NBER WORKING PAPER SERIES

VOLUNTARY ACCREDITATION AND HEALTHCARE QUALITY: EVIDENCE FROM A RANDOMIZED CONTROLLED TRIAL IN U.S. JAILS

Marcella Alsan
Crystal Yang

Working Paper 33357
http://www.nber.org/papers/w33357

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
January 2025, Revised May 2026

We gratefully acknowledge funding from JPAL North America, the William T. Grant Foundation, the Preparedness and Treatment Equity Coalition, Arnold Ventures, and the Malcolm Wiener Center at the Harvard Kennedy School. We thank seminar participants at NBER Law and Economics, NBER Crime, NBER Economics of Health, Harvard Law School, Harvard Kennedy School, Vancouver School of Economics, UCLA Anderson, MIT Economics, Minnesota Department of Management and Budget, Boston University Law School, Princeton, Stanford, MIT Organizational Economics and Dartmouth. We thank the National Commission on Correctional Health Care (NCCHC), especially Chief Executive Officer Deborah Ross, Vice President of Accreditation Amy Panagopoulos, and Vice President of Program Development Rich Forbus. We also thank jail custody and health staff, incarcerated individuals, and county sheriffs for their time and participation. We thank Jim Martin, Meghan Beal, Dr. Cara Buchanan Muñoz, George Hutchins, and Dr. Nicholos Joseph for sharing their expertise. This work would not have been possible without the effort of our dedicated project coordinators: Tasfia Jahangir and Nadia Huffman. Emily Deng, Alessandro Perga, Miguel Purroy, Qing Liu, Shipra Karan, Xingyou Ye, Vanessa Jia, Nada Shalash, Nick Shankar, and Lukas Leister provided excellent research assistance. Finally, we thank the numerous college, law, public policy, and medical students who worked on the project: James Jolin, Lucy Tu, Gwen Byrne, Aisha Oshilaja, Pierre Jolin, Josh Alpert, Daniel Tochimani-Hernandez, Farrah Kalach, Janine Robredo, Sophia Fend, Bolin Shen, Yedith Ortiz, Emily Shah, Susan Takang, Kate De Groote, Worthy Cho, Israa Alzami, Anne Deng, Christine Smith, Aydin Alsan, Achele Agada, Cana Nakase, Sarah Kliment, Kathryn Li, Kaitlyn Ham, Tatheer Adnan, Anita Onyimah, Folukemi Olufidipe, Meisui Liu, Phoebe Canagarajah, Lea Han, Allie Oh, Louis Lin, Jennifer Ding, Zach Channen, Nik Zappia, Fiona Doyle, Christian Carson, Drew Lane, Jack Haney, Haley Dougherty, Ashil Jhaveri, Abigail Sassoon-Dondero, and Isaac Harris. Angelique Aquatella, Caroline Hoxby, and Karine Van Der Straeten provided helpful feedback. The experiment was registered on the AEA RCT registry (RCT ID: AEARCTR-0007516). The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2025 by Marcella Alsan and Crystal Yang. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Voluntary Accreditation and Healthcare Quality: Evidence from a Randomized Controlled Trial in U.S. Jails

Marcella Alsan and Crystal Yang

NBER Working Paper No. 33357
January 2025, Revised May 2026
JEL No. I1, I14, I18, I3, K10, K14

ABSTRACT

We study whether voluntary accreditation improves healthcare quality in U.S. jails, where publicly-financed care is often delivered by private contractors and overseen by elected sheriffs with limited medical expertise. We conduct the first randomized controlled trial of healthcare accreditation in the U.S., randomizing the offer of accreditation to 46 jails. Assignment to accreditation improves compliance with quality standards, particularly in training and patient care, without increasing capital and labor inputs. It also reduces 12-month mortality by approximately 60 percent, with declines concentrated in jails that experienced the greatest quality improvements. Evidence suggests effects operate through improved coordination between staff, timely screening, and increased information for oversight. These findings highlight the potential for accreditation to mitigate incentive and information frictions in contracted public services.

Marcella Alsan
Stanford University
Department of Economics
and NBER
malsan@stanford.edu

Crystal Yang
Harvard University
Harvard Law School
and NBER
cyang@law.harvard.edu

A randomized controlled trials registry entry is available at
https://www.socialscienceregistry.org/trials/7516

I Introduction

Serving as the primary point of entry into the U.S. criminal justice system, jails process roughly 7 million admissions annually and detain over 600,000 individuals on any given day, the majority of whom are unconvicted (Zeng 2023). Beyond detention, jails function as a critical component of the U.S. healthcare system, providing initial screening and care to individuals who are often in crisis or acutely ill. Incarcerated individuals have a constitutional right to health care, yet the quality of care in correctional settings remains a substantial and growing concern: most jails rely on privatized healthcare provision, and mortality rates have increased over the past two decades (Carson 2021).

One commonly used mechanism to assure quality in the provision of publicly-financed services in the U.S. is voluntary accreditation. Although accreditation is typically modeled as quality signaling for uninformed consumers, its use in health and education sectors is commonplace but understudied. Broadly defined, service accreditation comprises three major components: development and dynamic review of evidence-based standards, training on the implementation of such standards, and compliance assessment. Thus, the overall accreditation process, by combining the positive effects of inspections with educational and managerial components, has the potential to yield broad organizational changes and quality improvements (Martinez et al. 2015; Gibbons and Henderson 2013).

As a practical matter, the process of service accreditation is handled by third-party agencies who develop proprietary standards, and these agencies are typically compensated by the entities they evaluate. In education, for example, the federal government designates "reliable authorities" to accredit postsecondary schools, thus providing institutional eligibility for federal financial aid (Martin 1994). A similar process operates in the U.S. civilian healthcare system, where accreditation is most commonly conducted by The Joint Commission (TJC); hospitals and many other acute care settings must obtain "deemed status" through TJC in order to be eligible for federal reimbursement from the Centers for Medicare and Medicaid Services (CMS).[1] As a result, nearly all U.S. patients receive care from accredited hospitals.

Importantly, this institutional structure differs sharply from correctional healthcare. Unlike hospitals and nursing homes, correctional facilities generally do not receive CMS reimbursement and are therefore not subject to the same federal oversight or accreditation-linked financial incentives. Instead, care is financed and administered at the state or local level, often through contracts with private vendors – a structure that has been linked to higher mortality (Bedard and Frech 2009).

Despite its widespread use, evidence on the effectiveness of accreditation in improving healthcare quality remains mixed and largely non-causal, according to recent systematic reviews (Hussein et al. 2021; Alhawajreh, Paterson and Jackson 2023). Our project addresses this gap in the literature and makes several contributions. First, we provide the first causal evidence on whether accreditation in the U.S. improves quality in healthcare. The need for mechanisms to ensure quality healthcare for incarcerated populations is critical. Drug- or alcohol-related deaths in jails have increased by a factor of four from 2000 to 2019, and the official suicide death rate in jails in 2019 was twice the national average (Carson 2021). As wards of the county or state during confinement, incarcerated individuals depend entirely on correctional institutions

[1] Every year a small fraction of accredited hospitals are also inspected by state agencies on behalf of CMS. For historical reasons, nursing homes do not generally undergo TJC accreditation; instead they are inspected annually by state agencies on behalf of CMS (Association of Health Care Journalists n.d.; Nursing Home Law Center 2025; Morford 1988). An important recent literature has shown these inspections are important in improving short-run mortality benefits (Chen and Dillender 2026; Gandhi, Olenski and Shi 2025) though this may reflect in part effort shifting rather than organizational change (Cheng 2023).

for medical care, and since the Supreme Court's landmark decision in *Estelle v. Gamble* (*1976*), correctional facilities have been constitutionally obligated to provide it.

Second, to the best of our knowledge, this is the first study to conduct a randomized controlled trial across U.S. jails and to experimentally manipulate jail healthcare systems. Legal and logistical barriers have historically limited rigorous evaluation of inmate populations. Over several years, we randomized the offer of healthcare accreditation across 46 jails, a setting characterized by weak external monitoring and non-expert principals. Jails are typically administered by county sheriffs with limited medical expertise, and care is largely delivered through contracts with for-profit vendors under capitated payment structures, which may weaken incentives for quality provision relative to cost containment. These features make the implementation of large-scale experimental research in this setting challenging but important.

Third, we test the impact of accreditation on both process-based measures of quality and mortality. Our data indicate that there is considerable scope for improvement along both margins within U.S. jails, where multiple frictions exist and where health outcomes are much worse than widely believed and officially reported. For example, a report from the Senate Permanent Subcommittee on Investigations (2022) found that the Department of Justice failed to identify hundreds of deaths in custody and those reported frequently omitted key information. These patterns reflect longstanding challenges with accurate and transparent reporting of deaths behind bars (Mitchell Jr. and Aronson 2023; Grant Smith 2020). Our study data comport with the Senate's findings, with mortality rates in the 2022-2024 period being three times higher than those officially reported in the most recently released statistics (Carson 2021).[2]

To carry out the study, we recruited a representative sample of small to medium size jails from across the U.S. Jails were randomized using a phased implementation design, such that 24 facilities were randomized to initiate accreditation immediately and received a subsidy to do so, while 22 were assigned to begin the process approximately 24 months after enrollment with a smaller subsidy. Accreditation was administered by the National Commission on Correctional Health Care (NCCHC), a nonprofit accrediting body for correctional facilities. NCCHC evaluates facilities across seven core domains, ranging from Governance and Administration to Patient Care and Treatment.

To assess compliance, NCCHC conducts on-site inspections and issues a report, after which a committee of experts – blind to facility identity – determines accreditation status. To maintain accreditation, NCCHC also requires jails to provide yearly evidence of compliance and to be re-inspected every three years. Accreditation is attached to the jail itself, rather than its healthcare vendor. Among practitioners and policymakers, NCCHC accreditation is often regarded as a "gold standard" and, in some cases, as evidence of constitutionally adequate care, although critics argue that it may be ineffective and even counterproductive.

In order to independently measure quality, we collected data at baseline and endline using multiple instruments. These include surveys of facility leadership and custody and medical staff, independent audits of administrative records (including meeting notes, protocols, death logs, and patient medical charts), and qualitative interviews with jail leadership and incarcerated individuals. These data were then coded by teams with clinical and legal training to ensure consistent interpretation. In addition, we fielded an expert survey of jail administrators, healthcare providers, policymakers, academics, and formerly incarcerated individuals, eliciting beliefs about the importance of different standards as well as predictions about average effects and heterogeneity.

[2]Following Carson (2021), we calculate the mortality rate by dividing deaths for the control facilities by their ADP and multiplying by 100,000.

By the end of the study, 13 of the 24 treatment facilities achieved accreditation, while the remaining treated facilities completed varying portions of the process. Assignment to the NCCHC accreditation process improves compliance with composite measures of quality standards by 8 to 13 percent, driven primarily by improvements in standards governing Personnel Training (15 percent) and Patient Care and Treatment (12 percent). Treatment assignment reduces the 12-month mortality rate at endline by over 60 percent, with declines concentrated in jails that experienced the greatest quality improvements. These effects are precisely estimated, robust across specifications, and observed in two independent data sources, limiting concerns about measurement error or reporting bias. These magnitudes likely reflect the fact that, unlike typical inpatient settings serving chronically ill elderly patients, jail populations are younger, and deaths more often arise from acute, potentially preventable conditions.

To examine the mechanisms through which accreditation improves quality and health outcomes, we consider a principal–agent framework with team production (Alchian and Demsetz 1972; Holmstrom 1982). In jail settings, healthcare delivery requires joint effort by custody and medical staff and is overseen by a principal with limited medical expertise. Accreditation introduces standardized protocols, expands documentation, and clarifies responsibilities, thereby acting as a coordinating device for quality-assured practices. In doing so, it reduces the scope for non-contractible quality and facilitates more effective monitoring. Consistent with this mechanism, we document improved coordination between custody and medical staff, more systematic screening and documentation at intake, and reductions in mortality. We find no evidence of structural changes in inputs such as staffing or capital investment, suggesting that accreditation operates primarily through organizational and informational channels rather than increased resources.

Our project contributes to three major literatures. First, we contribute to the theoretical and empirical literature on quality assurance mechanisms. Existing evidence largely focuses on nursing home inspections or disclosure and certification of consumer goods. These effects depend on market structure and the extent to which consumers can respond (Dranove and Jin 2010; Chen and Dillender 2026; Gandhi, Olenski and Shi 2025). Teacher certification has been well-studied (see Clotfelter, Ladd and Vigdor (2010); Kane, Rockoff and Staiger (2008); Goldhaber and Anthony (2007)), but causal evidence on school accreditation is relatively thin. We depart from the existing literature by studying accreditation in a setting without consumer choice, where publicly-financed services are provided by private vendors.

Second, we contribute to a growing literature on managerial practices and organizational culture.[3] Prior work has shown that management interventions can generate large productivity gains, in part by reducing informational barriers and improving organizational processes. We extend this literature by examining the impact of managerial interventions in a high-stakes environment – publicly-owned correctional institutions – where management challenges are amplified by information frictions including moral hazard in teams.

Third, we contribute to a growing literature in economics and public health on health and healthcare in carceral settings.[4] For example, the public health literature generally reports harmful effects of jails, including higher rates of overdose deaths immediately upon release (Binswanger et al. 2007) and higher rates of deaths due to carceral exposure over the life course (NASEM 2019; Daza, Palloni and Jones 2020). More recently, Norris, Pecenco and Weaver (2024), using an event study framework, find detrimental effects of release from incarceration in Ohio on mortality. The authors interpret this as evidence of high mortality risk

[3]See Atkin et al. (2017); Bloom et al. (2013, 2020); Bruhn, Karlan and Schoar (2018); Giorcelli (2019); Gosnell, List and Metcalfe (2020); Iacovone, Maloney and McKenzie (2022); Martinez et al. (2015) and Bianchi and Giorcelli (2022).

[4]See Batistich, Evans and Phillips (2021); Patterson (2010); Spaulding et al. (2011); Wildeman et al. (2016); Norris, Pecenco and Weaver (2024); Binswanger et al. (2007); Smith (2013); Bovell-Ammon et al. (2021) and Ruch et al. (2021).

among community-dwelling men that is lowered by the incapacitation effects of incarceration. Other work suggests that incarceration can improve health for some groups in settings with higher-quality care, such as Scandinavian prisons (Hjalmarsson and Lindquist 2022; Bhuller, Khoury and Løken 2025). We contribute experimental evidence on whether improving healthcare quality inside U.S. jails affects both care processes and mortality.

The remainder of the paper is organized as follows. Section II provides information on accreditation and the specific complications associated with delivering healthcare behind bars. Section III provides details on the intervention and experimental design. Section IV describes data collection and coding, Section V describes our empirical strategy, and Section VI presents our main findings. Section VII analyzes potential mechanisms including qualitative evidence from staff and inmate interviews. Section VIII concludes.

II Background

This section provides background on two areas: voluntary accreditation in healthcare and unique features of U.S. jails that make it difficult to deliver high quality healthcare.

II.1 Voluntary Accreditation

Voluntary Accreditation in Healthcare: Accreditation is a process designed to ensure quality of organizations in settings of limited information.[5] Typically, the accreditation process involves the development of standards and an independent, on-site audit by a third-party (Sechrest 1976). The theoretical effects of accreditation on quality and associated outcomes are ambiguous and depend on whether the third-party is an honest broker and whether the standards are evidence-based, points we return to below (Dranove and Jin 2010).

Empirical findings on the effects of accreditation on quality in healthcare, where it is incredibly common, are mixed and causal evidence is generally lacking (Hussein et al. 2021; Alhawajreh, Paterson and Jackson 2023). Despite the lack of compelling evidence on effectiveness, voluntary accreditation is a prerequisite for receipt of reimbursement from CMS in acute-care settings (Jha 2018; The Joint Commission 2024).[6]

Voluntary Accreditation in Correctional Health Care: There are several key differences between healthcare provision for the civilian and incarcerated populations that make the carceral setting theoretically distinct and empirically important to investigate. First, instead of voluntary accreditation for private facilities overseen by executives trained in management in the context of civilian healthcare, accreditation is for public facilities managed by officials who may have limited domain knowledge or managerial expertise when it comes to healthcare. This is certainly the case for elected sheriffs who are often responsible for healthcare provision in their jails. In this setting, accreditation may increase and improve the information available to sheriffs that allow them to better monitor their staff.

