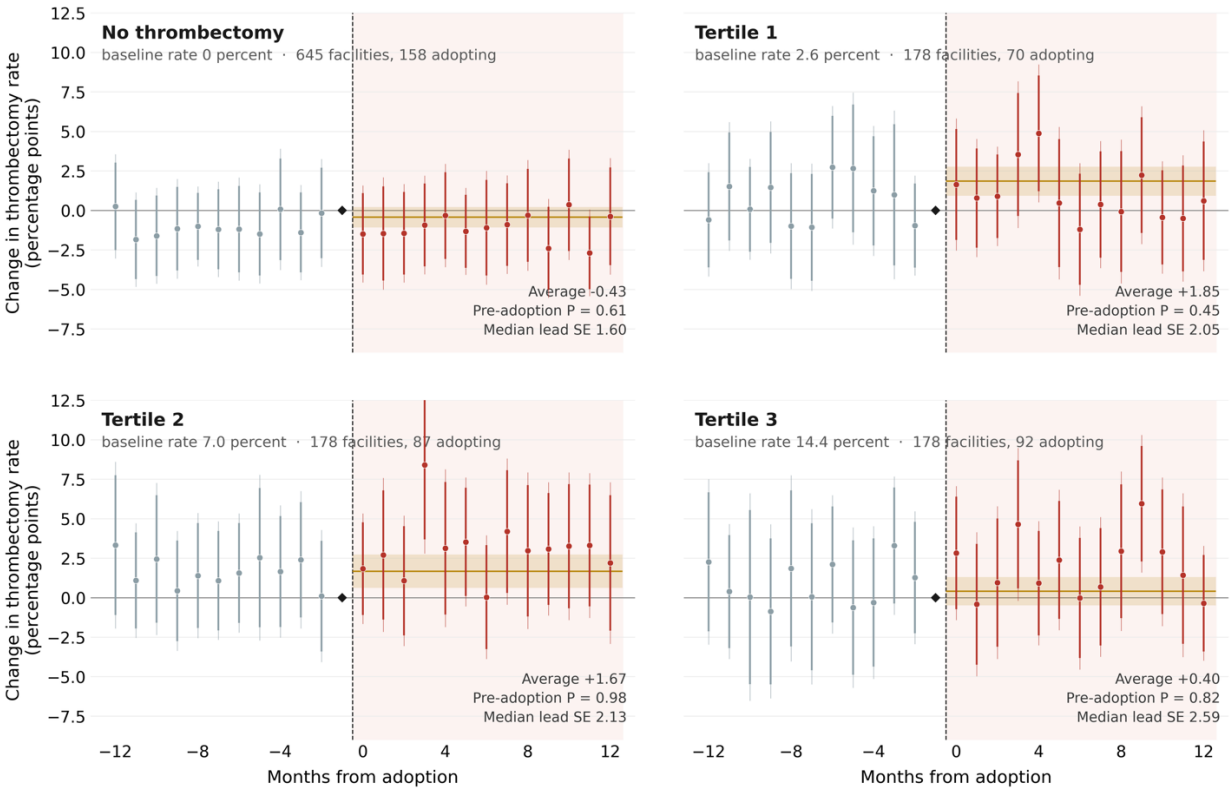


Figure 4: Event Studies by Baseline Capability

Notes: Points show the estimated change in the thrombectomy rate in each month relative to the month before adoption, marked by the black diamond. Thick bars show 90 percent confidence intervals and thin lines 95 percent; shading marks months after adoption. All four panels come from a single model in which each month indicator is interacted with baseline capability group, so the gold band in each panel is that group's average effect. Groups are fixed on 2017 through 2019 thrombectomy rates among the 1,179 facilities with at least 30 ischemic stroke admissions in that period: 645 performed none, and the remaining 534 are divided into equal tertiles. Pre-adoption coefficients are jointly consistent with zero in all four groups, but the median standard error on a pre-adoption coefficient ranges from 1.60 to 2.59 against effects of 1.7 to 1.8 points, so a flat pre-adoption path is weak evidence in three of the four; Tertile 2 is the exception.

