ARKA
Three ways to change what a physician orders. Only two of them work.
==========
And what happens when you try the same thing on imaging.
Most imaging software implements the one that didn't. Including, until recently, ours.

We also said, until this month, that nobody had studied this on imaging. They had. We were wrong and this page is the correction.

Structure: What we know about imaging; What we know about the mechanism; How the studies compare; What nobody knows; The study we intend to run

Generated from lib/tcoc/content/nudge-evidence.ts + mechanisms-catalog + never-do-catalog.
Ungated. No form.
Version 1.1 ? published 16 August 2026 ? engine tcoc-1.0.3.

?0 What this evidence is, and what it is not

The randomised evidence in this field is not imaging. Meeker 2016: 47 practices, 248 clinicians, 18 months, four arms, antibiotics. Accountable justification DiD ?7.0 pp; peer comparison DiD ?5.2 pp; suggested alternatives no effect vs control; control ?11.0 pp on secular trend alone. Linder 2017: All arms decayed after removal; peer comparison alone remained significantly below control. Domain: not imaging.

The imaging evidence is the Duke Primary Care peer-comparison dashboard, reported on two denominators ? volume and cost ? not two independent findings. Baseline top-decile:bottom-decile variation 4.2?; year-one median imaging rate ?17.3% (within-network pre-post). Median estimated radiology costs ?19.5%; >$3M saved in year one (within-network pre-post). Design: observational, no control group, no risk adjustment, effect not separable from secular trend. These are within-network pre-post changes; an unknown share is secular trend.

No randomised, controlled study of peer-comparison feedback on advanced-imaging ordering has been published. ARKA intends to run it: interrupted time series with a non-equivalent concurrent control group, a pre-registered analysis plan, and a fixed pre-period (Volume III ?17 Study 2). The pre-period lock is the precondition. The study is intended; it needs a data agreement. See the study designs.

Comparison table (generated from lib/sources/data.json)
Study Domain Design Control Risk adj GRADE
---
halpern-2021 imaging observational pre-post no no low
Effect: Baseline top-decile:bottom-decile variation 4.2?; year-one median imaging rate ?17.3% (within-network pre-post)
clark-randall-2021 imaging observational pre-post no no low
Effect: Median estimated radiology costs ?19.5%; >$3M saved in year one (within-network pre-post)
meeker-2016 analogous-domain (not cluster-randomised clinical tr yes no moderate
Effect: Accountable justification DiD ?7.0 pp; peer comparison DiD ?5.2 pp; suggested alternatives no effect vs control; control ?11.0 pp on secular trend alone
linder-2017 analogous-domain (not prespecified secondary analysi yes no low
Effect: All arms decayed after removal; peer comparison alone remained significantly below control

The one sentence to reuse in outreach. The best imaging-specific evidence in this category has no control group, which means nobody can tell a finance team how much of it was the intervention. That is the problem we built the product around.

Study designs: https://arkahealth.com/evidence/methodology

[meeker-2016] Meeker D et al. Effect of Behavioral Interventions on Inappropriate Antibiotic Prescribing. *JAMA.* 2016.
https://pubmed.ncbi.nlm.nih.gov/26864410/
[linder-2017] Meeker D et al. durability follow-up. *JAMA.* 2017.
https://pubmed.ncbi.nlm.nih.gov/29049577/
[halpern-2021] Halpern DJ, Clark-Randall A, Woodall J, Anderson J, Shah K. Reducing Imaging Utilization in Primary Care Through Implementation of a Peer Comparison Dashboard. J Gen Intern Med. 2021 Jan;36(1):108?113.

https://pubmed.ncbi.nlm.nih.gov/32885372/
[clark-randall-2021] Clark-Randall A, Halpern DJ, Taylor J, Roth CJ, Gupta RT, Woodall J, Shah KP. Implementation of a Radiology Utilization Dashboard Yields Significant Cost Savings in a Large Primary Care Network. J Am Coll Radiol. 2021 Jul;18(7):947?950. https://pubmed.ncbi.nlm.nih.gov/33745853/

?1 The four-arm trial

Meeker et al., JAMA 2016 ? cluster-randomised across primary-care practices. Raw within-arm drops and difference-in-differences are reported in separate columns. Do not conflate them.