[5] Compared to other quality assurance mechanisms such as warranties or licensing, quality disclosure systematically measures and disseminates information about product quality, is typically conducted by a third-party certifier, and relies on standardized quality standards so that results are comparable across producers (Dranove and Jin 2010).

[6] The Joint Commission (TJC), which accredits most U.S. hospitals, briefly attempted to accredit correctional health care facilities but ultimately discontinued the effort after recognizing that correctional health care was fundamentally different from traditional hospital settings (Chassin 2021).

Second, the public sector often contracts out services to for-profit entities.[7] Over the past few decades, correctional healthcare has been struggling with cost increases due to prisoner aging, increased mental health needs, and medical inflation.[8] Yet unlike the civilian healthcare system, jails do not receive federal funding: the 1965 Medicaid Inmate Exclusion Policy (MIEP) prohibits Medicaid from reimbursing most health care delivered to the incarcerated (Edmonds 2020).[9] In the case of prisons, Bedard and Frech (2009) find that privatized services have increased mortality – suggesting that the private provider is responding to capitated payment structures by reducing non-contractible quality (Hart, Shleifer and Vishny 1997). Thus, forcing the contracted healthcare provider to participate in the accreditation process could plausibly reduce harmful cost-cutting.[10]

Another point of distinction is that the delivery of healthcare in jails is tightly intertwined with custody operations. Clinical providers cannot examine, counsel, or treat patients at their discretion, and high-quality care may be viewed as conflicting with security priorities. This creates scope for low effort on both sides. Our qualitative interviews with incarcerated individuals highlight that custody staff do not always implement medical recommendations, as "officers are more concerned about the disciplinary and correctional aspects than whether they are receiving the proper care." At the same time, under capitated contracts, medical providers are dissuaded from providing costly interventions. Together, these frictions can undermine the delivery of high-quality healthcare.

Further, quality signaling to a consumer shopping for healthcare is not a motivation in the jail setting. Instead, findings from our pre-experimental survey of nearly 200 jails suggest that the main motivation for jail leadership to pursue voluntary healthcare accreditation is concern about liability and meeting constitutional obligations, with another substantial fraction valuing input from medical experts.[11] These obligations stem from the 1976 landmark decision *Estelle v. Gamble*, where the Supreme Court held that "deliberate indifference" to the serious medical needs of incarcerated people violates the Eighth Amendment's prohibition against cruel and unusual punishment (American Medical Association 1979; Greifinger 2007).[12]

Appendix B formalizes the potential effects of accreditation. The setup is a principal (the sheriff) who seeks to minimize excess (greater than expected) inmate mortality in a setting of team-based moral hazard. There are two agents – custody and medical staff – who must work together to reduce the "aggregate output" of inmate deaths (Alchian and Demsetz 1972; Holmstrom 1982). As we show in this model, in a setting where effort to improve quality is unobservable and contracts with vendors are incomplete, both medical and custody staff are incentivized to exert low effort, to the detriment of inmate health outcomes. By providing

[7] Comprehensive data are difficult to obtain, but reports suggest approximately 60-70 percent of jails contract with private healthcare vendors, including private-equity backed vendors Wellpath and Corizon, as well as Naphcare (Szep et al. 2020; McLeod 2019; Coll 2019).

[8] Currently, corrections represent the largest expenditure for state budgets after Medicaid, with one-sixth of prison budgets dedicated to health care (Pew Charitable Trusts 2014; Williams, DiTomas and Pachynski 2021).

[9] In recent years, several states have received Section 1115 Waivers which allow Medicaid funds to be used for a limited set of health care delivered to the incarcerated prior to release (Saloner 2023; Howell et al. 2023).

[10] This concern is echoed in a recent *Guardian* article, where an attorney notes that "when private healthcare companies receive a fixed fee per person, their financial incentive is to spend less on the care they provide to incarcerated patients; the impact on quality is predictable" (Green 2025).

[11] Accreditation could potentially reduce liability insurance premiums, with some accreditation agencies advertising that the receipt of accreditation can reduce premiums by up to 10 percent (see https://blog.guardianrfid.com/blog/why-you-should-care-about-the-aca-accreditation-process). Conversely, our own conversations with jail underwriters suggest that insurance premiums for off-site/catastrophic inmate medical insurance might increase in response to accreditation if medical claims rise because facilities may spend more on inmate healthcare.

[12] Further rulings expanded on this decision to include prevention and extended liability to contracted medical providers. See Alsan et al. (2023) for further details. Despite facing potential liability, these private vendors have in practice limited their financial exposure through bankruptcy reorganizations.

evidence-based standards and enabling monitoring of agent behavior, a non-colluding accrediting body acts as a public signal that facilitates coordination across agents. It also improves attribution, allowing the principal to target punishment for shirking. As a result, accreditation can sustain a high-effort equilibrium and improve inmate health outcomes.

NCCHC Accreditation: In order for accreditation to improve quality, it has to be conducted by a non-colluding third-party agency that carefully crafts and then enforces evidence-based standards. In the space of corrections, there are a number of agencies that provide accrediting services. Some of these agencies have been the subject of investigations stemming from accusations that they fail on both criteria.[13]

Our study randomizes jails to accreditation by NCCHC, a non-profit company of about 40 full-time employees based out of Chicago, IL (ProPublica 2026). Founded in 1983, the NCCHC grew out of a joint effort between legal and medical organizations to develop standards that would help jails comply with the *Estelle* ruling (Gibson and Phillips 2016).

NCCHC standards are often referred to as the "gold standard" in correctional healthcare. Leading public health scholars have advocated for mandating NCCHC accreditation for all correctional facilities (Berwick, Beckman and Gondi 2021). Courts also rely on NCCHC accreditation as a signal of adequate health care that complies with the Eighth Amendment. For example, in *Balla, et al. v. Idaho State Board of Correction*, a federal judge noted that "The NCCHC accreditation..., while not determinative, constitute substantial evidence of adequate medical care."[14]

Today, NCCHC administers consensus-driven and evidence-based proprietary standards covering 59 topics grouped into seven major categories: (1) Governance and Administration; (2) Health Promotion, Safety, and Disease Prevention; (3) Personnel and Training; (4) Ancillary Health Care Services; (5) Patient Care and Treatment; (6) Special Needs and Services; and (7) Medical-Legal Issues. For example, Governance and Administration standards address the establishment of a healthcare system that ensures access to care, professional administration, and monitoring and quality improvement policies. As another example, Patient Care and Treatment standards are designed to ensure the delivery of healthcare from booking upon arrival through discharge in a timely and evidence-based manner. Each standard is categorized, numbered and named, and is followed by a short description about why it is important and how it fits into overall correctional healthcare. For each standard, there are a series of compliance indicators that must be met to receive accreditation.

NCCHC emphasizes that the standards can be met by "jails of any size" and that it is not a membership organization. Rather, it is an independent entity that receives input and support from a variety of professional organizations to develop its standards including the American Medical Association (AMA), the American Bar Association, the American Nurses Association, and the National Sheriffs' Association. Specific examples of standards within each category can be found in Table 1. Each standard is also classified as either Essential or Important (see Appendix Table A1 for additional examples). Assessing compliance with standards is a multi-stage process that often requires jails to resubmit materials and in some instances undergo multiple on-site audits.

[13]The American Correctional Association (ACA) has been roundly criticized by policymakers, including Senator Elizabeth Warren, who reported that "it appears the ACA is a conflicted party with twisted incentives, a lack of transparency, and lax inspection policies that appear to have turned accreditation into a rubber-stamp process that does little to hold facilities accountable" (The Office of Elizabeth Warren 2019).

[14]Balla v. Idaho State Bd. of Corrections, 595 F. Supp. 1558 (D. Idaho 1984).

The NCCHC accreditation process emphasizes the importance of establishing protocols that meet clinical standards and requiring staff to document that such protocols are being followed. Such documentation may take the form of screening checklists and records, medical administration records, quality improvement meetings and minutes, logs of both sick call visits and deaths, and autopsy reviews. These protocols and documentation practices expand the set of observable and verifiable actions within correctional health care. They help jail administrators – such as sheriffs – better understand what care should be delivered and, ex post, attribute responsibility when adverse outcomes occur.

The centrality of collaboration between custody and medical staff is also directly addressed in various NCCHC standards (National Commission on Correctional Health Care 2021). For example, the NCCHC standard on medical autonomy states that "delivery of health care in a correctional facility is a joint effort of custody and health staff and is best achieved through trust and cooperation." Another NCCHC standard requires health training for correctional officers so that they can "recognize the need to refer an inmate to a qualified health care professional."

Critics of accreditation argue that its voluntary nature, and the fact that facilities must pay to be accredited, undermine its effectiveness. They contend that this structure weakens oversight incentives, pointing to instances of preventable deaths in accredited facilities as evidence in their favor.

II.2 Jail Healthcare Quality and Barriers to Improvement

Scale and Health Burden in U.S. Jails: Despite disagreement on how to improve quality of healthcare in corrections, few would dispute the view that preventable deaths in U.S. jails warrant intervention. The scale of the system is substantial: in 2022, U.S. jails recorded approximately 7.3 million admissions, compared to roughly 34 million hospital admissions (Zeng 2023; Health Forum 2024). On any given day, more than 600,000 individuals are detained in jails (Zeng 2023), a roughly similar number to the average daily census for U.S. hospitals in 2019 (Health Forum 2021). Jails are therefore a critical point of intervention, serving as the entry point to the criminal justice system and the setting where many individuals are first screened and triaged upon booking.

Individuals who enter the jail system often suffer from a dual burden of chronic and infectious disease (Puglisi and Wang 2021). The incarcerated population generally comes from lower-income households, with substantially greater physical and mental health needs than the general population; they have higher rates of tuberculosis, HIV, hepatitis, diabetes, and psychiatric illness (Maner et al. 2022; Curran et al. 2023; Puglisi and Wang 2021). Incarcerated individuals experience mental health problems at rates five times higher than the general population, as well as elevated rates of geriatric conditions compared to the age-matched general population (Bronson and Berzofsky 2017; Greene et al. 2018).

The occurrence of adverse events is commonplace and rising in jails. For example, mortality rates have increased by 11 percent from 2000 to 2019, with a *reported* average mortality rate in U.S. jails of 1.5 deaths per 1,000 and a mortality rate of almost 2 deaths per 1,000 among the unconvicted jail population in 2019 (Carson 2021; Adler and Chen 2023). During the same time period, drug- or alcohol-related deaths in jail increased by a factor of four, and the average suicide death rate remained high at two times the national average (Wang 2021; Carson 2021).

Barriers to Quality Improvement: Several barriers impede improvements in healthcare quality in U.S. jails. First, federal financing is virtually nonexistent and financial incentives are weak. Because incarcerated indi-

viduals are excluded from receiving federal benefits under the MIEP, jail healthcare systems are not subject to the same high-powered reimbursement incentives or statutory mandates that drive quality improvement and accreditation in other healthcare settings.

Second, legal accountability is constrained. Although incarcerated individuals have a constitutional right to adequate medical care, the Prison Litigation Reform Act (PLRA) of 1996 imposes substantial procedural hurdles to filing suit, including strict exhaustion requirements, short filing windows, and hefty financial costs (Alsan et al. 2023). These barriers reduce the likelihood that poor-quality care is challenged through litigation.

Third, transparency in mortality reporting is limited. Despite reauthorization of the Deaths in Custody Reporting Act (DCRA) in 2013, reporting remains incomplete and inconsistent (McCann 2024; Congressional Research Service 2023). Penalties for noncompliance – up to a 10 percent reduction in Edward Byrne Memorial Justice Assistance Grant (JAG) funding – are modest and applied at the state rather than local level. In addition, responsibility for data collection was recently shifted from the Bureau of Justice Statistics to the Bureau of Justice Assistance, which is less well suited to statistical reporting. Even when data are released, they contain substantial deficiencies.[15] Consistent with these concerns, mortality rates calculated from our own facility-level death logs substantially exceed official statistics published by the Department of Justice.

Taken together, the limited financial incentives, weak legal accountability, and lack of transparency reduce external pressure to improve healthcare quality in jails. In this context, we test whether voluntary accreditation could meaningfully improve the quality of care and health outcomes for the incarcerated.

III Intervention and Experimental Design

We randomized 46 jails to undergo NCCHC accreditation. Figure 1 depicts the NCCHC assessment process which often requires multiple stages of evaluation and corrective action. After a facility initially applies to begin the accreditation process, the NCCHC sends the jail the proprietary standards with a self-survey questionnaire. The jails use these materials to begin improvements and consult with NCCHC staff on how to implement changes. NCCHC then conducts an on-site visit, during which health administrators, nurses and physicians review custody and medical records and interview jail staff and patients. Findings from this on-site visit are incorporated into a report that is then presented to the Accreditation Committee. This Committee consists of physicians and nurses who are affiliated with professional societies (*e.g.,* AMA), industry (*e.g.,* Naphcare, health care executives) and government (*e.g.,* local county jails). The Committee is blinded to the facility's identity and determines compliance with standards based on the report.

To become accredited, a facility must be compliant with 100 percent of applicable Essential Standards and at least 85 percent of applicable Important Standards. If a facility is compliant with more than 90 percent but less than 100 percent of applicable Essential Standards and at least 85 percent of applicable Important Standards, the facility receives a status of "continuing accreditation with verification (CAV)." CAV status requires a jail to undergo corrective action. If corrective action is completed within the required four month time frame and verified by the Committee, the jail then receives full accreditation. Other facilities who are farther from the compliance thresholds are placed on "provisional accreditation" or "probation," which

[15]The GAO identified approximately 1,000 unreported deaths in 2021 relative to alternative data sources and found that roughly 70 percent of submitted records were missing key information (GAO 2022).

automatically requires another on-site visit (*i.e.*, focus survey). Jails that require these visits are then re-reviewed by the Committee. Facilities that do not submit corrective action in a timely manner, or fail to resolve compliance findings, are placed on probation and those that remain non-compliant are ultimately denied accreditation. At the end of our study, nine of 24 treated facilities did not successfully attain accreditation and two facilities attrited.

III.1 Study Timeline, Recruitment, and Randomization

Figure 2 shows our study timeline, with the study commencing in early 2020 and ending in late 2025. Baseline measurement occurred in late 2021 through 2023 and endline measurement from 2023 to 2025.[16] Additional details on study design can be found in Appendix C.

Recruitment and Enrollment: The study began with the recruitment of jails at national meetings and conferences targeting sheriffs and jail administrators. A member of our research team attended either in person or virtually. For in-person meetings, team members set up vendor booths with informational flyers and business cards for our project coordinators.

Sixty jails expressed initial interest in the study. We conducted an informational session with custody and/or medical leadership from each interested jail. During the informational session, we shared additional details about the study and information on our Certificate of Confidentiality from the National Institutes of Health (NIH) which we obtained to further protect the information provided by participants.

Jails were eligible for our study if they had an ADP between 100 and 3,000 (midsized jails, which represent approximately one-third of all jails) and had not previously received accreditation by the NCCHC (Zeng and Minton 2021). Privately operated jails, ICE facilities, and juvenile detention facilities were excluded from the study. Out of the 60 interested facilities, 12 declined to move forward following the informational session, one did not meet our study eligibility criteria, and one was not eligible for accreditation based on existing NCCHC standards. Non-participating and participating facilities are similar in observable characteristics, including mortality (Appendix Table A2).

To formally enroll in the study, we required written and informed consent signed by a representative of both custody and medical staff. Following enrollment, eligible jails received our baseline facility and staff survey. Completion of the facility survey was incentivized by $500, which could be used towards the cost of accreditation. Staff surveys were incentivized by a lottery of winning a $100 e-gift card.

All enrolled facilities were given the opportunity to become accredited at a highly subsidized rate, with the control facilities beginning the NCCHC accreditation process at the end of the data collection period. Accreditation fees are based on jail ADP. The subsidy for the treated group was much larger and available immediately compared to the control group, with a difference of about $5000 (Appendix Table A3).