increase at adoption (Figure A3). Volumes rose, but they were rising beforehand and rose equally for imaging the software cannot affect, consistent with a shift toward imaging all suspected stroke patients as thrombectomy eligibility widened (Mayer et al., 2020). The increase in thrombectomy therefore does not reflect a change in imaging practice specific to adoption. Second, thrombectomy may have displaced thrombolysis, but thrombolysis use did not change (-0.09 points; 95% CI, -0.62 to 0.43), and the lower bound of that interval rules out substitution for more than 57 percent of the increase. Third, adopting facilities may have retained patients they would previously have transferred, but transfer to other acute care hospitals did not change (0.11 points; 95% CI, -0.09 to 0.31). Among admissions for hemorrhagic stroke, for which thrombectomy is not indicated, neither thrombectomy nor thrombolysis changed (Table 2).

Forty-two facilities first billed the code after the NTAP had expired and within the study period, when no add-on was available to any hospital. Their estimate is 2.85 points, in the same direction as the main result and larger, though not conventionally significant: randomization inference across 1,000 permutations gives P = 0.065 and cluster-robust inference gives P = 0.094. None of the payers in these data received the NTAP.

The average effect over months 13 to 24 after adoption was 1.39 points (95% CI, -0.17 to 2.94), against 1.44 over months 0 to 12, and all 12 second-year coefficients were positive against 4 of 12 pre-adoption coefficients (Figure A1). Length of stay and recurrent ischemic stroke at 30 and 90 days did not change (Table 2). Because the estimated effect is a 1.08 point change in the probability of receiving thrombectomy, a change in any downstream outcome is scaled by that proportion.

5 Discussion and Policy Implications

In this national study of multi-payer claims, facility adoption of AI stroke-triage software was followed by an increase of 1.08 percentage points in mechanical thrombectomy, 23 percent of the pre-code baseline. The increase appeared in the month of adoption and showed no evidence of decay through two years. Applied to the 35,543 ischemic stroke admissions at adopting facilities in the final study year, the estimate corresponds to roughly 384 additional thrombectomies per year (95% CI, 192 to 576) among the admissions we observe, and, using the absolute difference in functional independence from the pooled trials (Goyal et al., 2016), roughly 75 additional patients