The control arm fell 11.0 percentage points on secular trend alone. The honest attributable effect is the difference-in-differences (about 5?7 points for the arms that worked). Quoting the raw 16?18 point within-arm drop as the intervention effect is what every vendor does ? and it is the fastest way to lose this buyer.

Every intervention figure below that claims an effect uses difference-in-differences vs control, not the raw drop.

Arm Raw change DiD vs control Prov
--- --- --- ---
Control 24.1%->13.1% (-11.0) secular only measured
Control secular decline: -11.0 pp [meeker-2016]
Suggested alternatives 22.1%->6.1% (-16.0) -5.0 pp DiD measured
Source: meeker-2016 ? Meeker D et al. Effect of Behavioral Interventions on Inappropriate An
Accountable justification 23.2%->5.2% (-18.0) -7.0 pp DiD measured
Source: meeker-2016 ? Meeker D et al. Effect of Behavioral Interventions on Inappropriate An
Peer comparison 19.9%->3.7% (-16.2) -5.2 pp DiD measured
Source: meeker-2016 ? Meeker D et al. Effect of Behavioral Interventions on Inappropriate An

?2 The persistence result

Peer comparison also decayed after the interventions stopped ? but it was the sole arm still significantly below control twelve months later. Accountable justification produced the largest effect during the intervention and was statistically indistinguishable from control a year after it was removed.

Never claim the accountable-justification effect persists after removal. It does not. Never claim peer comparison persisted without noting that it also decayed.

Arm End?12mo post DiD vs control Prov
--- --- --- ---
Control 14.2%?11.8% (-2.4) secular only measured
Source: linder-2017 / linder-2017
Suggested alternatives 7.4%?8.8% (+1.4) +3.8 pp DiD measured
Source: linder-2017 / linder-2017 ? 95% CI -10.3?17.9 ? DiD p = 0.55
Accountable justification 6.1%?10.2% (+4.1) +6.5 pp DiD measured
Source: linder-2017 / linder-2017 ? 95% CI 4.2?8.8 ? DiD p = 0.001 ? vs control p = 0.99
Peer comparison 4.8%?6.3% (+1.5) +3.9 pp DiD measured
Source: linder-2017 / linder-2017 ? 95% CI 1.1?6.7 ? DiD p = 0.005 ? vs control p = 0.001

?3 The wider evidence

? Peer comparison also moves statin prescribing in primary care ? the mechanism is not antibiotic-specific. [measured]
Implication: Imaging ordering is a fair extension of the same behavioural channel; it has not yet been trialled inside a risk-bearing population with persistence measured after withdrawal.
persell-2016: https://pubmed.ncbi.nlm.nih.gov/27438316/

? The settled nudge taxonomy separates defaults, social norms, and justification ? they are not interchangeable. [modeled]
Implication: Suggested alternatives (defaults / order sets) are the arm that failed vs control. Building only that layer is implementing the null result.
thaler-sunstein: https://yalebooks.yale.edu/book/9780300122231/nudge/

? Clinicians' self-estimates of their own CT utilisation are systematically inaccurate relative to measured rates. [modeled]

Implication: Peer comparison works partly because it replaces a biased self-view with an external, risk-adjusted number.
ct-self-estimate: https://pubmed.ncbi.nlm.nih.gov/26864410/

? Status-quo bias makes "do nothing different" the default clinical decision even when evidence favours change. [measured]
Implication: A monthly packet that names a specific alternative pathway addresses the bias; a silent CDS alert does not.
samuelson-1988: https://link.springer.com/article/10.1007/BF00055564