Randomization: Once jails completed the baseline facility survey, they were randomized into the treatment ("Accredit Now") or control ("Accredit Later") arm and informed of their assignment. Stratified randomization was conducted based on ADP and timing of enrollment ("batch"). Given that larger facilities might have more resources and/or more health challenges, we stratified on high versus low ADP to improve statistical power (Athey and Imbens 2017). Because there were capacity constraints in the accreditation agency, we also randomized in three batches.

[16]The timing of both baseline and endline measurement occurred after the peak of the COVID-19 pandemic. We also stratify treatment assignment by both jail size and calendar time of enrollment, as described below.

Our initial recruitment goal was to recruit 40 facilities with a wait list for additional interested facilities in the event of attrition per our pre-analysis plan. We invoked randomization off the wait list after two facilities randomized to treatment attrited: in one instance the consenting party suffered a heart attack, and the other was put on leave for sexual assault. We randomized six additional facilities from the wait list, yielding a total of 24 facilities allocated to treatment (including the two attriting facilities) and 22 to control.

NCCHC Accreditation Starts: After randomization, NCCHC began its accreditation process depicted in Figure 1. NCCHC did not deviate from its usual process for the purposes of our study. Once the treatment facilities received an initial accreditation decision from NCCHC, both control and treatment groups completed incentivized endline facility and staff surveys, as well as medical audits and inmate and leadership interviews. These survey instruments are described in Section IV.1 below.

III.2 Representativeness of Sample

Our recruitment efforts yielded a sample of jails from all over the United States, including regions where incarceration rates are highest (**?**). Figure 3 illustrates the U.S. states where the 46 facilities in our sample are located, with each circle representing the total ADP across all participating facilities in each state.[17]

To assess the representativeness of our study sample, we use data from the 2019 Bureau of Justice's Census of Jails (U.S. Department of Justice 2022). The Census is designed to capture all jails although reporting is not mandatory. Comparing the 42 out of 46 jails in our sample who responded to the 2019 Census of Jails against 1,085 other medium-sized (ADP 100-3,000) county jails serving adult populations, we find that our sample is representative of these other jails in terms of jail characteristics (Table 2 Panel (A)) and characteristics of the incarcerated population (Panel (B)), although jails in our study had slightly higher yearly admissions with shorter average lengths of stay and a lower share of White incarcerated individuals. Approximately 4 percent of our sample and a similar percent of out-of-sample U.S. jails were under a consent decree for conditions of confinement, and about 67 percent of all inmates at in-sample and out-of-sample jails were convicted of felony offenses.

While broadly representative of medium-sized jails in the U.S., the jails in our sample are likely negatively selected relative to other jails that voluntarily seek NCCHC accreditation, consistent with Einav et al. (2022). Anecdotal information from NCCHC prior to the experiment was that most jails that applied for accreditation eventually received it, although a small set of extremely negatively selected facilities were court-ordered to go through the process as part of a consent decree or monitoring process (Douds and Ahlin 2016).

By contrast, our recruitment efforts revealed many jails had never heard of NCCHC. Among our sample of 24 treated facilities, only 13 successfully received accreditation. All 22 non-attriting treated facilities received NCCHC standards and completed the self-assessment. Among those not accredited, most underwent an on-site visit but either declined to implement further changes recommended by the Accreditation Committee or were denied accreditation despite follow-up efforts. Although the jails often had little prior information about NCCHC before their recruitment, each of the 13 treatment facilities that received accreditation publicized their status on jail websites or social media accounts.

[17] Information on the exact identity of each county is protected pursuant to our data use agreements.

IV Data Collection, Coding, and Integrity

This section describes our survey instruments and administrative data, coding methodology, and pre-specified construction of indices to address multiple hypothesis testing concerns. We also discuss steps taken to ensure the accuracy of results to nonresponse, attrition, and demand bias. Details on the data collection, along with links to all study instruments, are provided in Appendix D. Details on variable coding and index construction are provided in Appendix E.

IV.1 Survey Instruments and Administrative Data

Our data collection strategy included: (1) facility survey, (2) medical audit, (3) anonymous staff survey, and (4) qualitative interviews with incarcerated individuals and facility leadership. We also conducted an expert survey to obtain weights on the importance of each quality standard and baseline features of facilities predicted to impact the effect of accreditation.

Our research team analyzed the NCCHC quality standards to determine whether and to what extent each standard could be objectively assessed, as well as the best manner to obtain accurate information. For example, NCCHC's Privacy of Care standard requires that health care encounters and exchanges of information are conducted in private (*e.g.*, private area, privacy screen, curtain). To measure whether a facility complies with this standard, our facility survey required the facility to upload photo documentation of their patient care areas (see Appendix Figure D9). As another example, NCCHC's Receiving Screening standard requires medical personnel to inquire about illness histories, allergies, recent communicable symptoms, prescription medications, drug use, and other conditions shortly upon intake. In our medical audit, our physician surveyor visually inspected randomly selected medical documentation of receiving screenings in order to assess compliance. Our anonymous staff survey also enabled us to collect systematic information on staff sentiment and cross-validate the other metrics.

We describe each of these data collection elements below:

(*1*) *Facility Survey*: The facility survey was administered at both baseline and endline via a secure Qualtrics link for each facility. The facility survey had two major components: a questionnaire that was answered by the facility administrator, and document uploads of requested procedures, protocols, and minutes from specified meetings. In the questionnaire component which included 62 question items, facilities were asked about a range of topics related to the quality standards, such as the availability of specific physical equipment and medical supplies, the availability of mental health and dental services, the FTE of each type of medical staff, and how incarcerated individuals were screened upon booking in general and for specific conditions. The upload component of the survey required facilities to provide a copy of their official polices and procedures manual, along with 43 other requested documents (including photo documentation). These included documentation on management of COVID-19 and protocols for Medication Assisted Treatment (MAT), treatment protocols for patients with conditions such as hypertension, diabetes, and HIV, as well as CPR and other health certification for all personnel. In addition, we requested blank copies of health assessment forms, mental health screenings, sick call forms, and death logs, among others. Finally, we requested health staff meeting and administrative meeting minutes and attendance records from the past 12 months.

(*2*) *Medical Audit*: The medical audit was administered at endline with a retrospective look-back to the

study baseline period. The audits were performed on Zoom with an independent physician who was blind to the treatment status of the jail alongside a similarly blinded scribe, and took approximately three hours per facility. For the audit, each jail was asked to pull a random set of medical records for two patients covering both the endline and baseline periods, along with two records per condition and per period for suicidality, opioid use disorder, and hypertension. In addition, jails provided information from their death logs (specifically, the month, year, and cause of death for individuals who died while in custody), clinical death reviews, sick calls, and psychological autopsy reviews for the past 12 months. Due to Health Insurance Portability and Accountability Act (HIPAA) protections, the research team was prohibited from collecting personally identifiable information (PII) for the deceased, such as name, demographic information, or exact dates.[18] All records (whether electronic or paper) were shown to the physician virtually, with PII redacted, and were recorded by the scribe.

(*3*) *Staff Survey*: The staff survey was also administered at both baseline and endline via a secure Qualtrics link for each respondent. The survey was sent to all personnel including both custody and medical staff. The survey included approximately 120 questions related to staff satisfaction, the relationship between custody and medical staff, and perceptions of care. For example, we asked respondents whether they agreed with statements like "Custody staff support the implementation of clinical decisions and partner with clinical staff," "Custody and clinical team staff work collaboratively to give good health care to inmates," and "There is open communication between custody and health staff," which gauge the degree of "horizontal coordination" between custody and medical staff. To measure "vertical coordination" between staff and supervisors, we also asked respondents whether they agree that "Staff feel comfortable speaking up when they see something that may negatively affect inmate care," and whether that feedback is acted upon by the supervisor; questions were asked of both staff and supervisors.

(*4*) *Interviews with Incarcerated and Leadership:* We also conducted qualitative interviews with incarcerated individuals and facility leadership at endline only. For each facility, we requested interviews with three incarcerated individuals (at least two of whom had used medical services in the last month) and with one senior health services administrator and one senior custody administrator. These interviews were conducted on Zoom, took approximately 45 minutes each, and were conducted in private with the study coordinator and a scribe. These interviews were considered optional and the study coordinator was not blinded to the treatment condition, unlike the medical audit. Incarcerated individuals were asked questions in both multiple choice and open response format regarding their perceptions of the care they received. Administrators were also asked questions about their perceptions of health care delivery.

(*5*) *Expert Survey:* From October 2022 to March 2024, we collected incentivized predictions on our study from 145 experts across various domains, including jail administrators and health care providers, policy/academic experts, and the formerly incarcerated. Recruitment occurred at national subject matter conferences, through cold calls using Safety Source's National Directory on Law Enforcement Administrators, and through recruitment methods on Prolific (Safety Source 2021). We use expert weights to construct an aggregate quality standards index, as described below. Respondents were asked to rank order quality standards in terms of perceived importance. We also adopted a principled approach to assessing heterogeneity by relying on predictions from this survey: respondents were asked to select five factors that they believed would most affect

[18]The HIPAA Privacy Rule applies to the individually identifiable health information of a decedent for 50 years following the date of death of the individual.

whether accreditation of jail health care services by NCCHC has an effect on outcomes, including healthcare quality and inmate mortality. We use these predictions in our heterogeneity analysis as described below. See Appendix D.6 for characteristics of the expert survey participants.

IV.2 Coding, Construction, and Measurement of Main Outcomes

Our two main outcomes are (1) compliance with quality standards and (2) inmate health outcomes as measured by mortality. As described above, we specifically asked domain experts to predict the impact of accreditation on these outcomes in our expert survey.

Per our published pre-analysis plan, we did not initially anticipate being able to obtain administrative data on deaths, given privacy restrictions. We planned to use death logs as a measure of quality in reporting practices: what fields were available and how completely those fields were filled out. However, we quickly realized that we could obtain standardized death information from the logs in a manner consistent with HIPAA requirements. Specifically, scribes copied the administrative record of the count and cause of deaths. We updated our pre-analysis plan in summer 2024, prior to analyzing these data, to include mortality as an outcome. We also use qualitative interviews and staff surveys to examine mechanisms and related outcomes.

Measuring Quality: To measure compliance with quality standards, we used data collected from the facility survey and medical audit, which were mapped into each of the seven NCCHC quality standard categories. We used a detailed coding process to independently and objectively measure compliance on the basis of the documents uploaded by facilities in the facility survey. Specifically, a Coding Committee consisting of Harvard students and chaired by the study coordinator developed a coding manual to map raw data elements into a dataset. The manual consists of 43 separate spreadsheet pages corresponding to the major compliance indicators within each of the seven quality standards. Each sheet provided a step-by-step guide for coders to determine whether the facility was complying with the standard. As one example, the current NCCHC standards require facilities to have a grievance policy for incarcerated individuals that provides clear guidance on days to first response, formal documentation of grievances, and an appeals process. The coding manual instructed our coders to read uploaded protocols to determine whether such criteria were met, as can be seen in Figure 4.

As an alternative to the seven NCCHC quality standards, we developed and pre-specified a production function approach to the categorization of outcomes. This approach groups outcomes based on five categories of health production: (1) Inputs (*e.g.*, staff FTE, staff qualifications, equipment), (2) Formal Procedures (*e.g.*, screening, emergency response plan, grievance policy), (3) Administrative Procedural Outcomes (e.g., administrative meetings, annual reviews, custody staff training), (4) Health Procedural Outcomes (e.g., screening schedule, medication administration, sick call schedule), and (5) Medical Outcomes (*e.g.*, continuation of medication for MAT, treatment plan for suicidal ideation).

For each jail, the coding process was performed by two independent college or law students who were blinded to the treatment assignment. Students were only eligible for coding after undergoing an online training session and practice exam which included intentional errors to ensure their accuracy and attention to detail. If there was disagreement between the two coders as to whether a facility complied with a standard, a third independent student (again blinded to treatment assignment) would adjudicate. Another Coding Committee consisting of Harvard medical students, the study coordinator, and one of the PIs mapped questions in the medical audit to each of the seven quality standard categories.

At the end of the coding process, we created an index for each of the seven NCCHC standards categories (and each of the five production function categories). All elements in each index ranged from 0 to 1, with either binary indicators to measure compliance with the particular part of a standard, or a fraction to indicate percent compliance when that step required multiple elements. We construct the index for each quality standard as the arithmetic average across all such indicators associated with that standard. Thus, each index can be interpreted as the aggregate percent compliance with the standard, with 0 indicating complete non-compliance and 1 indicating full compliance.

Measuring Mortality: Our primary measure of mortality comes from death logs obtained directly from jails. These data cover all deaths occurring within twelve months prior to our medical audit.[19] Death logs were obtained for all 46 facilities either during the medical audit or later by request.

In addition, we collected data on deaths in custody from media reports for all 46 facilities. This served two purposes: (i) to provide an independent measure of mortality, and (ii) to construct a baseline measure, as we did not have IRB approval to obtain administrative death log data beyond a twelve-month window. Administrative death logs generally record more deaths than media reports, as they capture events not covered by the media.

IV.3 Multiple Hypothesis Testing - Creation of Pre-Specified Indices

To address concerns with multiple hypothesis testing, we pre-specified constructing a meta-index by aggregating all the outcomes used to construct the individual indices of compliance with the NCCHC quality standards (and the five indices for each of the production function categories). We do this in three pre-specified ways. First, we aggregate by computing the simple average across the seven indices corresponding to the NCCHC quality standards. Second, we use a principal component analysis with all outcomes (224 questions in total), using the first component to create a meta-index. One caveat is that the weights with which the outcomes are aggregated to components are not necessarily positive. We rescale the first principal component so it lies in the [0,1] interval (Lise and Postel-Vinay 2020). Finally, we calculate a weighted average across the seven quality standard indices, where weights are determined by asking experts (health care, corrections, policy experts, and previously incarcerated individuals, detailed above in Section IV.1) the importance of different outcome categories.

The reciprocal rank weights are presented in Figure 5 and show that with respect to the NCCHC quality standards (Panel (A)), experts typically rank standards on Patient Care and Treatment as most important, with less weight on Governance and Medical-Legal Standards. With respect to the production function categories (Panel (B)), experts typically rank Health Procedural Outcomes and Medical Outcomes as most important, with less weight on Administrative Procedural Outcomes.[20]

[19] This time period covers the post-randomization period for all facilities except two. For these facilities, two and three months of data prior to randomization were included, respectively.

[20] Specifically, experts ranked standards from most important ($r_i = 1$) to least important ($r_i = 7$) for NCCHC quality standards (or $r_i = 5$ for production function categories) and these ranks are transformed to weights w_i as follows: $w_i = \frac{\frac{1}{r_i}}{\sum_{j=1}^{n} \frac{1}{r_j}}$. With respect to the NCCHC quality standards, the formerly incarcerated (lived experience experts) place more weight on Personnel Training whereas the academics, administrators and health care providers (professional experts) place more weight on Safety and Prevention. Our main aggregation index combines both sources of expertise, averaging weights computed for each of the four types of experts. Further details on our aggregation methods are provided in Appendix E.8.

IV.4 Measurement Integrity

Our survey instruments rely on documentation and information provided by facilities that were informed of their treatment assignment, raising the possibility of demand bias or Hawthorne effects. We take several steps to ensure the integrity of the data we collected.

As described above, we designed our survey instruments to target quality standards that could be objectively measured and verified by the research team. For example, we measure compliance with quality standards using components from the facility survey including both facility "self-reported" answers and independent coding of uploaded documents by students, as well as data from the medical audit. Figure 6 Panel (A) demonstrates the composition of the seven NCCHC indices by data source, illustrating substantial variation in the extent to which a given standard relies on each source. Notably, quality standards on Safety and Prevention and Patient Care and Treatment utilize a high share of data from the medical audit, which relied on direct observation and inspection by our medical personnel and scribe, and rely minimally on facility self-reports. Thus, we view the measurement of these standards as objective and difficult to falsify. We also intentionally chose to exclude staff survey responses in our measurement of compliance with quality standards. Instead, we use the staff survey to create separate metrics of staff sentiment.