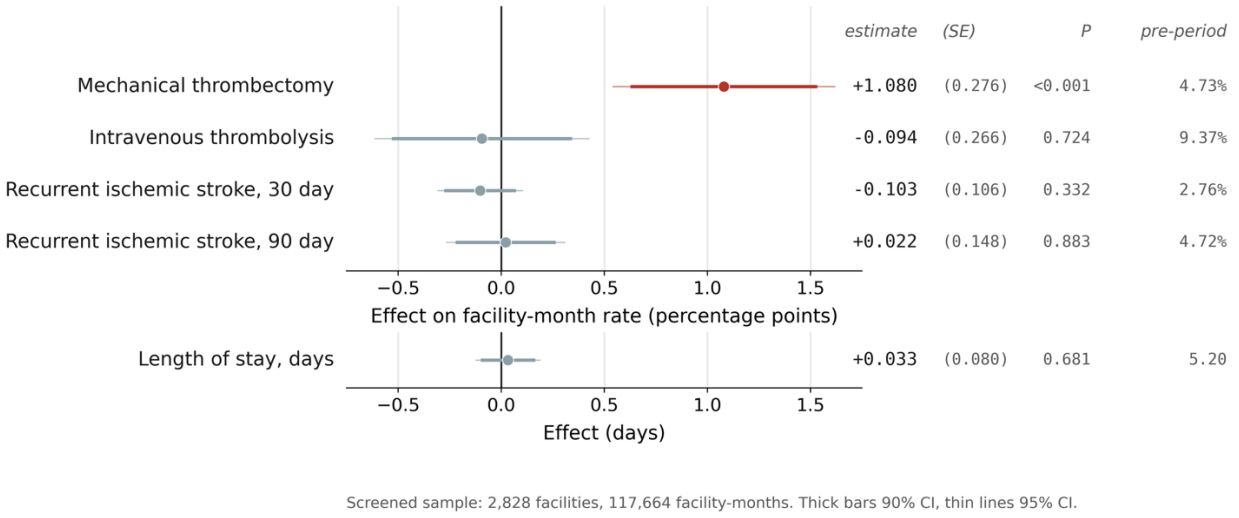


Figure 5: Secondary Outcomes and Sample Screens

Notes: Main-panel estimates use 2,828 facilities and 117,664 facility-months. Thick bars show 90 percent confidence intervals and thin lines 95 percent. Values in the upper panel are changes in the share of ischemic stroke admissions receiving each treatment; length of stay in the lower panel is in days on its own scale. Mechanical thrombectomy is shown for reference. Recurrent ischemic stroke is the next ischemic stroke admission for the same patient within the stated interval, not all-cause readmission. Because adoption changes the probability of thrombectomy by about one percentage point, any downstream effect is diluted by the same proportion, so these results do not rule out effects among treated patients. The 12-month screen does not move the estimate: without it the main specification gives 1.13 rather than 1.08 percentage points. Discharge disposition was recorded for 90 percent of admissions, but adoption was associated with a 1.78 point reduction in the proportion unrecorded, so this outcome was not carried forward.

per year who survive with functional independence (95% CI, 37 to 112). These figures describe our sample rather than the country, because the data do not include traditional Medicare, and the disability arithmetic is illustrative.

Because adoption was measured from the first admission carrying the AI code, billing could instead mark a coincident expansion of stroke-program capacity. Several findings weigh against this explanation. Imaging utilization rates show no increase at adoption, and rising volumes predate adoption and appear equally for imaging the software cannot affect (Mayer et al., 2020). Transfers to other acute care hospitals did not change. The nonmonotonic pattern by baseline capability also differs from a simple program-expansion explanation: the increase was not detected at facilities with no prior thrombectomy use, was largest at intermediate-capability facilities, and was small at the highest-capability facilities. No change was detected in the hemorrhagic-stroke negative controls. These findings are more consistent with faster recognition at facilities with existing thrombectomy capacity, although time-varying confounding cannot be excluded. Facilities first billing after the NTAP expired had a point estimate at least as large, and no payer in these data received the NTAP, weakening a billing-incentive explanation. In the study supporting authorization, notification reached the specialist earlier than standard workflow in 42 of 44 cases, with a mean difference of 36 to 59 minutes (U.S. Food and Drug Administration, 2018), a reduction of the magnitude associated with better outcomes in pooled trial data (Saver et al., 2016).

Direct NTAP payment is unlikely to explain the estimated increase because our data exclude traditional fee-for-service Medicare, to which the add-on payment applied. Its expiration nevertheless matters for future diffusion. The NTAP provided temporary payments for qualifying technologies for two to three years but ended without a permanent successor; a dedicated outpatient reimbursement mechanism for AI-enabled devices has only recently been proposed (Centers for Medicare and Medicaid Services, 2026). The estimated thrombectomy increase was largest among managed Medicaid admissions, suggesting AI adoption narrowed a baseline treatment gap within adopting facilities. But adoption was associated with facility size and prior capability rather than payer mix, and non-adopting facilities served more managed Medicaid. Population-level benefits will therefore depend on whether and to what extent AI triage reaches Medicaid-serving facilities.

Our findings extend the limited evidence on AI stroke triage and treatment. A four-hospital randomized trial of one of these products, in which sites adopted the software in a staggered sequence, reported faster treatment but was not powered to assess utilization (Martinez-Gutierrez

et al., 2023). At 26 English hospitals adopting a similar tool within a national stroke network, thrombectomy use roughly doubled, from 2.3 to 4.6 percent, in a single-payer system without payment variation (Nagaratnam et al., 2025). A prior U.S. study using the same billing code focused on which hospitals adopted, using a 5 percent Medicare sample restricted to admissions receiving thrombolysis (Pelzl et al., 2026), while a national claims study of AI billing considered treatment effects from this application unmeasurable in its data (Wu et al., 2024). The present study therefore complements trial evidence on treatment speed with national evidence on treatment use.