? Small, frequent incentives ("peanut effect" / mental accounting) change behaviour differently from large lump-sum bonuses. [modeled]
Implication: Module 5 incentive design must respect mental accounting ? do not assume a year-end shared-savings cheque substitutes for monthly peer feedback.
kahneman-peanut: https://pubsonline.informs.org/doi/10.1287/mksc.4.3.199

? Defaults save lives in organ donation ? and they can also lock in low-value imaging pathways when the default order set is wrong. [measured]
Implication: Under-used defaults are a design opportunity; over-trusted defaults are how suggested-alternatives software ships the arm that did not work.
johnson-goldstein-2003: https://pubmed.ncbi.nlm.nih.gov/14631022/

? Audit and feedback produces a small-to-moderate, highly variable effect (Cochrane). [measured]
Implication: Quote this before a buyer finds it. Peer comparison is the arm that worked and, though it also decayed after removal, remained significantly below control at twelve months ? not a guarantee of a large effect in any given organisation.
ivers-cochrane: https://www.cochranelibrary.com/cdsr/doi/10.1002/14651858.CD000259.pub4/full

? Brehaut’s 15 suggestions: recommend a specific action the recipient controls; align with their own goals; deliver repeatedly; keep key messages short. [measured]
Implication: Encoded in the peer-packet schema (two gap cohorts, one named alternative, monthly cadence, outside the EHR).
brehaut-2016: https://pubmed.ncbi.nlm.nih.gov/26903136/

?4 What each finding implies for every CDS product ? including ARKA’s
---
Suggested alternatives had the large raw drop everyone quotes and was not significant vs control. That is the arm most imaging software ships. The Mechanism Ledger below is our own position, not a competitor scorecard.

? Accountable justification (implemented): attributable DiD -7 pp (measured). ARKA position: use the text for performance measurement, credentialing or compensation; store the text.
? Peer comparison (implemented): attributable DiD -5.2 pp (measured). ARKA position: rank clinicians; show a bottom-N list with names; attach personal dollar figures to a clinician-facing packet.
? Suggested alternatives (deliberately-not-implemented-as-such): attributable DiD -5 pp (measured). ARKA position: ship a thin draft labelled as a default; claim an effect for a suggestion that still needs clinician input.
? Defaults and pre-orders (implemented): no attributable DiD published for this mechanism in the four-arm table. ARKA position: emit a default that still leaves required fields blank; silently degrade a failed default into a suggestion.
? Goal setting (planned): no attributable DiD published for this mechanism in the four-arm table. ARKA position: pretend a planned mechanism is live; substitute education or alerts for a negotiated goal.
? Feedback (implemented): no attributable DiD published for this mechanism in the four-arm table. ARKA position: build a justification-analytics dashboard; show clinician-facing ranked lists.
? Information transparency (not-implemented): no attributable DiD published for this mechanism in the four-arm table. ARKA position: claim cost or radiation transparency as a shipped mechanism.
? Active choice (deliberately-not-implemented-as-such): no attributable DiD published for this mechanism in the four-arm table. ARKA position: replace free-text justification with a forced binary or taxonomy choice; gate the order on selecting an option.
? Alerts and reminders (deliberately-limited): no attributable DiD published for this mechanism in the four-arm table. ARKA position: fire above the published ceilings; treat an alert without a record write as accountable justification.
? Environmental cueing (not-implemented): no attributable DiD published for this mechanism

in the four-arm table. ARKA position: omit this mechanism from the ledger to look more complete than we are.
? Education (not-implemented): no attributable DiD published for this mechanism in the four-arm table. ARKA position: sell education as the intervention.
? Hard stops (will-never-implement): no attributable DiD published for this mechanism in the four-arm table. ARKA position: block, delay, or condition an order signature; require justification before the order can proceed.

?5 Our correction

The intervention we demonstrated to a former CMS Chief Medical Officer in August 2026 was suggested alternatives. It is the arm of the trial that did not work. We rebuilt. Accountable justification writes into the patient’s record and we verify the write; peer comparison arrives monthly outside the EHR; a default is executable with zero additional clinician input or we label it a suggestion. We publish the fire-rate ceilings and the measured rate. We do not sell the control arm.