Second, we do not use any of the data NCCHC collected in its on-site visit or final report to the Accreditation Committee as our outcomes. All data used in this study to determine whether facilities met standards were collected and coded independently by the Harvard research team. Figure 6 Panel (B) compares our study's compliance score for all Essential standards with NCCHC's compliance score.[21] Recall that a facility must meet 100 percent compliance with Essential Standards in order to receive NCCHC accreditation. While the two measures are highly correlated among treatment facilities at 0.71 (p-value = 0.0004), Harvard compliance scores are substantially lower on average, with average scores of 0.58 and 0.74 for the study and NCCHC, respectively.

Lastly, we verify the accuracy of our mortality measure given potential concerns about underreporting of deaths. As described above, we cross-check each death reported in the death logs with media reports using our PII-compliant information. In general, we find that the death logs are more complete than media reports, although the correlation coefficient between the number of endline deaths as reported in death logs and media is 0.82.

IV.5 Missing Data

Data for Quality: As discussed above, our first primary outcome of compliance with quality standards is constructed using data obtained from the facility survey and medical audit. Data from the baseline and endline facility survey are non-missing for all 46 jails in our study. Ten (including the two attriting jails) out of 46 facilities refused to complete the medical audit, however. Importantly, we find that accreditation assignment is not correlated with completing this medical audit (see Appendix Table A4). In our primary analysis, we penalize for non-response on the medical audit, coding non-responses as failure to comply with the required standard as per our pre-analysis plan. Further, in Section VI.1 below, we show that our findings are robust to measuring compliance with quality standards using *only* the facility survey data and that our estimates are statistically similar (albeit noisier) if we exclude the ten facilities with missing medical audits.

[21] This comparison is only feasible for treatment facilities because NCCHC does not compute scores for control facilities as they do not interact with them until study completion.

Data for Mortality: For the second primary outcome of mortality, there are no missing data. Through both death logs obtained at the time of the medical audit, as well as those obtained through subsequent data collection even for jails that refused to complete a medical audit, our administrative death log data are complete for all 46 jails. In addition, we use in-jail deaths from media reports which are available for all facilities.

Data for Secondary Analyses: Staff surveys and qualitative interviews with inmates and jail leadership were optional instruments in our study. As a result, 5 out of 46 facilities are missing endline staff survey responses and 14 out of 46 are missing qualitative interviews. Accreditation assignment is not correlated with completing these optional surveys (see Appendix Table A4). We used these sources of data as secondary outcomes in support of mechanisms in Section VII.

IV.6 Summary Statistics and Balance

Table 3 presents baseline means of key variables and assesses for balance between treatment and control facilities in our sample. Panel (A) reports baseline compliance with NCCHC quality standards as measured by meta-indices, and Panel (B) reports baseline mortality as measured by the number of total deaths in the 12 months prior to randomization divided by the facility's baseline ADP per 1,000. Panel (C) reports baseline jail characteristics. The first column reports the overall sample mean, the second column represents the coefficient on the treatment assignment indicator, and the third column reports the associated standard error, adjusting for randomization strata fixed effects.

Panel (A) of Table 3 shows that at baseline, jails in our study meet 48-56 percent of NCCHC quality standards on average. Importantly, there are no significant differences between treatment and control facilities on any of the quality standard meta-indices.[22] Panel (B) shows that the mean mortality rate at baseline is 2.16 deaths per 1,000 individuals, with balance by treatment status. Finally, Panel (C) demonstrates balance on baseline jail characteristics such as number of medical staff (FTE), whether the jail utilizes a for-profit health vendor (65 percent), whether the jail was previously accredited by a non-NCCHC organization (10.9 percent), and political affiliation of the county and sheriff. Joint F-tests for all baseline characteristics and each meta-index yield p-values of approximately 0.37.

Figure 7 presents additional descriptive evidence on the deaths among control jails using the administrative death log data. Panel (A) presents the distribution of deaths by cause of death. The most common cause of death in jails is undetermined (43 percent) followed by suicide/homicide (28 percent), illness (20 percent), and drug overdose (9 percent). While some of these undetermined deaths may still be actively under investigation, our independent review of media reports associated with these deaths indicates that many undetermined deaths involved situations alleging homicide, medical emergencies, and drug overdose. The prominence of deaths due to illnesses is concerning given that the average age of incarcerated individuals in the sample is 37.[23] Panel (B) of Figure 7 demonstrates a marked seasonality of deaths in jails – with a spike in the third quarter when 35 percent of all deaths occur. These patterns are consistent with lack of air conditioning and other heat-mitigation systems that may worsen chronic conditions and/or contribute to acute dehydration and heat stroke (Skarha et al. 2022, 2023; Cloud et al. 2023; Tuholske et al. 2024; Cowan,

[22] We also find that there are no significant differences between treatment and control facilities when using individual baseline NCCHC quality standards as well.

[23] This is the mean age for inmates in the 34 facilities for which we obtained jail booking records that contain defendant demographic information.

Lemasters and Brinkley Rubinstein 2024) and increase violence while incarcerated (Mukherjee and Sanders 2021).

V Empirical Strategy

Following our pre-analysis plan, we estimate the reduced form effect of assignment to treatment (initiation of the NCCHC accreditation process) of the form:

$$Y_j = \alpha_0 + \alpha_1 Treat_j + X_j\Gamma + \mu_{s(j)} + \epsilon_j \tag{1}$$

where j indexes facility. Y_j comprise our key pre-specified outcome of compliance with quality standards under both NCCHC and production function approaches. For secondary staff-level outcomes, Y_{ij} is indexed by individual i and facility j.

$Treat_j$ is an indicator for being randomly assigned to the treatment group and $\mu_{s(j)}$ represent strata fixed effects. Baseline measures of the relevant outcome, X_j, are included as pre-specified. Heteroskedasticity-robust standard errors are used throughout for facility-level outcomes and we cluster standard errors at the facility level for individual-level staff outcomes. We also compute *p*-values using randomization inference (RI) obtained by permuting accreditation assignments for each facility 1,000 times.

For mortality, the statistical properties of the outcome motivate a different analytical approach. Unlike the quality condition measures, which rely on a scale-invariant threshold applied uniformly across facilities, the occurrence of fatal events scales with exposure (i.e., facility size). Although we stratify by size, this approach does not fully account for the increased risk associated with greater exposure time in larger facilities, characterized by higher average daily populations (ADP) and longer lengths of stay (Adler and Chen 2023). Accordingly, our preferred specification is a Poisson model, as shown in Equation 2:

$$D_j = P_j \exp(\alpha_0 + \alpha_1 Treat_j + \mu_{s(j)}), \tag{2}$$

where D_j is the count of deaths in jail j over the relevant 12-month window and log P_j is an offset for person-time in custody, proxied by the facility's ADP. We report average marginal effects of treatment, which can be interpreted as the average change in the expected death count attributable to accreditation assignment, along with incidence-rate ratios (IRRs), corresponding to the average treatment effect in levels expressed as a percentage of the control mean. Both approaches yield consistent, point-identified estimates even when many observations equal zero (see Chen and Roth (2023) and Mullahy and Norton (2024)).

We also estimate linear models following Equation 1 where Y_j is either a rate (i.e., D_j divided by ADP multiplied by 1,000) or a threshold metric similar to the quality standards (i.e., $1(D_j \leq \text{x})$). For the rate outcome, we present ordinary and weighted least squares specifications with baseline ADP as analytic weights. The former provides the average effect across equally-weighted facilities, while the latter approximates the average experience of an inmate by placing greater weight on facilities that contribute more person-time exposure. Alternatively, we use a threshold metric where $Y_j = 1(D_j <= 1)$ over the 12-month endline period as one death is the median number of deaths among control jails.

VI Results

We now describe our reduced form estimates of assignment to accreditation on our primary outcomes of quality standards and mortality.

Effects on Quality Measures. Figure 8 presents the reduced form effects of assignment to treatment on the seven NCCHC quality standards in Panel (A) and the meta-indices described above in Panel (B). Appendix Table A5 presents the corresponding point estimates in table form.

Across all seven NCCHC categories, treatment assignment is associated with improvements in quality. Notably, the standards for which the improvements reach statistical significance at the 10 percent level are standards governing Personnel Training and Patient Care and Treatment. With respect to Personnel Training, for example, assignment to accreditation causes a 8.3 percentage point increase in compliance, a 15 percent increase from the control mean. For Patient Care and Treatment, treatment assignment causes a 6.6 percentage point increase in quality, a 12 percent increase relative to the control mean.

Moving to the composite indices in Panel (B) of Figure 8, we find that treatment assignment yields statistically significant and medically meaningful increases across all three meta-indices. Although both treatment and control groups improved in two out of three meta-indices relative to the baseline averages, assignment to accreditation causes an increase in aggregate quality compliance of 4.5 percentage points using a simple mean (8 percent), 7.4 percentage points using the principal component analysis (13 percent), and 4.9 percentage points using expert weights (9 percent) relative to the control group mean.[24] Results are also qualitatively similar using weights constructed separately for the formerly incarcerated – experts with lived experience – versus all other experts (see Appendix Table A6).

Appendix Table A7 demonstrates similar findings using the alternative production function approach. Consistent with our findings using the NCCHC quality standards, we find that treatment increases compliance with outcomes governing Health Procedural Outcomes. Recall that this category of outcomes includes items like the screening schedule, medication administration, and sick call schedule. Specifically, assignment to accreditation increases this Health Procedural Outcomes by 6.0 percentage points, a 12 percent increase from the control mean.

While not statistically significant at the 10 percent level, treatment also improves outcomes governing Formal Procedures and Medical Outcomes. For example, treatment increases medical outcomes, including continuation of medication and treatment for suicidal ideation, by 6.4 percentage points, a 19 percent increase from the control mean (p-value = 0.27).

Effects on Mortality. Figure 9 presents the mortality rate distribution for treatment and control groups using administrative death logs. There is a pronounced difference between the two groups and a leftward shift in the jails assigned to accreditation. To assess whether these results are driven by differences at baseline, we report both baseline and endline mortality rates and counts from media reports in Appendix Figure A1. Panels (A) and (C) of Appendix Figure A1 demonstrate that treatment and control jails have similar baseline mortality rates and counts prior to randomization, consistent with the balance presented in Table 3. In contrast, endline mortality rates and counts using media reports (Panels (B) and (D)) show a pronounced leftward shift for treatment jails, consistent with the patterns using administrative death logs.

[24] Control group improvement could be consistent with all recruited jails undergoing an informational session during recruitment.

Table 4 reports reduced form estimates of the effect of accreditation assignment on mortality using count and linear models. Data are from administrative death logs from all 46 jails. All specifications include randomization strata fixed effects, and odd columns also include a baseline control. The outcome varies across specification, accordingly.

Columns (1) and (2) report average marginal effects of -1.333 (s.e. = 0.438) and -1.246 (s.e. = 0.495) estimated using Equation 2. These effects correspond to an approximately 63 to 65 percent reduction in mortality relative to the control mean, calculated as $1 - \text{IRR}$. Columns (3) and (4) show that accreditation assignment increases the probability that a jail meets the one or fewer death threshold by approximately 28 percentage points (s.e. = 0.141) and 26 percentage points (s.e. 0.142), respectively, estimated using Equation 1. The last four columns present findings for deaths per 1,000 ADP estimated using Equation 1. In columns (5) and (6), treatment effects are negative, but estimates are smaller and lack precision. Weighting by ADP in columns (7) and (8), however, yields a larger and more precisely-estimated reduction of 3.5 to 3.6 deaths per 1,000 inmates, significant at the 5 percent level. Relative to a control mean of 5.1 deaths per 1,000, this corresponds to a decline in mortality of roughly two-thirds, closely aligned with the magnitude from the count model. We find similar results using deaths as reported in the media (see Appendix Table A8).

We also explore the effect of assignment to treatment on cause of death using administrative death logs. Across both count and linear specifications (Appendix Tables A9 and A10), we find reductions in mortality attributable to suicide and homicide, drug overdose, illness, and undetermined causes, the latter two being the largest and most significant.

Taken together, we find that accreditation assignment substantially reduces mortality in jails, with reductions on the order of 60 percent over a twelve month period. In Appendix Figure A2, we show that our preferred Poisson estimates are qualitatively similar when dropping each facility one at a time. The results are also robust to different functional forms and alternative data sources.[25] The similarity of estimates across both media- and facility-reported deaths mitigates concerns that the findings are driven by experimenter demand.

Heterogeneity. We explore heterogeneity along dimensions identified in our expert survey, including medical staff FTE per ADP, baseline quality of care, and management quality (Appendix Figure A3). Appendix Figure A4 reports interactions with treatment for endline quality and mortality. We find limited evidence of heterogeneity overall along baseline dimensions. In contrast, we find strong heterogeneity along endline quality of care. Appendix Table A11 presents reduced form estimates of the effect of accreditation assignment on mortality using our preferred count model. Column (1) presents our main average marginal effects and column (2) presents average marginal effects on both accreditation assignment and also an interaction between assignment and endline quality of care (as measured by an indicator for above median quality of care, constructed using our expert-weighted meta-index). These findings indicate that mortality reductions are driven by treatment facilities with higher endline quality, consistent with the view that quality improvements are a key mechanism for improving health outcomes.

[25] Our mortality results using facility-reported deaths are robust to using the generalized ridge regression estimator from Kwon and Sun (2025), regulaTE, which constructs valid confidence intervals under treatment effect heterogeneity across strata even when one stratum lacks overlap between treated and control facilities. Using bounds on the degree of treatment effect heterogeneity across strata, OLS specifications yield an average treatment effect of 3.7 fewer deaths per 1,000 inmates (s.e. = 0.861) and statistical significance at the 5% level, consistent with the corresponding OLS estimate in column (8) of Table 4. Similar results are obtained when running regulaTE on media-reported deaths.

VI.1 Robustness

We evaluate three potential sources of selection that could affect the interpretation of our findings.

Robustness to Missing Data: As discussed previously in Section IV.2, our measure of compliance with quality standards is constructed using data from the facility survey and medical audits. While we have facility survey data for all 46 jails in our sample, ten facilities did not complete medical audits. In our primary analysis above, we penalize their missing responses, coding non-response as failure to comply with the corresponding standards. To assess the sensitivity of our results to missing data, in Appendix Table A12, we construct the seven NCCHC quality standards (Panel (A)) and three meta-indices (Panel (B)) using *only* the facility survey, dropping the medical audit from the construction of the quality outcomes. In Panel (A), we continue to find that treatment assignment leads to positive improvements in compliance across most domains, although precision varies. We find statistically significant improvements again for Personnel Training and Patient Care and Treatment, with effect sizes of 8.3 percentage points (15 percent relative to control mean) and 4.1 percentage points (7 percent relative to control mean) respectively. Panel (B) aggregates the seven standards into the aforementioned three composite indices. Across all variations, overall compliance improves by approximately 5.0 to 6.3 percentage points, which is very similar to our main estimates presented in Figure 8 and Appendix Table A5.[26]

Inmate Selection: Treatment versus control jails could select different individuals to book at endline, changing the composition of those incarcerated. In Appendix Table A13, however, we find no differences in the composition of individuals by number of charges, length of stay, or demographic characteristics such as gender, race, and age based on booking records, which we obtained for 34 out of 46 facilities.

Shifting Deaths Out of Jails: Another concern is that deaths are being shifted from the jail to the hospital or another external location in order for the jail to avoid responsibility for a "death in custody." Appendix Table A14 investigates this possibility by measuring mortality as all media reports of deaths related to a jail, even if the individual was pronounced dead at the hospital or if the location of death was undetermined. Our reduced form effects continue to show even more precise mortality reductions on the order of 50 to 60 percent using this more expansive measure of deaths. Most importantly, we find no evidence of an *increase* in non-jail deaths, ruling out the possibility that jails are strategically shifting deaths out of jails to avoid responsibility.

VII Mechanisms

In this section, we assess potential mechanisms that explain our findings on quality compliance and mortality. For these secondary analyses, we leverage several data sources, including the staff survey, facility survey, our review of inmate health records during the medical audit, and qualitative interviews. Overall, this evidence highlights organizational change among treatment facilities driven by increased staff coordination, sharing of information with administrators, and early screening of inmate populations.