This study has several limitations. It does not estimate the NTAP's effect on adoption, assess cost-effectiveness, or observe traditional Medicare outcomes, and therefore cannot determine whether permanent payment would increase adoption. The disability translation assumes that trial effects apply to marginal patients who may differ from trial populations. Adoption is inferred from billing, so unbilled and pre-code use may be misclassified, likely attenuating the estimate; the NTAP-window analysis addresses declining billing after expiration. Race and ethnicity are missing for 53 percent of admissions in our data, preventing subgroup analyses. Our data do not contain stroke severity, so we could not assess whether the increase reached the patients most likely to benefit, or in-hospital mortality.

Despite these limitations, our findings have practical implications. For stroke teams, adding an AI fast-notification layer to existing imaging may move a meaningful number of patients to definitive treatment without a change in imaging practice specific to adoption. For payers, they quantify the treatment changes associated with adoption during the temporary-payment period, while leaving unresolved how permanent reimbursement would affect adoption and future diffusion.

References

Albers, G. W., Marks, M. P., Kemp, S., et al. (2018). Thrombectomy for stroke at 6 to 16 hours with selection by perfusion imaging. *New England Journal of Medicine*, 378(8):708–718. doi:10.1056/NEJMoa1713973.

Callaway, B. and Sant'Anna, P. H. C. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics*, 225(2):200–230.

Centers for Medicare and Medicaid Services (2020). Hospital inpatient prospective payment systems, FY2021 final rule. 85 Federal Register 58432.

Centers for Medicare and Medicaid Services (2026). Hospital outpatient prospective payment system, CY2027 proposed rule: interim payment policy for software as a medical service. Proposed rule (CMS-1850-P), 91 Federal Register 41734, July 7, 2026.

de Chaisemartin, C. and D'Haultfœuille, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, 110(9):2964–2996.

Goyal, M., Menon, B. K., van Zwam, W. H., et al. (2016). Endovascular thrombectomy after large-vessel ischaemic stroke: a meta-analysis of individual patient data from five randomised trials. *The Lancet*, 387(10029):1723–1731. doi:10.1016/S0140-6736(16)00163-X.

Hassan, A. E. (2021). New technology add-on payment (ntap) for viz lvo: a win for stroke care. *Journal of NeuroInterventional Surgery*, 13(5):406–408.

Inovalon (2024). Real-world data (rwd) data dictionary (closed claims — consolidated). Accessed August 7, 2026.

Kamel, H., Parikh, N. S., Chatterjee, A., et al. (2021). Access to mechanical thrombectomy for ischemic stroke in the United States. *Stroke*, 52(8):2554–2561. doi:10.1161/STROKEAHA.120.033485; PMID 33980045.

Martinez-Gutierrez, J. C. et al. (2023). Automated large vessel occlusion detection software and thrombectomy treatment times: a cluster randomized clinical trial. *JAMA Neurology*, 80(11):1182–1190. doi:10.1001/jamaneurol.2023.3206.

Mayer, S. A., Viarasilpa, T., Panyavachiraporn, N., et al. (2020). CTA-for-All: impact of emergency computed tomographic angiography for all patients with stroke presenting within 24 hours of onset. *Stroke*, 51(1):331–334. doi:10.1161/STROKEAHA.119.027356.

Nagaratnam, K., Neuhaus, A. A., Fensome, L., et al. (2025). Artificial intelligence imaging decision support for acute stroke treatment in England: a prospective observational study. *The Lancet Digital Health*, 7(12):100927. doi:10.1016/j.landig.2025.100927.

Nogueira, R. G., Jadhav, A. P., Haussen, D. C., et al. (2018). Thrombectomy 6 to 24 hours after stroke with a mismatch between deficit and infarct. *New England Journal of Medicine*, 378(1):11–21. doi:10.1056/NEJMoa1706442.

Pelzl, C. E., Christensen, E. W., Drake, A. R., Rula, E. Y., Sanelli, P. C., and Sanmartin, M. X. (2026). Patient and facility factors associated with uptake of NTAP-billed AI to identify suspected large vessel occlusion in ischemic stroke. *American Journal of Neuroradiology*. doi:10.3174/ajnr.A9494; online June 24, 2026.