Accountable justification works because the justification is written into the patient’s record. If it is not written, it is not accountable justification ? it is an alert. Our software checks that the write happened, tells the customer when it did not, and does not claim an effect for deployments where it could not.

Before / after ServiceRequest JSON from the default conversion (Vol III item 6). A suggestion that fails assertIsDefaultOrder is labelled a suggestion; a default is executable with zero additional clinician input.

BEFORE ? suggestion ServiceRequest (fails assertIsDefaultOrder):

```
{
"resourceType": "ServiceRequest",
"status": "draft",
"intent": "order",
"subject": {
"reference": "Patient/patient-1"
},
"code": {
"text": "X-ray - lumbar spine",
"coding": [
{
"display": "X-ray",
"code": "x-ray"
}
]
}
}
```


AFTER ? default ServiceRequest (passes assertIsDefaultOrder):

```
{
"resourceType": "ServiceRequest",
"id": "default-order-1",
"status": "draft",
"intent": "order",
"priority": "routine",
"subject": {
"reference": "Patient/p-1"
},
"requester": {
"reference": "Practitioner/prac-1"
},
"encounter": {
"reference": "Encounter/enc-1"
},
"occurrenceDateTime": "2026-08-09T15:00:00.000Z",
"code": {
"text": "X-ray chest",
"coding": [
{
```

"system": "http://www.ama-assn.org/go/cpt",
"code": "71046",
"display": "Radiologic examination, chest; 2 views"
}
]
},
"reasonCode": [
{
"coding": [
{
"system": "http://hl7.org/fhir/sid/icd-10-cm",
"code": "R05.9",
"display": "Cough, unspecified"
}
]
}
],
"supportingInfo": [
{
"reference": "https://arkahealth.com/evidence/chest-xr"
}
]
}

?6 What we will never build, and why

What we will never do

Every negative reaction a clinical audience has to this category is a fear about one of these. Each commitment is generated from the firewall that enforces it ? a named check, not a marketing promise.

? No leaderboard, no rank, no bottom-N.
Feared: You will put me on a leaderboard or shame the bottom quartile.
Enforced by: lint: CHECK-I3-RANK in scripts/lint-tcoc.ts ? no function may return clinicians sorted by rate to a clinician-facing surface

? No personally attributed dollar figure in a clinician-facing artefact.
Feared: You will put a dollar amount next to my name.
Enforced by: type: PeerPacket has no dollar field (lib/tcoc/peer/packet-types.ts); CHECK-I4-DOLLARS in scripts/lint-tcoc.ts ? PeerPacket cannot carry a personally attributed dollar field at the type level

? Justification text is never used for discipline, credentialing, or compensation.
Feared: Anything I type will be used against me in HR, credentialing, or bonus.
Enforced by: lint: CHECK-I10-JUSTIF in scripts/lint-tcoc.ts; CHECK-J5-JOIN / CHECK-J5-BRIDGE in scripts/lint-aj-firewall.ts; aj_ schema forbid ? justification text and its hash cannot be joined to peer, ledger, stats, group, or incentive aggregates

? ARKA can auto-approve; it can never auto-deny.
Feared: The computer will deny my patient’s scan.
Enforced by: runtime-guard: evaluateReviewerAuthorization in lib/ins/reviewer-queue.ts; lint:never-deny (CHECK-ND-1?CHECK-ND-8); lib/arka-ip/never-deny.ts; IRE decision-surface test ? adverse determinations require a human licensed-physician reviewer identity with a verified NPI; automation may only auto-approve or route to clinical review; ARKA-IP may recommend a second review and may never cancel, convert, hold or set the status of an order

? No hard stops, ever.
Feared: This will block me from signing when I need to order.
Enforced by: runtime-guard: CHECK-J1-BLOCKING / CHECK-J1-NO-BLOCK in scripts/lint-aj-firewall.ts; order-sign-while-AJ-500 acceptance test (J1) ? no AJ code path may prevent, delay past budget, or condition an order signature