Staff Sentiment. We first explore changes in staff sentiment, with a particular focus on changes to both horizontal and vertical coordination as described in our stylized model in Appendix B. Appendix Table A15

[26] We also estimate the effect of accreditation assignment on quality standards dropping the ten facilities that did not complete medical audits. While estimates are noisier given the smaller sample, estimated treatment effects are statistically indistinguishable from our main estimates.

presents reduced form treatment effects on staff attitudes and perceptions.[27] We find that accreditation assignment leads to meaningful and statistically significant improvements in job satisfaction and horizontal coordination. The estimated treatment effect on the job satisfaction index is 2.8 percentage points (s.e. = 1.4 pp), an approximately four percent increase relative to the control mean of 67.1 percent. Similarly, horizontal coordination – capturing communication and collaboration between custody and clinical staff – increases by 4.7 percentage points (s.e. = 2.2 pp) on a control mean of 64.3 percent, a roughly 7 percent improvement. Within these indices, we find that assignment to accreditation increased whether custody and medical staff coordinate and communicate and increased meaningful sub-components of the satisfaction index, including whether staff feel they are properly trained, are equipped with sufficient tools, and feel valued at work (Appendix Table A16).

While we find a positive although imprecise effect on an aggregate index for vertical coordination, individual components of the index such as whether staff feel comfortable speaking up on inmate care issues and whether authority is open to such feedback are positively affected by treatment (Appendix Table A16). These results also echo what we find in the qualitative comments: many in facility leadership noted the importance of knowing what the correct processes are and the importance of staff buy-in. Collectively, these staff findings are consistent with accreditation resolving information frictions to improve team coordination and enhance monitoring by administrators.

Inmate Screening. Given our finding that assignment to accreditation substantially reduces mortality, particularly deaths attributable to preventable illnesses, we now study the potential role of early stage screening. Within each of the two NCCHC standards for which we find detectable evidence of improvements in Figure 8 (Personnel Training, Patient Care and Treatment), we observe increased compliance along the majority of substandards, with notable increases in compliance with personnel credentials, staff training, health orientation, and the use of receiving screenings and mental health screenings for all inmates after admission (see Appendix Figure A5).

These findings are buttressed by our review of inmate medical records. We observe notable increases in documentation for health conditions and general and specific screening for medical conditions. Appendix Figure A6 presents reduced form estimates on required receiving screening components (the screening required "as soon as possible upon acceptance into custody"), mental health screening (a screening specific to mental health required within 14 days after admission), and initial health assessment (the screening that is more thorough and includes a physical exam by a medical professional required within 14 days after admission). Panel (A) of Appendix Figure A6 indicates that accreditation assignment increased screening for several acute and mental health–related conditions, with specific positive effects for suicide risk, withdrawal symptoms, and chronic disease indicators. Panel (B) shows that accreditation substantially improved documentation and screening related to mental health history, suicide risk, and prior psychiatric and addiction treatment, although effects on medication management and certain diagnostic categories are more mixed. Panel (C) suggests that accreditation strengthens some components of the initial health assessment tied to medical history, vital signs, and structured evaluations.

We also find notable increases in screening for particular conditions. Appendix Table A17 demonstrates that accreditation assignment affected morbidity screening for three high-risk conditions: suicidal ideation

[27] We construct indices of staff sentiment using endline staff surveys obtained from 41 out of 46 facilities, imputing baseline staff survey responses for the five facilities that did not complete the endline staff survey. All facilities completed the baseline staff survey. Specifications control for baseline facility-level means of the respective outcome and include randomization strata fixed effects. Outcomes are normalized to range from 0 to 1, with higher values indicating more positive sentiment.

(SI), opioid use disorder (OUD), and hypertension (HTN). Panel (A) presents results for screening indices for each of these three conditions. We find that accreditation assignment increases screening for SI by 14.2 percentage points, a 33 percent increase from the control mean that is marginally significant at the 10 percent level. Effects on screening for OUD and HTN are positive but not statistically significant at conventional levels. Panel (B) reports effects on treatment indices. Across all three conditions, the estimated impacts are small and statistically insignificant. Overall, the results suggest that accreditation strengthens intake and screening processes – particularly for mental health and high-risk conditions – by improving early identification, documentation, and structured health assessments. These results comport with our finding that treatment substantially reduces mortality associated with acute and preventable causes.

Inputs. In contrast to the staffing sentiment and screening changes documented above, we do not find that accreditation assignment yields structural changes in other inputs and investments (Appendix Table A18). For example, assignment to accreditation did not lead to an increase in the number of staff, health care management (*i.e.,* a change in vendor), use of staff (*i.e.,* change in referral patterns for offsite hospital or specialty care),[28] or capital investments. In contrast, treatment facilities increased the services (*e.g.,* lab and pharmacy) they provided to incarcerated individuals (see Appendix Table A18), consistent with our above findings on enhanced screenings both at intake and for high-risk conditions.

Qualitative Evidence. To better understand the nature of changes that occurred at treatment facilities as they underwent the NCCHC accreditation process, we now turn to qualitative evidence collected during our interviews with facility leadership. Among all treatment facilities, we asked several questions at endline on their experience with the accreditation process, the challenges they faced, the biggest change implemented, and the relationship between custody and medical staff.[29]

Word clouds depicting responses to these questions are shown in Appendix Figure A7. The word clouds demonstrate that staff find the experience helpful but also difficult, with some of the biggest challenges described as adopting new policies and processes. Changes that were implemented focus on improvement in protocols and documentation for medical and mental health, which could facilitate better coordination and monitoring. Almost all treatment facilities indicated that the biggest change was with respect to aspects of screening and initial health assessments. One facility noted "It was that intake for the initial and it going immediately going to medical." Another facility stated that "receiving screening is one of the bigger ones. Way more contact with patients when they come in now. Also coordination with custody staff is more comprehensive now."

Mental health screening and processes of care were also subject to substantial change, according to some facilities. For example, one facility stated, "Staying on top of mental health assessments was important. With the new protocols, the facility can ensure psychiatrists are staying on top of what they need to be doing." Other facilities shared similar changes, such as "Really focused on turnaround time for health assessments and review," and "chronic disease management, screening everyone on admission, accreditation gave teeth

[28] NCCHC standards require "appropriate and timely access to hospital and specialist care when necessary." Our survey instruments indicate that baseline levels of referrals are already high, with staff typically responding that health care staff refer inmates out for care "Often."

[29] Specifically we asked facilities to (1) "Please describe your facility's experience undergoing NCCHC accreditation process," (2) "Were there any challenges your facility faced during the NCCHC accreditation process? If so, what were they?" (3) "In your opinion, what is the biggest change your facility has implemented as a result of the NCCHC accreditation process?" and (4) "Do you think that the relationship and coordination between correctional and health care staff has changed due to the NCCHC accreditation process? If so, how has it changed?"

to move these forward on the admin end," and "Increased timeline of how fast we are seeing people when they come in. Increasing the care they are receiving."[30]

Lastly, although some custody and medical staff emphasized they already worked well together, others noted the accreditation process spurred conversations and improved understanding (Panel (D) of Appendix Figure A7). About half of the treatment facilities noted substantial changes in the relationship between medical and custody. Medical leadership at one facility told us that "The staff had to have conversations about how processes will work, which meant discussing how custody versus medical execute certain processes." Medical staff reported that "Showing the deputies all of the different regulations and rules we need to follow that they need to help us with. Understand why we are asking for certain things. Really good education for us and custody. Makes it easier to do our work." As one facility leader noted: "Yeah, when you look at officers they have a better understanding why they are making sure sick call slips are being given timely and being seen timely after submission rather than just giving aspirin." Another facility commented that the "process in general is awesome... Its about finding shortfalls and closing the gap. Staff has become more involved in the medical process and have better understanding of why we are doing those things. Building relationships between medical and security was another piece."

VIII Conclusion

In this study, we evaluate whether accreditation – a widely-used but under-studied quality assurance mechanism – can improve healthcare delivery in jails. Over several years, we conducted, to the best of our knowledge, the first randomized controlled trial across U.S. jails and the first experimental evaluation of healthcare accreditation. Jails occupy a distinct role in the healthcare system: they function as de facto emergency departments, admitting individuals who are often young but acutely ill or in crisis. Deaths in custody frequently arise from the interaction between these acute conditions and the constraints of confinement – for example, suicidality exacerbated by segregation, substance use complicated by withdrawal, or treatable illnesses progressing due to delays in care. In this setting, how facilities screen, process, and triage individuals in the initial hours and days of incarceration has first-order implications for survival.

Using original survey instruments and administrative data, we find that assignment to accreditation has statistically significant and medically meaningful effects on both process measures and health outcomes. Assignment to accreditation improves compliance with quality standards, particularly in personnel training and patient care. We also find reductions in mortality, with declines concentrated in jails that experienced the greatest quality improvements.

These effects appear to operate through organizational changes rather than structural changes in inputs. We find little evidence of increased staffing or capital investment. Instead, accreditation improves coordination, training, and adherence to protocols, particularly in screening and intake. These empirical patterns are consistent with a model in which accreditation reduces information frictions, facilitates coordination, and enables more effective monitoring in a setting characterized by team production and incomplete contracts with private service providers.

Taken together, our findings suggest that accreditation can meaningfully improve healthcare quality in correctional settings. More broadly, the results highlight the potential for standardized oversight and man-

[30]Dental care was also mentioned by several facilities, and it was noted to be particularly important given the lengthening stay of inmates in jails.

agerial interventions to improve performance in public services delivered under contracting, particularly in environments with weak external monitoring and non-expert principals.

Bibliography

Adler, Jessica L., and Weiwei Chen. 2023. "Jail Conditions and Mortality: Death Rates Associated with Turnover, Jail Size, and Population Characteristics." *Health Affairs*, 42(6): 849–857.

Alchian, Armen A, and Harold Demsetz. 1972. "Production, Information Costs, and Economic Organization." *American Economic Review*, 62(5): 777–795.

Alhawajreh, Mohammad J., Audrey S. Paterson, and William J. Jackson. 2023. "Impact of Hospital Accreditation on Quality Improvement in Healthcare: A Systematic Review." *PLoS One*, 18: 1–20.

Alsan, Marcella, Crystal S. Yang, James R. Jolin, Lucy Tu, and Josiah D. Rich. 2023. "Health Care in U.S. Correctional Facilities—a Limited and Threatened Constitutional Right." *New England Journal of Medicine*, 388(9): 847–852.

American Medical Association. 1979. *American Medical Association Standards for Health Services in Jails.* American Medical Association Program to Improve Medical Care and Health Services in Correctional Institutions.

Association of Health Care Journalists. n.d.. "A Q&A with CMS: Getting Up to Speed on Inspection Reports." [Internet] Accessed April 11, 2026. Available from https://www.hospitalinspections.org/qanda.

Athey, Susan, and Guido W. Imbens. 2017. "Chapter 3 - The Econometrics of Randomized Experiments." In *Handbook of Economic Field Experiments*. Vol. 1, 73–140. Elsevier.

Atkin, David, Azam Chaudhry, Shamyla Chaudry, Amit K. Khandelwal, and Eric Verhoogen. 2017. "Organizational Barriers to Technology Adoption: Evidence from Soccer-Ball Producers in Pakistan." *Quarterly Journal of Economics*, 132(3): 1101–1164.

Bai, Jushan. 2003. "Inferential Theory for Factor Models of Large Dimensions." *Econometrica*, 71(1): 135–171.

Batistich, Mary Kate, William N. Evans, and David C. Phillips. 2021. "Reducing Re-Arrests through Light Touch Mental Health Outreach." Technical report, University of Notre Dame.

Bedard, Kelly, and Harry E. Frech. 2009. "Prison health care: is contracting out healthy?" *Health Economics*, 18(11): 1248–1260.

Berwick, Donald M., Adam L. Beckman, and Suhas Gondi. 2021. "The Triple Aim Applied to Correctional Health Systems." *JAMA*, 325(10): 935–936.

Bhuller, Manudeep, Laura Khoury, and Katrine V Løken. 2025. "Mental Health Consequences of Correctional Sentencing." *American Economic Journal: Economic Policy*, 17(1): 70–105.

Bianchi, Nicola, and Michela Giorcelli. 2022. "The Dynamics and Spillovers of Management Interventions: Evidence from the Training within Industry Program." *Journal of Political Economy*, 130(6): 1630–1675.

Binswanger, Ingrid A., Marc F. Stern, Richard A. Deyo, Patrick J. Heagerty, Allen Cheadle, Joann G. Elmore, and Thomas D. Koepsell. 2007. "Release from Prison — A High Risk of Death for Former Inmates." *New England Journal of Medicine*, 356(2): 157–165.

Bloom, Nicholas, Aprajit Mahajan, David McKenzie, and John Roberts. 2020. "Do Management Interventions Last? Evidence from India." *American Economic Journal: Applied Economics*, 12(2): 198–219.

Bloom, Nicholas, Benn Eifert, Aprajit Mahajan, David McKenzie, and John Roberts. 2013. "Does Management Matter? Evidence from India." *Quarterly Journal of Economics*, 128(1): 1–51.

Bovell-Ammon, Benjamin J., Ziming Xuan, Michael K. Paasche-Orlow, and Marc R. LaRochelle. 2021. "Association of Incarceration with Mortality by Race from a National Longitudinal Cohort Study." *JAMA Network Open*, 4(12): e2133083–e2133083.

Bronson, Jennifer, and Marcus Berzofsky. 2017. "Indicators of Mental Health Problems Reported by Prisoners and Jail Inmates, 2011-12." Special Report. NCJ 250612. *US Department of Justice, Office of Justice Programs, Bureau of Justice Statistics*.

Bruhn, Miriam, Dean Karlan, and Antoinette Schoar. 2018. "The Impact of Consulting Services on Small and Medium Enterprises: Evidence from a Randomized Trial in Mexico." *Journal of Political Economy*, 126(2): 635–687.

Carson, E. Ann. 2021. "Mortality in Local Jails, 2000–2019-Statistical Tables." Mortality in Local Jails and State Prisons. *US Department of Justice, Office of Justice Programs, Bureau of Justice Statistics*.

Chassin, Mark R. 2021. "Personal Communication." August, 2021. Notes available from https://www.dropbox.com/scl/fi/bb48ojpxcv23b2ht69p8q/Joint-Commission-Dr-Chassin_2aug2021.docx?rlkey=nffjk24l2yinugsjt5wqdam95&dl=0.

Cheng, Alden. 2023. "Barriers to Choosing High-Quality Healthcare Providers: Evidence from the Nursing Home Market." Available from: https://aldenpapers.s3.amazonaws.com/nh_choice_info_fric.pdf.

Chen, Jiafeng, and Jonathan Roth. 2023. "Logs with Zeros? Some Problems and Solutions." *The Quarterly Journal of Economics*, 139(2): 891–936.

Chen, Yiqun, and Marcus Dillender. 2026. "Government Monitoring of Health Care Quality: Evidence from the Nursing Home Sector." NBER Working Paper 34037.

Clotfelter, Charles T, Helen F Ladd, and Jacob L Vigdor. 2010. "Teacher Credentials and Student Achievement in High School: A Cross-Subject Analysis with Student Fixed Effects." *Journal of Human Resources*, 45(3): 655–681.

Cloud, David H., Brie Williams, Regine Haardörfer, Lauren Brinkley-Rubinstein, and Hannah L.F. Cooper. 2023. "Extreme Heat and Suicide Watch Incidents among Incarcerated Men." *JAMA Network Open*, 6(8): e2328380–e2328380.

Coll, Steve. 2019. "The Jail Health-Care Crisis." *The New Yorker*. [Internet] Accessed November, 2024. Available from https://www.newyorker.com/magazine/2019/03/04/the-jail-health-care-crisis.

Congressional Research Service. 2023. "Death in Custody Reporting Act: Background and Legislative Considerations." *Congressional Research Service*

Cowan, Kristen N., Katherine Lemasters, and Lauren Brinkley Rubinstein. 2024. "Temperatures Extremes and Mortality during Summer Months in Federal Prisons across the United States, 2010-2021." Vol. 2024.

Curran, Jill, Brendan Saloner, Tyle Winkelman, and G. Caleb Alexander. 2023. "Estimated Use of Prescription Medications among Individuals Incarcerated in Jails and State Prisons in the U.S." *JAMA Health Forum*, 4(4): e230482–e230482.