Powers, W. J., Rabinstein, A. A., Ackerson, T., et al. (2018). 2018 guidelines for the early management of patients with acute ischemic stroke: a guideline for healthcare professionals from the American Heart Association/American Stroke Association. *Stroke*, 49(3):e46–e110. doi:10.1161/STR.0000000000000158.

Roth, J. (2022). Pretest with caution: event-study estimates after testing for parallel trends. *American Economic Review: Insights*, 4(3):305–322.

Saver, J. L., Goyal, M., van der Lugt, A., et al. (2016). Time to treatment with endovascular thrombectomy and outcomes from ischemic stroke: a meta-analysis. *JAMA*, 316(12):1279–1288. doi:10.1001/jama.2016.13647.

Schmidheiny, K. and Siegloch, S. (2023). On event studies and distributed-lags in two-way fixed effects models: Identification, equivalence, and generalization. *Journal of Applied Econometrics*, 38(5):695–713.

U.S. Food and Drug Administration (2018). De Novo classification request DEN170073 (Viz.ai ContaCT), granted February 13, 2018. https://www.accessdata.fda.gov/cdrh_docs/reviews/DEN170073.pdf.

Wing, C., Freedman, S. M., and Hollingsworth, A. (2024a). Stacked difference-in-differences. Working Paper 32054, National Bureau of Economic Research.

Wing, C., Yozwiak, M., Hollingsworth, A., Freedman, S., and Simon, K. (2024b). Designing difference-in-difference studies with staggered treatment adoption: Key concepts and practical guidelines. *Annual Review of Public Health*, 45:485–505.

Wolfers, J. (2006). Did unilateral divorce laws raise divorce rates? A reconciliation and new results. *American Economic Review*, 96(5):1802–1820.

Wu, K., Wu, E., Theodorou, B., Liang, W., Mack, C., Glass, L., Sun, J., and Zou, J. (2024). Characterizing the clinical adoption of medical AI devices through U.S. insurance claims. *NEJM AI*, 1(1):AIoa2300030.

A Appendix

Table A1: Diagnosis and Procedure Codes

Element	Code system	Codes
Ischemic stroke cohort	ICD-10-CM	I63.x, principal diagnosis
Hemorrhagic stroke, negative control	ICD-10-CM	I60.x, I61.x, I62.x, principal diagnosis
AI stroke triage	ICD-10-PCS	4A03X5D
Mechanical thrombectomy	ICD-10-PCS	03CG3Zx, 03CH3Zx, 03CJ3Zx, 03CK3Zx, 03CL3Zx, 03CM3Zx, 03CN3Zx, 03CP3Zx, 03CQ3Zx
Intravenous thrombolysis	ICD-10-PCS	3E03317
Computed tomographic angiography, head and neck	CPT	70496, 70498
Non-contrast computed tomography, head	CPT	70450
Magnetic resonance imaging, brain	CPT	70551, 70552, 70553
Magnetic resonance angiography, head and neck	CPT	70544–70549
Computed tomographic perfusion	CPT	0042T
Computed tomographic angiography, chest, abdomen, and pelvis	CPT	71275, 74174, 72191

Notes: Ischemic stroke is identified from a principal diagnosis of ICD-10-CM I63.x and hemorrhagic stroke from I60.x, I61.x, or I62.x. Imaging claims are counted across all care settings; all other outcomes are measured on inpatient admissions. Computed tomographic angiography of the chest, abdomen, and pelvis serves as a negative control, since the software analyzes only head and neck angiography. Magnetic resonance imaging of the brain carries the diffusion-weighted sequence and is the magnetic resonance counterpart to non-contrast computed tomography for diagnosing infarct. Abbreviations: AI (artificial intelligence), CPT (Current Procedural Terminology), CTA (computed tomographic angiography), ICD-10-CM (International Classification of Diseases, Tenth Revision, Clinical Modification), ICD-10-PCS (International Classification of Diseases, Tenth Revision, Procedure Coding System), MRA (magnetic resonance angiography), MRI (magnetic resonance imaging), NCCT (non-contrast computed tomography).