? ARKA does not store justification text.
Feared: You will keep a copy of my clinical reasoning and mine it, lose it, or hand it over.
Enforced by: schema: CHECK-J6-NOTEXT in scripts/lint-aj-firewall.ts; npm run aj:verify-no-text; migration column allowlist ? no aj_ table may contain a column named text, justification, narrative, comment, or body; sentiment free text (Prompt 32.4) is a separate artefact with a published retentionDays purge and must not store clinical

justification

? No claim of effect where the mechanism is absent.
Feared: You will claim the trial effect even when the chart write is not working at my site.
Enforced by: type: JustificationReceipt.recordWrite required discriminated union; CHECK-J2-RECEIPT in scripts/lint-aj-firewall.ts; MECHANISM DEGRADED below MECHANISM_FLOOR ? a deployment whose written-share is below floor is rendered DEGRADED and may not have an effect claimed for it

? No fire above the published ceilings.
Feared: You will fire so often that I ignore every alert.
Enforced by: runtime-guard: FIRE_CEILINGS / resolveFireCeilings / effectiveCeilingValue in lib/aj/governor/spec.ts; CHECK-CEIL-1 / CHECK-CEIL-2; governor suppression + recorded ceiling events (J4); published on /evidence/mechanisms ? cards above the published day / 30-day / org-share ceilings (and any tighter cohort override) are suppressed and recorded, never silently dropped

? No site may raise a fire-rate ceiling in exchange for a commercial concession.
Feared: You will trade my interruption budget for a discount or a deal term.
Enforced by: lint: CHECK-CEIL-1 / CHECK-CEIL-2 in lib/aj/governor/spec.ts and scripts/lint-aj-firewall.ts; Clause 2 in docs/arka-aj/DEPLOYMENT_AGREEMENT_CLAUSES.md; Math.min against CEILING_MAXIMA ? no configuration path ? including a future item-35 site config ? can raise a clinician above the global maximum or loosen an in-force ceiling for commercial reasons

? We will not touch your ordering behaviour until your LEAD benchmark is closed.
Feared: You will cut imaging volume in 2026 and I will carry a lower benchmark until 2036.
Enforced by: lint: CHECK-LEAD-MTI in scripts/lint-tcoc.ts; Clause 5 in docs/arka-aj/DEPLOYMENT_AGREEMENT_CLAUSES.md ? ARKA will not intervene on ordering until LEAD Performance Year 1 (1 January 2027); the commitment is a deployment-agreement term, and the lint proves the published /lead copy and the clause cannot drift

? ARKA does not compute, influence, or recommend changes to beneficiary attribution or to risk-adjustment coding. Attribution and risk scores are read-only inputs supplied by your organisation. We measure avoided utilisation against the attribution and risk model you give us, and we publish the specification of how we use them.
Feared: This is a coding product wearing a quality costume, and it will put our Shared Savings or LEAD settlement under audit.
Enforced by: lint: CHECK-ATTR-1 in scripts/lint-tcoc.ts and the attribution scope rule in scripts/lint-scope-boundary.ts; Clause 6 in docs/arka-aj/DEPLOYMENT_AGREEMENT_CLAUSES.md ? No ARKA code path outside the ingest boundary writes to an attribution assignment, an HCC code, a risk score or a coding field. The build fails if one is added, and the deployment agreement makes the commitment contractual.

? ARKA measures which clinician’s ordering behaviour changed against a locked pre-period, and exports that measurement. ARKA never uses, alters or influences the assignment of beneficiaries to clinicians, and never derives a contribution figure from a field that determines attribution or risk score. The two are computed from disjoint inputs and the build proves it.
Feared: They are going to use our compensation feed to reverse-engineer attribution, or to reward clinicians for coding.
Enforced by: lint: CHECK-INC-1 and CHECK-INC-2 in scripts/lint-tcoc.ts; the disjoint-input proof in docs/attribution-firewall.json; Clause 8 in docs/arka-aj/DEPLOYMENT_AGREEMENT_CLAUSES.md ? No contribution figure may be derived from a beneficiary-attribution assignment, HCC code, risk score or coding field. The two field sets are a computed set-difference in docs/attribution-firewall.json, and the build fails if they overlap or if an identifier under lib/tcoc/incentive conflates the two.