Daza, Sebastian, Alberto Palloni, and Jerrett Jones. 2020. "The Consequences of Incarceration for Mortality in the United States." *Demography*, 57(2): 577–598.

Douds, Anne S., and Eileen M. Ahlin. 2016. "Do NCCHC Dental Standards Have any Teeth?" *Journal of Correctional Health Care*, 22(3): 180–188.

Dranove, David, and Ginger Zhe Jin. 2010. "Quality Disclosure and Certification: Theory and Practice." *Journal of Economic Literature*, 48(4): 935–963.

Edmonds, Mira. 2020. "The Reincorporation of Prisoners into the Body Politic: Eliminating the Medicaid Inmate Exclusion Policy." *Georgetown Journal on Poverty Law and Policy*, 28: 279.

Einav, Liran, Amy Finkelstein, Yunan Ji, and Neale Mahoney. 2022. "Voluntary Regulation: Evidence from Medicare Payment Reform." *The Quarterly Journal of Economics*, 137(2): 565–618.

Estelle v. Gamble. 1976. "Estelle v. Gamble." 429 U.S. 97.

Gandhi, Ashvin, Andrew Olenski, and Maggie Shi. 2025. "Predictably Unpredictable Inspections." NBER Working Paper 34491.

GAO. 2022. "Deaths in Custody: Additional Action Needed to Help Ensure Data Collected by DOJ Are Utilized." *United States Government Accountability Office*, 2022.

Gibbons, Robert, and Rebecca Henderson. 2013. "What Do Managers Do?" In *The Handbook of Organizational Economics*. , ed. Robert Gibbons and John Roberts, 680–731. Princeton, NJ: Princeton University Press.

Gibson, Brent R., and Gary Phillips. 2016. "Challenges and Opportunities in Correctional Health Care Quality: a Descriptive Analysis of Compliance with NCCHC Standards." *Journal of Correctional Health Care*, 22(4): 280–289.

Giorcelli, Michela. 2019. "The Long-Term Effects of Management and Technology Transfers." *American Economic Review*, 109(1): 121–152.

Goldhaber, Dan, and Emily Anthony. 2007. "Can Teacher Quality be Effectively Assessed? National Board Certification as a Signal of Effective Teaching." *Review of Economics and Statistics*, 89(1): 134–150.

Gosnell, Greer K, John A List, and Robert D Metcalfe. 2020. "The Impact of Management Practices on Employee Productivity: A Field Experiment with Airline Captains." *Journal of Political Economy*, 128(4): 1195–1233.

Grant Smith. 2020. "Dying Inside: Data & Key Findings." *https://www.reuters.com/investigates/special-report/usa-jails-graphic/*, Data reporting by Grant Smith, Peter Eisler, Linda So, Ned Parker, Brad Heath, and Jason Szep; additional contributors listed on the page. A Reuters Investigation. Filed October 16, 2020. Accessed March 13, 2026.

Greene, Meredith, Cyrus Ahalt, Irena Stijacic-Cenzer, Lia Metzger, and Brie Williams. 2018. "Older Adults in Jail: High Rates and Early Onset of Geriatric Conditions." *Health & Justice*, 6: 1–9.

Green, Hannah Harris. 2025. "Private-equity backed prison health companies continue despite decade of alleged constitutional violations." *The Guardian*.

Greifinger, Robert B. 2007. "Thirty Years Since Estelle v. Gamble: Looking Forward, Not Wayward." In *Public Health Behind Bars: From Prisons to Communities*. 1–10. Springer.

Hart, Oliver, Andrei Shleifer, and Robert W. Vishny. 1997. "The Proper Scope of Government: Theory and an Application to Prisons." *The Quarterly Journal of Economics*, 112(4): 1127–1161.

Health Forum. 2021. "AHA Hospital Statistics." 2021 Edition. *AHA Data & Insights. American Hospital Association.*

Health Forum. 2024. "AHA Hospital Statistics." 2024 Edition. *AHA Data & Insights. American Hospital Association.*

Hjalmarsson, Randi, and Matthew J. Lindquist. 2022. "The Health Effects of Prison." *American Economic Journal: Applied Economics*, 14(4): 234–270.

Holmstrom, Bengt. 1982. "Moral Hazard in Teams." *The Bell Journal of Economics*, 13(2): 324–340.

Howell, Benjamin A., Laura C. Hawks, Lilanthi Balasuriya, Virginia W. Chang, Emily A. Wang, and Tyler Winkelman. 2023. "Health Insurance and Mental Health Treatment Use Among Adults with Criminal Legal Involvement after Medicaid Expansion." *Psychiatric Services*, 74(10): 1019–1026.

Hussein, Mohammed, Milena Pavlova, Mostafa Ghalwash, and Wim Groot. 2021. "The impact of hospital accreditation on the quality of healthcare: a systematic literature review." *BMC Health Services Research*, 21(1): 1057.

Iacovone, Leonardo, William Maloney, and David McKenzie. 2022. "Improving Management with Individual and Group-Based Consulting: Results from a Randomized Experiment in Colombia." *Review of Economic Studies*, 89(1): 346–371.

Jha, Ashish K. 2018. "Accreditation, Quality, and Making Hospital Care Better." *JAMA*, 320(23): 2410–2411.

Kane, Thomas J, Jonah E Rockoff, and Douglas O Staiger. 2008. "What does Certification Tell us about Teacher Effectiveness? Evidence from New York City." *Economics of Education Review*, 27(6): 615–631.

Kwon, Soonwoo, and Liyang Sun. 2025. "Estimating Treatment Effects Under Bounded Heterogeneity." *arXiv preprint arXiv:2510.05454*.

Lise, Jeremy, and Fabien Postel-Vinay. 2020. "Multidimensional Skills, Sorting, and Human Capital Accumulation." *American Economic Review*, 110(8): 2328–2376.

Maner, Morgan, Marisa Omori, Lauren Brinkley-Rubinstein, Curt G. Beckwith, and Kathryn Nowotny. 2022. "Infectious Disease Surveillance in U.S. Jails: Findings from a National Survey." *PLoS One*, 17(8): e0272374.

Martinez, Elizabeth A., Nancy Beaulieu, Robert Gibbons, Peter Pronovost, and Thomas Wang. 2015. "Organizational Culture and Performance." *American Economic Review: Papers & Proceedings*, 105(5): 331–335.

Martin, Jeffrey C. 1994. "Recent Developments Concerning Accrediting Agencies in Postsecondary Education." *Law and Contemporary Problems*, 57(4): 121–150.

McCann, Sam. 2024. "The Hidden Deaths in American Jails." *Vera*, 2024. [Internet] Accessed November, 2024. Available from https://www.vera.org/news/the-hidden-deaths-in-american-jails#:~:text=They%20exist%20for%20one%20simple,death%20data%20is%20deliberately%20opaque.

McLeod, Marsha. 2019. "The Private Option." *The Atlantic*. [Internet] Accessed November, 2024. Available from https://www.theatlantic.com/politics/archive/2019/09/private-equitys-grip-on-jail-health-care/597871/.

Mitchell Jr., Roger A, and Jay D Aronson. 2023. *Death in Custody: How America Ignores the Truth and What We Can Do about It*. Johns Hopkins University Press.

Morford, Thomas G. 1988. "Nursing Home Regulation: History and Expectations." *Health Care Financing Review*, 1988.

Mukherjee, Anita, and Nicholas J Sanders. 2021. "The Causal Effect of Heat on Violence: Social Implications of Unmitigated Heat among the Incarcerated." NBER Working Paper 28987.

Mullahy, John, and Edward C. Norton. 2024. "Why Transform Y? The Pitfalls of Transformed Regressions with a Mass at Zero." *Oxford Bulletin of Economics and Statistics*, 86(2): 417–447. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/obes.12583.

NASEM. 2019. *The Effects of Incarceration and Reentry on Community Health and Well-Being: Proceedings of a Workshop*. National Academies Press (US). National Academies of Sciences, Engineering, and Medicine; Health and Medicine Division; Board on Population Health and Public Health Practice; Roundtable on the Promotion of Health Equity. Karen M. Anderson and Steve Olson, editors.

National Commission on Correctional Health Care. 2021. "The Importance of Custody-Health Care Collaboration, Part 2." [Internet] Accessed November, 2024. Available from https://www.ncchc.org/the-importance-of-custody-health-care-collaboration-part-2/.

Norris, Samuel, Matthew Pecenco, and Jeffrey Weaver. 2024. "The Effect of Incarceration on Mortality." *Review of Economics and Statistics*, 106(4): 956–973.

Nursing Home Law Center. 2025. "Nursing Home Reform Act." *https://www.nursinghomelawcenter.org/news/nursing-home-reform-act-obra/*, [Internet] Accessed April 11, 2026.

Patterson, Evelyn J. 2010. "Incarcerating Death: Mortality in U.S. State Correctional Facilities, 1985–1998." *Demography*, 47: 587–607.

Pew Charitable Trusts. 2014. "Managing Prison Health Care Spending." Pew Charitable Trusts and John D. and Catherine T. MacArthur Foundation.

Pew Charitable Trusts. 2017. "Prison Health Care: Costs and Quality."

ProPublica. 2026. "Nonprofit Explorer: Form 990 Filing for Organization 36-3221830 (2025)." *https://projects.propublica.org/nonprofits/organizations/363221830/202543219349306149/full*, Accessed: 2026-04-06.

Puglisi, Lisa B., and Emily A. Wang. 2021. "Health Care for People who are Incarcerated." *Nature Reviews Disease Primers*, 7(1): 50.

Ruch, Donna A., Danielle L. Steelesmith, Guy Brock, Samantha J. Boch, Camille R. Quinn, Jeffrey A. Bridge, John V. Campo, and Cynthia A. Fontanella. 2021. "Mortality and Cause of Death Among Youths Previously Incarcerated in the Juvenile Legal System." *JAMA Network Open*, 4(12): e2140352–e2140352.

Safety Source. 2021. "National Directory of Law Enforcement Administrators."

Saloner, Brendan. 2023. "A Chance to Modernize Health Care Behind Bars-Section 1115 Medicaid Inmate Exclusion Waivers." *New England Journal of Medicine*, 389(16): 1449–1451.

Sechrest, Dale K. 1976. "The Accreditation Movement in Corrections." *Federal Probation*, 40: 15. *U.S. Department of Justice. National Institute of Justice. Office of Research Programs.*

Senate Permanent Subcommittee on Investigations. 2022. "Uncounted Deaths in America's Prisons & Jails: How the Department of Justice Failed to Implement the Death in Custody Reporting Act." Staff Report.

Skarha, Julianne, Amite Dominick, Keith Spangler, David Dosa, Josiah D. Rich, David A. Savitz, and Antonella Zanobetti. 2022. "Provision of Air Conditioning and Heat-Related Mortality in Texas Prisons." *JAMA Network Open*, 5(11): e2239849–e2239849.

Skarha, Julianne, Keith Spangler, David Dosa, Josiah D. Rich, David A. Savitz, and Antonella Zanobetti. 2023. "Heat-Related Mortality in U.S. State and Private Prisons: A Case-Crossover Analysis." *PLoS One*, 18(3): e0281389.

Smith, Amy. 2013. *Health and Incarceration: A Workshop Summary.* National Academies Press. National Research Council; Division of Behavioral and Social Sciences and Education; Institute of Medicine; Committee on Law and Justice; Board on the Health of Select Populations; Committee on Causes and Consequences of High Rates of Incarceration.

Spaulding, Anne C., Ryan M. Seals, Victoria A. McCallum, Sebastian D. Perez, Amanda K. Brzozowski, and N. Kyle Steenland. 2011. "Prisoner Survival Inside and Outside of the Institution: Implications for Health-Care Planning." *American Journal of Epidemiology*, 173(5): 479–487.

Szep, Jason, Ned Parker, Linda So, Peter Eisler, and Grant Smith. 2020. "Dying Inside: The Hidden Crisis in America's Jails." *https://www.reuters.com/investigates/special-report/usa-jails-privatization/*, A Reuters Investigation. Filed October 26, 2020. Accessed April 17, 2026.

The Joint Commission. 2024. "Hospital Accreditation Fact Sheet." [Internet] Accessed November, 2024. Available from https://www.jointcommission.org/resources/news-and-multimedia/fact-sheets/facts-about-hospital-accreditation/.

The Office of Elizabeth Warren. 2019. "Warren Opens Investigation into Private Detention Accreditation System After Widespread Reports of Substandard Conditions & Safety Concerns in Multiple Facilities." [Internet] Accessed April 8, 2026. Available from https://www.warren.senate.gov/oversight/letters/warren-opens-investigation-into-private-detention-accreditation-system-after-widespread-reports-of-substandard-conditions-and-safety-concerns-in-multiple-facilities.

Tuholske, Cascade, Victoria D. Lynch, Raenita Spriggs, Yoonjung Ahn, Colin Raymond, Anne E. Nigra, and Robbie M. Parks. 2024. "Hazardous Heat Exposure among Incarcerated People in the United States." *Nature Sustainability*, 7(4): 394–398.

U.S. Department of Justice. 2022. "Census of Jails, 2019." United States Department of Justice. Office of Justice Programs. Bureau of Justice Statistics. Inter-university Consortium for Political and Social Research [distributor]. Accessed November, 2024. Data available from https://www.icpsr.umich.edu/web/NACJD/studies/38323.

Wang, Leah. 2021. "Rise in Jail Deaths is Especially Troubling as Jail Populations Become More Rural and More Female." *Prison Policy Initiative*, 2021. [Internet] Accessed November, 2024. Available from https://www.prisonpolicy.org/blog/2021/06/23/jail_mortality/#:~:text=New%20data%20show%20record%20high,ensures%20more%20deaths%20to%20come.

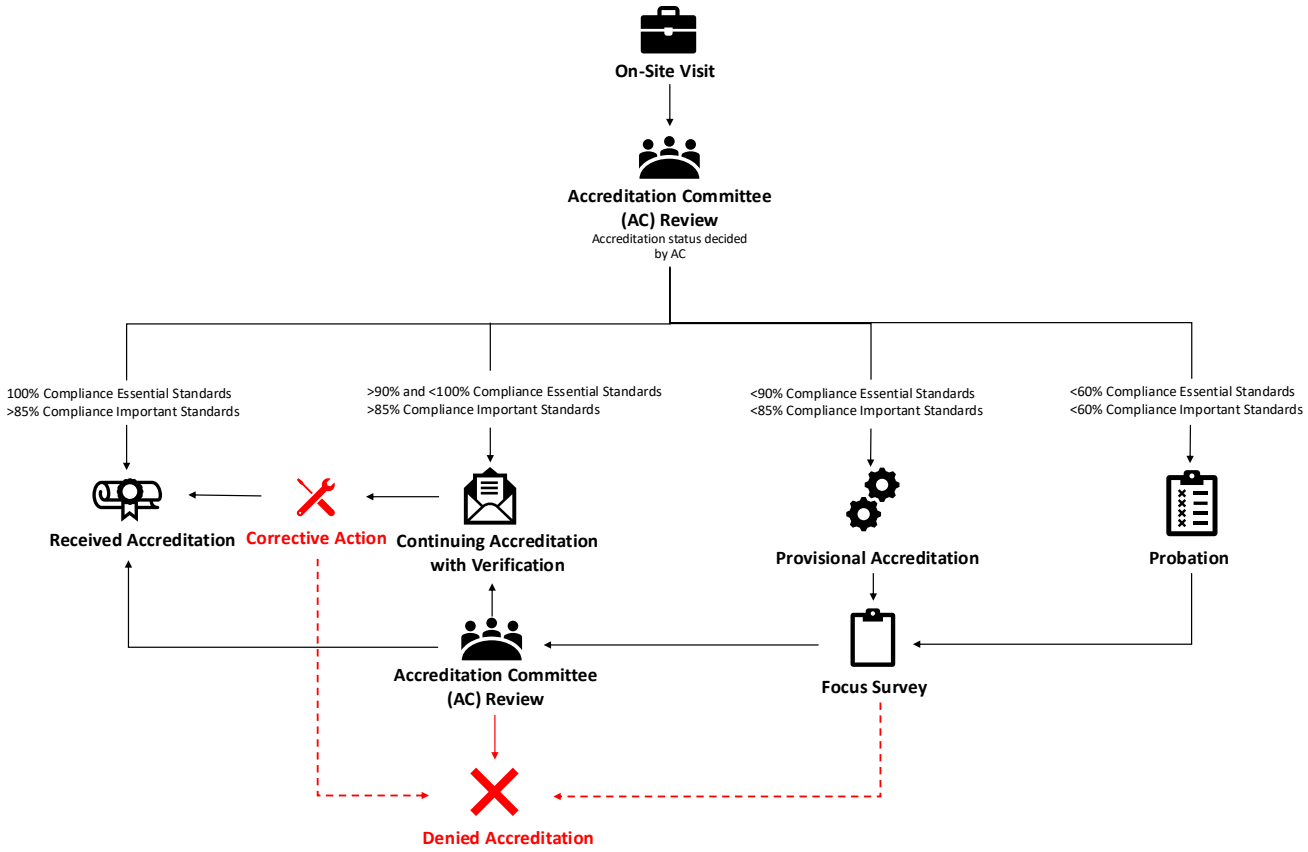
Wildeman, Christopher, E. Ann Carson, Daniela Golinelli, Margaret E. Noonan, and Natalia Emanuel. 2016. "Mortality Among White, Black, and Hispanic Male and Female State Prisoners, 2001–2009." *SSM-Population Health*, 2: 10–13.