Table A2: Determinants of Adoption

Characteristic, 2017 through 2019	Odds ratio	95% CI	P
Mean monthly ischemic stroke admissions, log	12.11	10.11 to 14.50	<0.001
Any mechanical thrombectomy performed	2.82	2.09 to 3.81	<0.001
Intravenous thrombolysis rate, per percentage point	1.010	1.005 to 1.015	<0.001
Mechanical thrombectomy rate, per percentage point	0.996	0.987 to 1.004	0.32
Managed Medicaid share of admissions, per percentage point	1.001	0.998 to 1.004	0.45

Notes: Logistic regression at the facility level. The outcome is any billing of the AI code during the study period. The model includes 24,959 facilities, of which 713 adopted; this is a wider set than the 2,828 facilities in the estimation sample, because adoption can be modelled for facilities too small to contribute to the monthly panel. Covariates are measured over January 2017 through December 2019, before the code became reportable. Facilities with no ischemic stroke admissions in that period are assigned a rate of zero.

Table A3: Pre-Adoption Diagnostics

Quantity	Value
Linear trend fitted to pre-adoption coefficients	-0.016 percentage points per month (SE 0.060)
Joint test of the eleven pre-adoption coefficients	F(11, 2827) = 0.57, P = 0.86
Trend detectable with 80 percent power	0.246 points per month
Trend detectable with 50 percent power	0.186 points per month
Trend that would generate the full estimate	0.068 points per month
Facility trend, adopting facilities, 2017m1-2020m9	0.042 points per month
Facility trend, non-adopting facilities, 2017m1-2020m9	0.045 points per month

Notes: The table asks whether a pre-existing divergence between adopting and comparison facilities could account for the estimated effect. The first five rows describe what the pre-adoption test can detect: a linear divergence of 0.068 percentage points per month, continued after adoption, would generate the entire estimate, which is about one third of the smallest divergence the test detects half the time, so a violation of that size would usually pass unnoticed (Roth, 2022). The last two rows measure the divergence directly. Facility-specific linear trends over January 2017 through September 2020 are 0.042 percentage points per month at adopting facilities and 0.045 at non-adopting facilities; the differential is far smaller than would be required and is in the direction that would understate the effect. Medians are zero in both groups, because most facilities perform no thrombectomy in this period. These trends are fitted over short windows for many facilities and are used only for comparison, not to adjust the outcome.