? Where ARKA writes a care-setting flag or a ’considered and deferred’ note back to the record, it is documented as a clinical decision. ARKA never suggests, implies or enables a documentation change whose effect is to increase a risk score or a billing level.
Feared: The write-back feature is a coding-intensity tool with a clinical label on it.
Enforced by: lint: CHECK-ATTR-2 in scripts/lint-tcoc.ts ? No write-back template or user-facing Module 3 copy may suggest, imply or enable a documentation change whose effect is to increase a risk score or a billing level. The build fails if uplift language is added.

? ARKA never sets, suggests, or ratchets a group utilisation target. Medical directors choose the goal; ARKA measures against it.

Feared: This is a utilisation-pressure tool dressed as quality improvement ? you will set the target and make me explain it to my board.
Enforced by: lint: CHECK-GOAL-1 and CHECK-GOAL-2 in scripts/lint-tcoc.ts ? setGroupGoal is only callable from an authenticated medical-director handler; lib/tcoc/group has no numeric ratePer1k default; the goal surface forbids recommended-range copy; no code path derives a new goal from prior performance

? A payer-authored or utilisation-management source may inform a coverage determination and may never inform clinical appropriateness. The two field sets are disjoint and the build proves it.
Feared: Your clinical score is the payer's criteria with a white coat on.
Enforced by: lint: CHECK-APPR-1 in scripts/lint-tcoc.ts; CHECK-FW-1..3 in lint:sources / lint:coverage / lint:scope; the disjoint-input proof in docs/appropriateness-firewall.json; Clause 9 in docs/arka-aj/DEPLOYMENT_AGREEMENT_CLAUSES.md ? No payer-authored or utilisation-management source may inform a clinical appropriateness rating. Verdict A (score.ts and the knowledge matrix) and Verdict B (lib/coverage) are computed from disjoint field sets in docs/appropriateness-firewall.json; the clinical engine cannot reach lib/coverage, lib/pa, lib/davinci or lib/ins even transitively; no ModalityRating carries a payer_um / payer_utilization anchor.

? Every evidence-panel decision is published, including adverse decisions.
Feared: You will bury the ratings you lowered and only publish the upgrades.
Enforced by: lint: CHECK-CHARTER-PUBLISH in scripts/lint-tcoc.ts; docs/governance/EVIDENCE_PANEL_CHARTER.md publication commitment ? the published charter requires every panel decision ? including returned, rejected, recused, and certainty-lowered rows ? to be published; an unpublished override is void

? The chair and any co-chair of the evidence panel are entirely free of conflict.
Feared: The panel chair works for you, or owns a piece of the company.
Enforced by: lint: CHECK-CHARTER-CHAIR in scripts/lint-tcoc.ts; IOM 2011 chair rule in docs/governance/EVIDENCE_PANEL_CHARTER.md ? the published charter adopts IOM 2011 by name and requires the chair and any co-chair to be entirely free of conflict; no person with an equity interest in ARKA may hold a seat

? No ARKA surface, artefact, metric, export, contract clause or piece of marketing may reference evaluation-and-management coding, billing level, code selection or reimbursement in connection with the considered-and-deferred note ? and the build fails if one does.
Feared: The write-back is a coding-intensity tool, and you will teach my clinicians to document for level.
Enforced by: lint: CHECK-CF-1 in scripts/lint-coding-firewall.ts; docs/coding-firewall.json computed disjointness; Clause 10 in docs/arka-aj/DEPLOYMENT_AGREEMENT_CLAUSES.md ? No write-back-touching module or copy may carry banned coding lexicon (CPT, E/M, level of service, MDM, data/risk element, complexity, documentation improvement, capture, uplift, reimbursement); clinical-note and billing-relevant field sets are a computed set-difference