Williams, Brie, Michele DiTomas, and Alison Pachynski. 2021. "The Growing Geriatric Prison Population: A Dire Public Health Consequence of Mass Incarceration." *Journal of the American Geriatrics Society*, 69(12): 3407–3409.

Zeng, Zhen. 2023. "Jail Inmates in 2022 – Statistical Tables." *U.S. Department of Justice, Bureau of Justice Statistics*.

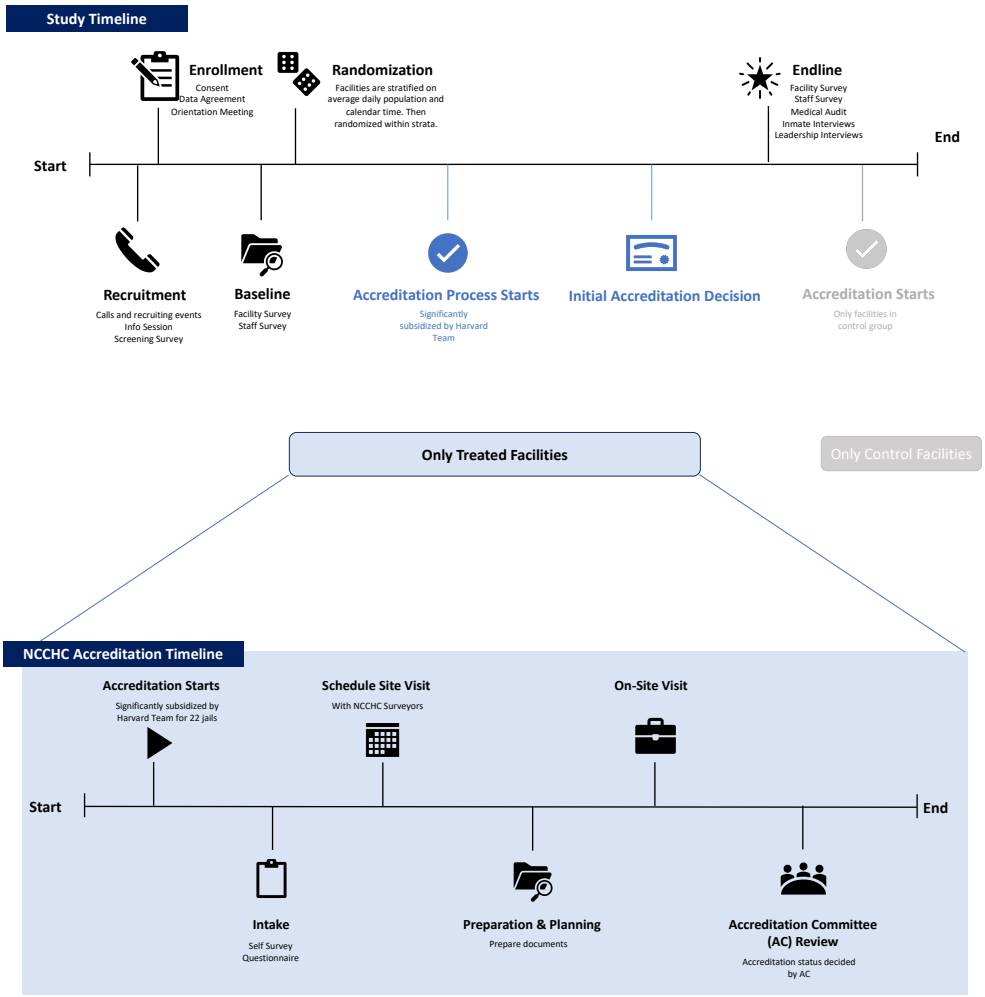
Zeng, Zhen, and Todd D. Minton. 2021. "Census of Jails, 2005–2019 – Statistical Tables." *U.S. Department of Justice, Bureau of Justice Statistics*.

Figure 1: Accreditation Decision Process



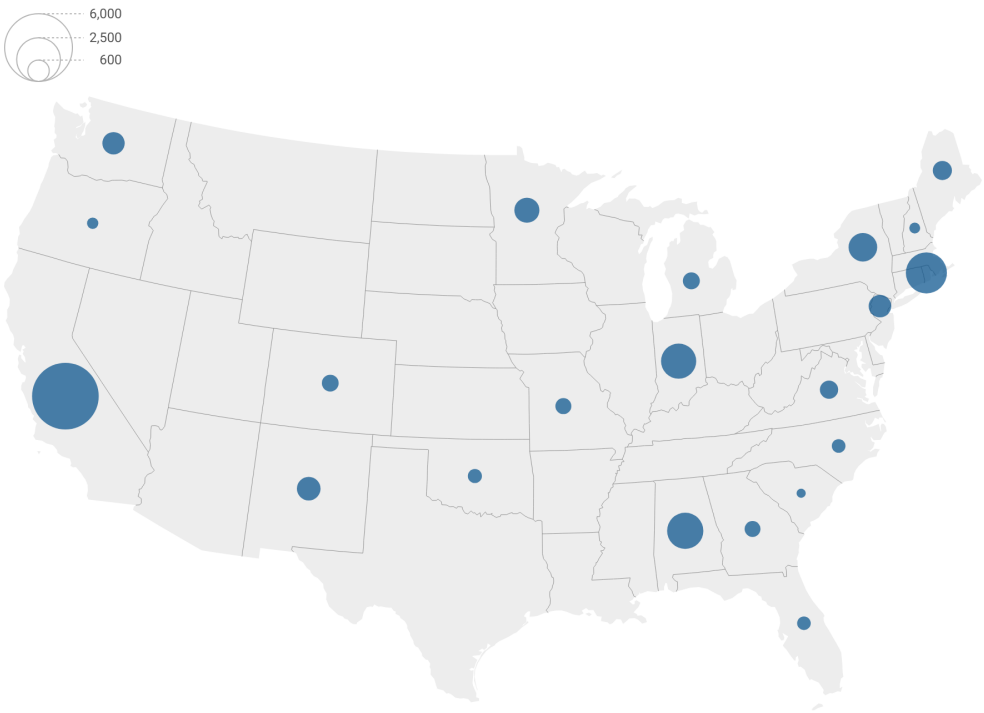
Notes: This figure depicts the NCCHC accreditation process. The process is based in part on compliance with Important and Essential Standards. Facilities that comply with less than 90 percent of Essential Standards and less than 85 percent of Important Standards have apparent deficiencies that significantly lower their compliance levels below that acceptable for accreditation and are placed on "Provisional Accreditation" status. Facilities that comply with less than 60 percent of Essential and Important Standards have excessive standard noncompliance and are placed on "Probation" status.

Figure 2: Study Timeline



Notes: This figure depicts the study timeline. The top panel illustrates the study timeline from recruitment to endline for all facilities. The bottom panel illustrates the NCCHC accreditation process for facilities randomized to the treatment group who initiated the accreditation process. See also Appendix Section C.1 for additional details.

Figure 3: Map of Study Jails and ADP



Notes: This map depicts the geographic location of the facilities in our sample (N=46). The size of the circle represents the total ADP across facilities within that state.

Figure 4: Example from Facility Survey Coding Manual

Table of Contents

Variable Name	Reason for Potential Skip	Format	Value if no Information	Question	Coder to Determine Instructions	Valid Entries	Variable Label	Variable Description
upl_griev_pol		*Text*		*Please upload the protocol for addressing health-related grievances*	***Filled by facility***		*Grievance Policy*	*Indicates status of facility uploaded document*
A10_GRV01_F	Skip if answer to upl_griev_pol is not "Document Uploaded" or "In policy and procedures manual"	Yes (1) - No (0)		Does the documentation uploaded by the facility include a protocol for addressing health-related grievances?	Coder to look at uploaded documentation/facility answer (Based on document upload - upl_griev_pol)	0 and 1	Grievance Policy: Check	Indicates whether there is a grievance policy for health care complaints
A10_GRV02_F	Skip if A10_GRV01_F was skipped, or the answer is No (0)	Integer	-99	Based on the grievance policy, once someone files a grievance, how many days until there is a formal response?	Determine Answer by Checking Grievance Policy. A simple verbal acknowledgement of the grievance does not count as a formal response (Based on document upload - upl_griev_pol)	Integers	Grievances: Days for Formal Response	States, based on grievance policy, how many days until there is a formal response after someone files a grievance
A10_GRV03_F	Skip if A10_GRV01_F was skipped, or the answer is No (0)	Yes (1) - No (0)		Does the grievance policy indicate that there is a process for appealing?	Determine Answer by Checking Grievance Policy (Based on document upload - upl_griev_pol)	0 and 1	Grievances: Appeal Process	Indicates, based on grievance policy, whether there is a process for appealing
A10_GRV04_F	Skip if A10_GRV01_F was skipped, or the answer is No (0)	Yes (1) - No (0)		Does the grievance policy indicate that there are formal records of health-related grievances?	Coder to look at uploaded documentation/facility answer (Based on document upload - upl_griev_pol)	0 and 1	Grievances: Formal Records	Indicates whether there are formal records of health-related grievances from the past month

J-A-10
important

GRIEVANCE PROCESS FOR HEALTH CARE COMPLAINTS

Standard

The facility protects a patient's right to disagree with or question the health care system.

Compliance Indicators

1. A grievance process is in place.
2. The grievance policy includes:
   a. A time frame for response
   b. The process for appeal
3. Responses to inmate grievances are:
   a. Timely
   b. Based on principles of adequate medical care
   c. Include documentation of response
4. All aspects of the standard are addressed by written policy and defined procedures.

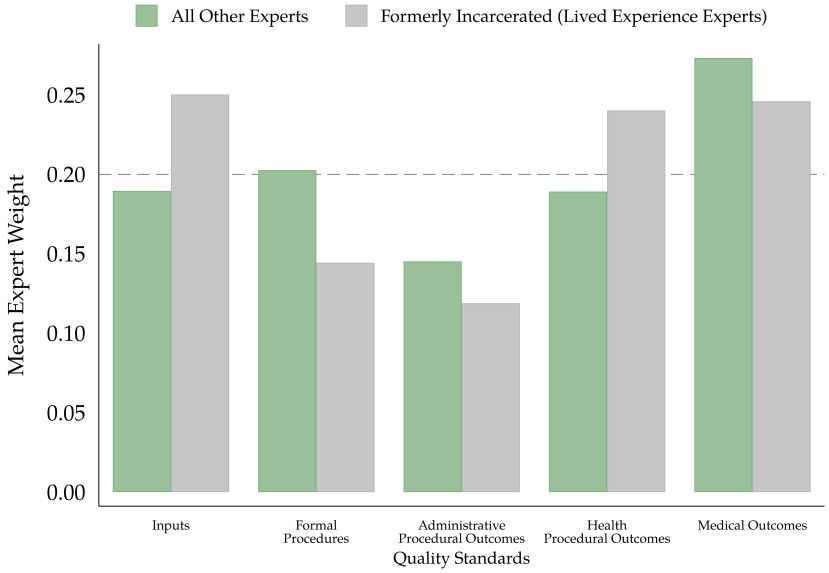
Notes: This figure depicts an example from the NCCHC standards and our coding manual that assists coders in determining whether a facility complies with the standards on grievance policies.

Figure 5: Expert Weights

Panel (A) NCCHC Indices



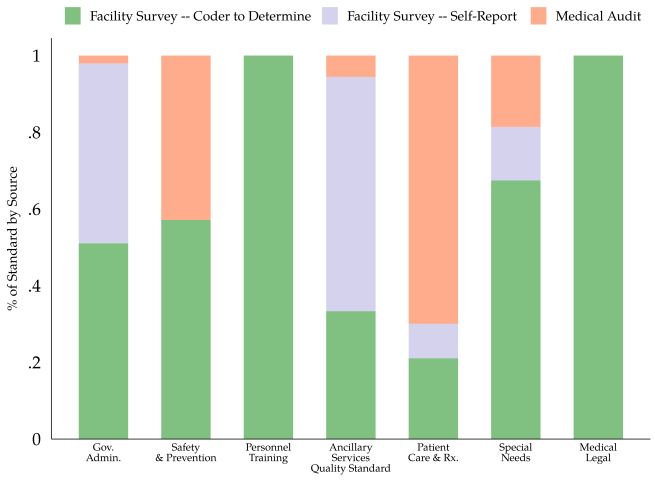
Panel (B) Production Function Indices



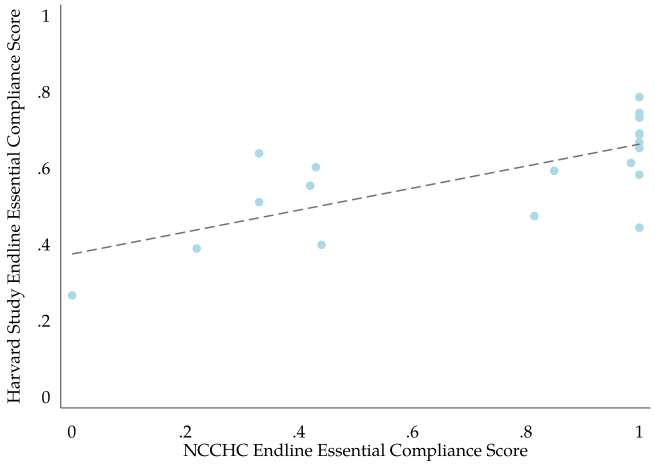
Notes: This figure shows the reciprocal rank weights placed on the seven NCCHC quality standards (Panel (A)) and on the five production function indices (Panel (B)) by the formerly incarcerated (lived experience experts) and other experts (jail administrators, jail health professionals, and policy and public health experts).