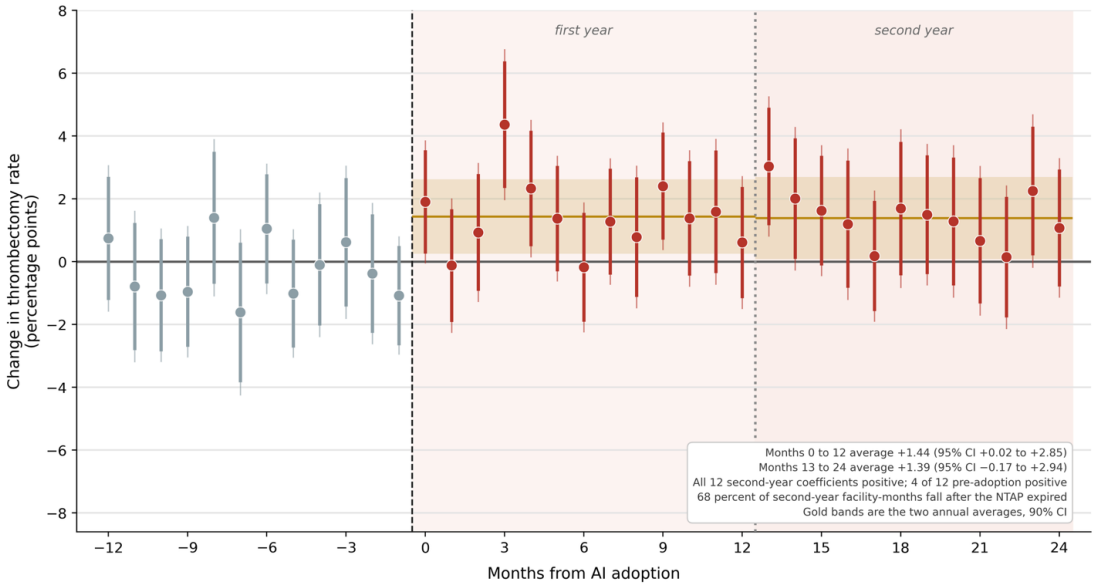


Figure A1: Change in Thrombectomy Use Through Two Years

Notes: Points show the estimated change in the thrombectomy rate in each month relative to the month before adoption. Thick bars show 90 percent confidence intervals and thin lines 95 percent. The two gold bands are the average effects over the first and second years, 1.44 and 1.39 percentage points, each shaded to its 90 percent confidence interval; they are drawn to show that the effect does not decay and are not summaries of the points beneath them. Fewer facilities are observed in later months: 623 of the 643 adopting facilities at month 12, 575 at month 18, and 445 at month 24. Sixty-eight percent of second-year facility-months fall after the NTAP expired. Abbreviations: NTAP (New Technology Add-on Payment).

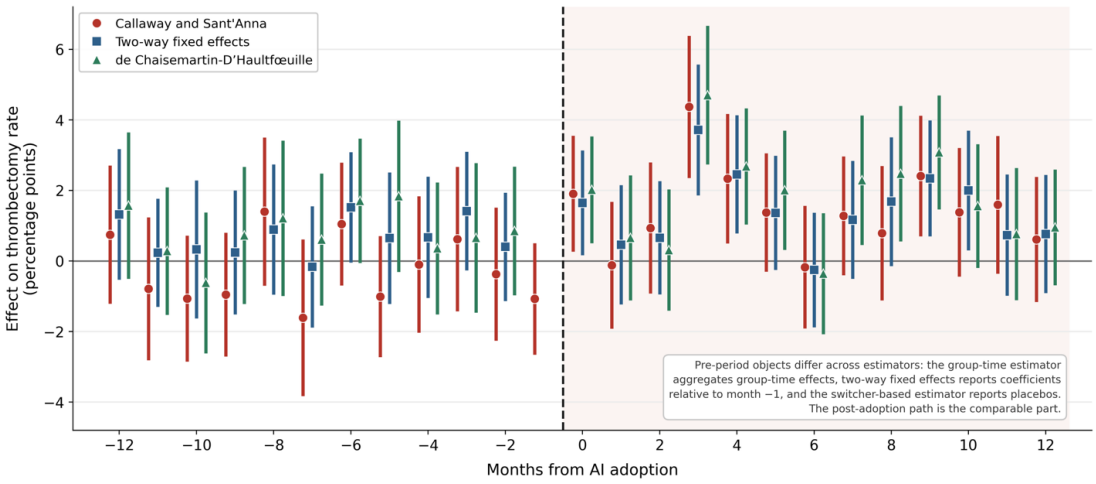


Figure A2: Event Studies by Estimator

Notes: Three estimators plotted on common axes, with points offset horizontally within each month so the series can be told apart. Bars show 90 percent confidence intervals. The months before adoption are not the same quantity across estimators: CSDID aggregates group-time effects, TWFE reports coefficients relative to the month before adoption, and DCDH reports placebo estimates. The months after adoption are comparable, and the three estimators trace them closely. For DCDH the placebo estimates are jointly consistent with zero (P=0.68) while the post-adoption effects are jointly significant (P<0.001). Abbreviations: CSDID (Callaway and Sant'Anna estimator), DCDH (de Chaisemartin and D'Haultfœuille estimator), TWFE (two-way fixed effects).