? Whether a considered-and-deferred note was written may not enter the attribution ledger, the peer-comparison packet, the contribution export, or any incentive computation. The write-back module graph is unreachable from attribution, ledger and incentive, and the build proves it.
Feared: Clinicians who write the note will look better on the measure than clinicians who do not ? you built a documentation incentive by accident.
Enforced by: lint: CHECK-CF-2 / CHECK-CF-3 in scripts/lint-scope-boundary.ts; CHECK-CF-4 / CHECK-CF-5 in scripts/lint-coding-firewall.ts and scripts/lint-tcoc.ts; Clause 10 in docs/arka-aj/DEPLOYMENT_AGREEMENT_CLAUSES.md ? No module under lib/tcoc/ may import lib/ehr/considered-deferred; no write-back field appears in peer / contribution / ledger / incentive schemas; no published figure is derived from note counts; attribution/ledger/incentive import graphs never reach write-back modules

? ARKA never assigns a less favourable appropriateness rating to an order because of the care setting in which it was placed.
Feared: You will down-score my telehealth orders because you assume I cannot examine the patient.
Enforced by: lint: CHECK-CS-3 in lib/aiie/care-setting.ts and scripts/lint-tcoc.ts; docs/governance/VIRTUAL_ENCOUNTER_POSITION.md ? a setting-specific rating less favourable than the base rating without a cited clinical reason fails the build; virtual encounters are a review dimension, not a default penalty

? ARKA never recommends an in-person evaluation in place of a study that the indication

supports.
Feared: You will use 'come in for an exam' as a deflection when imaging is indicated, especially on telehealth orders.
Enforced by: lint: CHECK-IPE-SPEND-1 in lib/cards/ipe-spend-firewall.ts and scripts/lint-tcoc.ts; buildInPersonEvaluationCard refuses when indicationSupportsStudy is true ? an in_person_evaluation card cannot be emitted when the indication supports the ordered study; IPE is never counted as avoided spend

---

Figures (every figure carries its qualifier):
? At Duke, the top decile of primary-care providers ordered 4.2? more imaging than the bottom decile ? before anyone intervened.: 4.2? [measured] ? OBSERVATIONAL ? NO CONTROL GROUP ? GRADE: low ? halpern-2021
? JGIM volume result: year-one median imaging rate ?17.3% (within-network pre-post) [measured] ? OBSERVATIONAL ? NO CONTROL GROUP ? GRADE LOW ? halpern-2021
? JACR cost result: Median estimated radiology costs ?19.5%; >$3M saved in year one (within-network pre-post) [measured] ? OBSERVATIONAL ? NO CONTROL GROUP ? GRADE LOW ? clark-randall-2021
? Control arm secular decline (raw within-arm drop): 11.0 percentage points [measured] ? raw within-arm drop; secular trend, not an intervention effect ? meeker-2016
? Suggested alternatives difference-in-differences vs control: -5.0 percentage points [measured] ? difference-in-differences vs control (not the raw drop) (DiD) ? meeker-2016
? Accountable justification difference-in-differences vs control: -7.0 percentage points [measured] ? difference-in-differences vs control (not the raw drop) (DiD) ? meeker-2016
? Peer comparison difference-in-differences vs control: -5.2 percentage points [measured] ? difference-in-differences vs control (not the raw drop) (DiD) ? meeker-2016

Cite this document

Documents that tell you how to cite them get cited. Suggested citation (v1.1, 16 August 2026, engine tcoc-1.0.3, generated August 25, 2026):

ARKA Health. Three ways to change what a physician orders. Only two of them work. And what happens when you try the same thing on imaging. What we know about imaging; What we know about the mechanism; How the studies compare; What nobody knows; The study we intend to run (v1.1, 16 August 2026, engine tcoc-1.0.3).
https://arkahealth.com/api/evidence/nudge-writeup. Generated August 25, 2026. Retrieved August 25, 2026.

? end of write-up ?