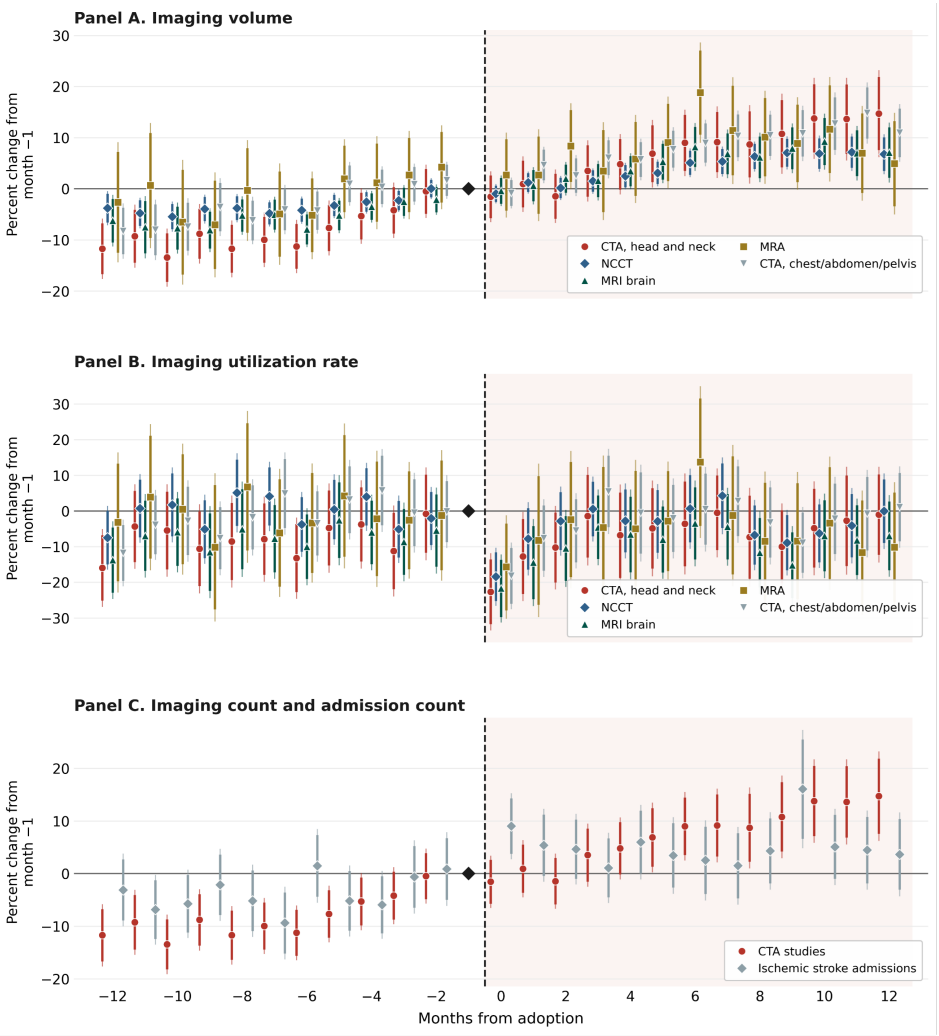


Figure A3: Imaging by Month from Adoption

Notes: Points show the estimated change in each measure in each month relative to the month before adoption, marked by the black diamond. Thick bars show 90 percent confidence intervals and thin lines 95 percent; shading marks months after adoption. Estimates are from TWFE with facility and calendar-month fixed effects and standard errors clustered by facility, and are expressed as a percentage of the mean at month −1 among the 450 adopting facilities observed in that month. All measures in this figure are unweighted, because weighting a count by admissions would give larger facilities influence through both the outcome and the weight. Panel A shows imaging volume as claims per facility-month. Panel B shows imaging utilization rate, defined as studies per ischemic stroke admission; emergency department imaging is not separately identifiable, so inpatient ischemic stroke admissions are the denominator throughout. CTA of the chest, abdomen, and pelvis serves as a negative control, since the software analyzes only head and neck angiography. Computed tomographic perfusion, listed in Table A1, is not shown: it is billed on 0.4 studies per facility-month before adoption, so on a percentage scale its confidence intervals span several times the range of the other measures. It follows the same pattern, with a pre-adoption trend in volume and in utilization. Pre-adoption coefficients are jointly non-zero for every volume measure except MRA, and for utilization of angiography and the negative control; these tests are underpowered, and the pre-adoption slopes are the more informative summary. Utilization falls at the adoption month for all measures in Panel B, including the negative control. Panel C shows why: admissions rise in that month while imaging counts do not, so the ratio falls mechanically. Because adoption is defined as the first month containing an admission that carries the AI code, that month is more likely to be one of higher admission volume. Abbreviations: CTA (computed tomographic angiography), MRA (magnetic resonance angiography), MRI (magnetic resonance imaging), NCCT (non-contrast computed tomography), TWFE (two-way fixed effects).