Figure 6: Measurement of Quality Standards

Panel (A) Composition of Standards

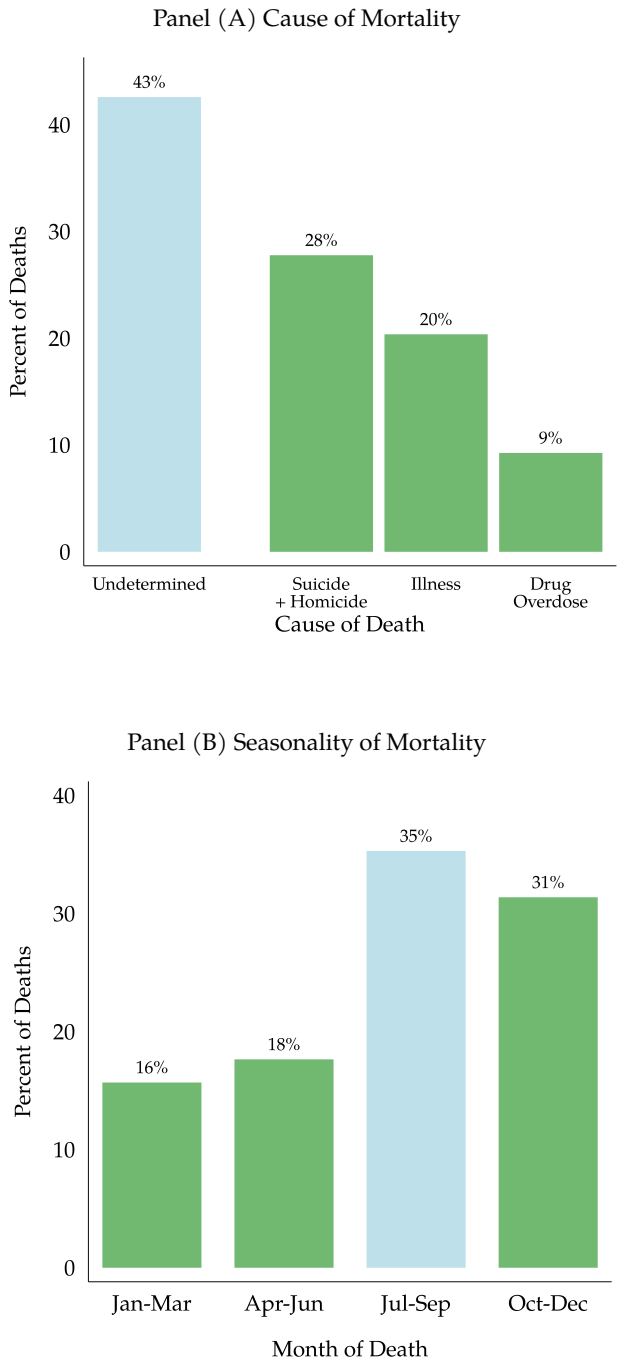


Panel (B) Corr. Harvard and NCCHC Compliance



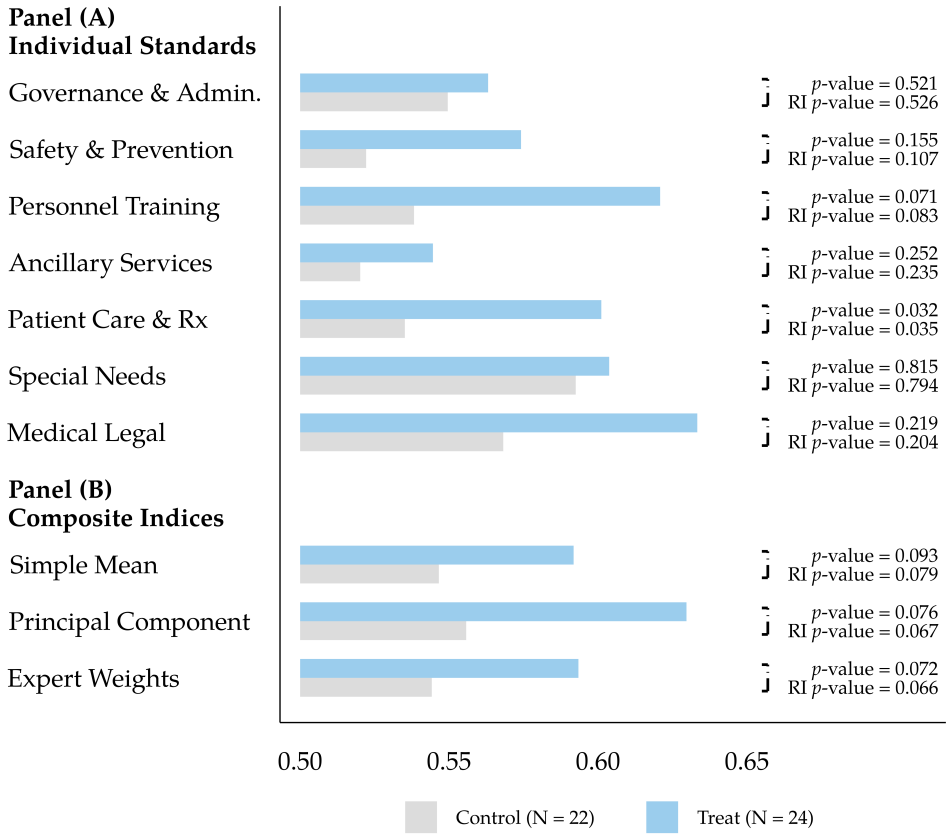
Notes: Panel (A) shows the proportion of variables used to construct the seven quality standards from the (i) facility survey – self-reported, (ii) the facility survey – coder to determine, and (iii) the medical audits, for all available facilities. Panel (B) shows the correlation between the Harvard study compliance score and the NCCHC compliance score. The study score follows the same approach as used in the NCCHC accreditation process, but uses the study-collected data instead.

Figure 7: Distribution of Mortality by Cause and Season



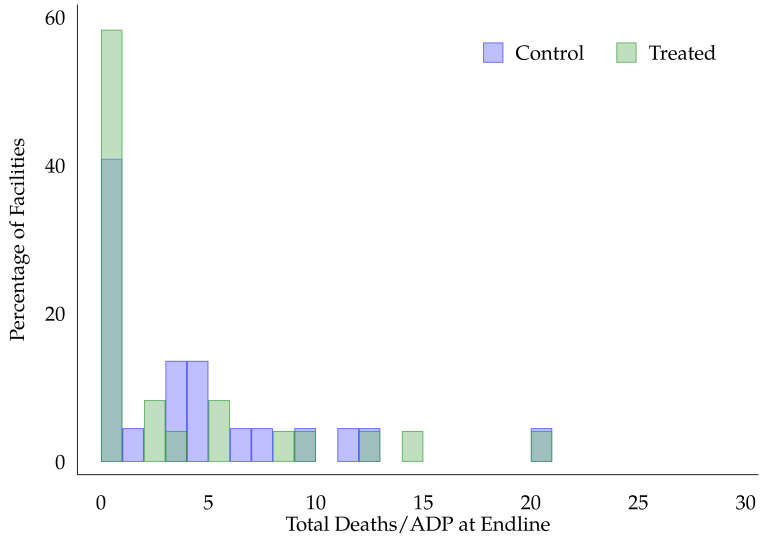
Notes: This figure shows cause and month of death using facility-reported deaths from administrative death logs in the 12 months prior to medical audit for facilities in the control group. Panel (A) disaggregates the total number of deaths in the sample by cause of death and plots the share of each cause as a percentage of total deaths. If the cause is unknown, undetermined, or if the report is pending, it is classified as Undetermined. Deaths by suicide or homicide are classified as such. If a pre-existing condition, infection, or any other physical ailment is indicated as the cause of death, the death falls under Illness. In the event of a drug toxicity, or overdose, the death is labeled as a Drug Overdose. Panel (B) shows the seasonality of deaths by plotting the mean number of deaths for each quarter of the year. The month-year of death indicated for each death in the death logs is used to identify the quarter of the year.

Figure 8: Effect of Accreditation Assignment on NCCHC Quality Standards



Notes: This figure shows the treatment effect of accreditation assignment on the seven NCCHC quality standards, both individually (Panel (A)) and aggregated into composite indices (Panel (B)). Quality standards are created using data from the baseline/endline facility survey and look-back medical audit. For detailed definitions of the standards and their calculation, see Appendix E.8. The composite indices are constructed by combining the seven quality standards in three pre-specified ways as described in the main text: equal weights, the first component using principal component analysis, and reciprocal rank weights using our expert survey responses. In all bars, the control mean (gray) is calculated as the mean of facilities in the control group. Treatment effects are obtained from estimating Equation 1, including controls for the baseline mean of the respective outcome variable at the facility level and randomization strata fixed effects. The treatment mean (blue) plotted is calculated as the control mean plus the estimated treatment effect. p-values based on robust standard errors and randomization inference (RI) are shown.

Figure 9: Distribution of Endline Mortality Rates by Treatment Status



Notes: This figure shows mortality rates at endline according to administrative death logs, separately by randomized treatment status. Similar histograms using media reports and covering the baseline period can be found in Appendix Figure A1. Mortality rates are computed as total deaths over the year prior to medical audit, divided by facility ADP at endline and multiplied by 1,000.

Table 1: Quality Standard Indices

Category	Description	Examples
Standard A *Governance Administration*	Standards in this section address the establishment of a health care system that ensures access to care, professional administration of all aspects of health care, and monitoring and quality improvement policies that effectively process health care issues from identification through resolution.	• medical record content • privacy • continuous quality improvement • grievance process • procedure in event of death
Standard B *Health Promotion, Safety and Disease Prevention*	Standards in this section address the need to optimize education, safety, and preventive care. Policies and procedures related to these standards require involvement by all facility staff.	• infectious disease prevention • suicide prevention • contraception
Standard C *Personnel Training*	Standards in this section ensure that appropriately trained personnel are in place to deliver health care to the inmate population and that qualified health care professionals are evaluated for continuing competency.	• custody staff training • health staff orientation • credentialing
Standard D *Ancillary Health and Care Services*	Standards in this section address the establishment and maintenance of all necessary procedures for the provision of ancillary health care services.	• medications • clinic space and supplies • emergency response plan
Standard E *Patient Care and Treatment*	Standards in this section ensure the delivery of health care from arrival through discharge for health care issues. All care is timely, appropriate, and continues until resolution of the problem or until discharge.	• health care screenings • medical & dental care • discharge planning • sick call requests • mental health
Standard F *Special Needs and Services*	Standards in this section address patients with special health care needs and establish compliance requirements specific to each health care issue.	• MAT • sexual abuse protocol • pregnancy care
Standard G *Medical-Legal Issues*	The standards in this section ensure that health services comply with legal requirements.	• use of restraints • segregation • informed consent/refusal

Table 2: Sample Representativeness

	Mean	Coefficient	SE
	(1)	(2)	(3)
Panel (A) Jail Characteristics			
ADP	372.281	25.600	(53.868)
Admissions	5,173.866	1,165.810	(739.517)
Avg. Stay Length (Months)	1.096	-0.204***	(0.077)
# Staff	104.820	30.691	(23.245)
Under Consent Decree	0.044	0.003	(0.033)
Maximum Jail Capacity	438.247	67.482	(84.800)
Misdemeanor and Felony Defendants	0.447	-0.094	(0.076)
Panel (B) Incarcerated Population Characteristics			
Share Male	0.833	0.009	(0.008)
Share White	0.596	-0.081**	(0.036)
Share Black	0.263	0.008	(0.034)
Share Hispanic	0.111	0.045	(0.031)
Share Felony	0.665	0.038	(0.031)
Share Misdemeanor	0.264	-0.015	(0.026)
F-stat (*p*-value)		1.455 (0.128)	
Analytical Sample Facilities		42	
Comparative Facilities		1085	

Notes: This table compares jails in the study sample to similar jails in the Census of Jails (COJ) (U.S. Department of Justice 2022). Forty-two of the 46 jails that report in the COJ are matched to 1,085 other county jails that house adults with an average daily population (ADP) between 100 and 3,000. Column (1) reports the sample mean. Column (2) reports the coefficient on an indicator equal to one if the facility is in the sample. Column (3) reports the associated robust standard error. ADP, the number of admissions, and the number of staff are reported at the reporting-unit level. A reporting unit may manage multiple facilities; these variables are adjusted to the facility level by dividing by the number of facilities managed by each reporting unit. Average length of stay (in months) is calculated as $(ADP \times 365/\text{Admissions})/30$. "Misdemeanor and Felony Defendants" is an indicator equal to 1 if the facility holds defendants convicted of both misdemeanor and felony offenses. An omnibus test of balance is reported by regressing in-sample status on all variables in the table and computing the *F*-statistic for joint significance. To preserve the full sample when computing the joint *F*-statistics, missing values in each variable are replaced with zero and a missing indicator is included. The *F*-statistic and corresponding *p*-value are reported at the bottom of the table. $*$, $**$, and $***$ denote statistical significance at the 10, 5, and 1 percent levels, respectively.

Table 3: Balance Across Outcomes and Jail Characteristics

	Mean	Coefficient	SE
	(1)	(2)	(3)
Panel (A) Baseline NCCHC Quality Standards			
Simple Mean Meta-Index	0.489	0.015	(0.049)
Principal Component Analysis Meta-Index	0.559	-0.000	(0.068)
Expert Weights Meta-Index	0.483	0.014	(0.047)
Panel (B) Baseline Mortality			
Mortality/ADP from Media Reports	2.160	-0.900	(1.578)
Panel (C) Baseline Jail Characteristics			
ADP	375.812	-173.462	(123.364)
Medical FTE	5.682	-1.243	(1.303)
Medical FTE Variety	5.739	-0.486	(0.970)
For-Profit Health Vendor	0.652	0.009	(0.145)
Previously Accredited	0.109	0.064	(0.096)
Share County Voting Republican	0.448	-0.021	(0.039)
Sheriff is Republican	0.326	0.018	(0.128)
Jail Located in South	0.283	0.147	(0.134)
F-Stat with Simple Mean Meta-Index (p-value)		1.136 (0.371)	
F-Stat with PCA Meta-Index (p-value)		1.223 (0.318)	
F-Stat with Expert Weights Meta-Index (p-value)		1.146 (0.364)	
Observations		46	
Strata FE		✓	

Notes: This table compares baseline outcomes and jail characteristics across treatment and control facilities. Column (1) reports the sample mean, and columns (2) and (3) report the coefficient on an indicator for treatment assignment and the associated robust standard error, respectively. Panel (A) compares baseline NCCHC quality standards (for detailed definitions of the standards and their calculation, see Appendix Section E.8). Panel (B) compares baseline mortality rates from media reports, (total deaths in the 12 months prior to randomization divided by the facility's baseline ADP and multiplied by 1,000). Panel (C) compares baseline jail characteristics. Medical FTE is the number of full-time equivalents of health staff. "For-profit health vendor" indicates that the facility has hired an external private company to provide health services. An omnibus test of balance is reported by regressing treatment assignment on all variables in the table, using each of the three meta-indices in Panel (A) in turn and controlling for randomization strata fixed effects. The F-statistics are computed from tests of joint significance. Joint F-statistics and the corresponding p-values are reported at the bottom of the table. $*$, $**$, and $***$ denote statistical significance at the 10, 5, and 1 percent levels, respectively.

Table 4: Effect of Accreditation Assignment on Mortality

	Count Model		Linear Model					
	Poisson		LPM		OLS		WLS	
	Total Deaths	Total Deaths	≤ 1 Deaths	≤ 1 Deaths	Deaths/ ADP	Deaths/ ADP	Deaths/ ADP	Deaths/ ADP
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treat	-1.333***	-1.246**	0.275*	0.264*	-0.552	-0.373	-3.623**	-3.472**
	(0.438)	(0.495)	(0.141)	(0.142)	(1.761)	(1.719)	(1.499)	(1.435)
RI p-value	.025	.038	.096	.08	.75	.822	.033	.053
Control Mean	2.455	2.455	0.545	0.545	4.201	4.201	5.060	5.060
Observations	46	46	46	46	46	46	46	46
Baseline Control		✓		✓		✓		✓

Notes: This table reports estimates of the treatment effect of accreditation assignment on mortality obtained from estimating Equations 1 and 2. Data are from administrative death logs. The outcome variable in columns (1) and (2) is the total number of deaths that occurred in the twelve months prior to each facility's medical audit. The outcome variable in columns (3) and (4) is an indicator equal to 1 if the number of deaths that occurred in the facility in the twelve months prior to the medical audit is less than or equal to 1. The outcome variable in columns (5)–(8) is the total number of deaths that occurred in the twelve months prior to the medical audit, divided by facility ADP at endline and multiplied by 1,000. Columns (2), (4), (6), and (8) control for baseline deaths, which are from media reports. For columns (2), (6), and (8), the baseline is the total number of deaths in the twelve months prior to accreditation assignment, divided by facility ADP at baseline and multiplied by 1,000. For column (4), the baseline measure is an indicator equal to 1 if the number of deaths that occurred in the facility in the twelve months prior to accreditation assignment is less than or equal to 1. Columns (1) and (2) report average marginal effects from Poisson regressions and the corresponding standard errors, using the logarithm of endline ADP as an offset variable. All specifications include randomization strata fixed effects. Robust standard errors are reported in parentheses. $*$, $**$, and $***$ denote statistical significance at the 10, 5, and 1 percent levels, respectively.

Appendix

Table of Contents

A Additional Figures and Tables **A.2**

B Model of Accreditation **A.27**

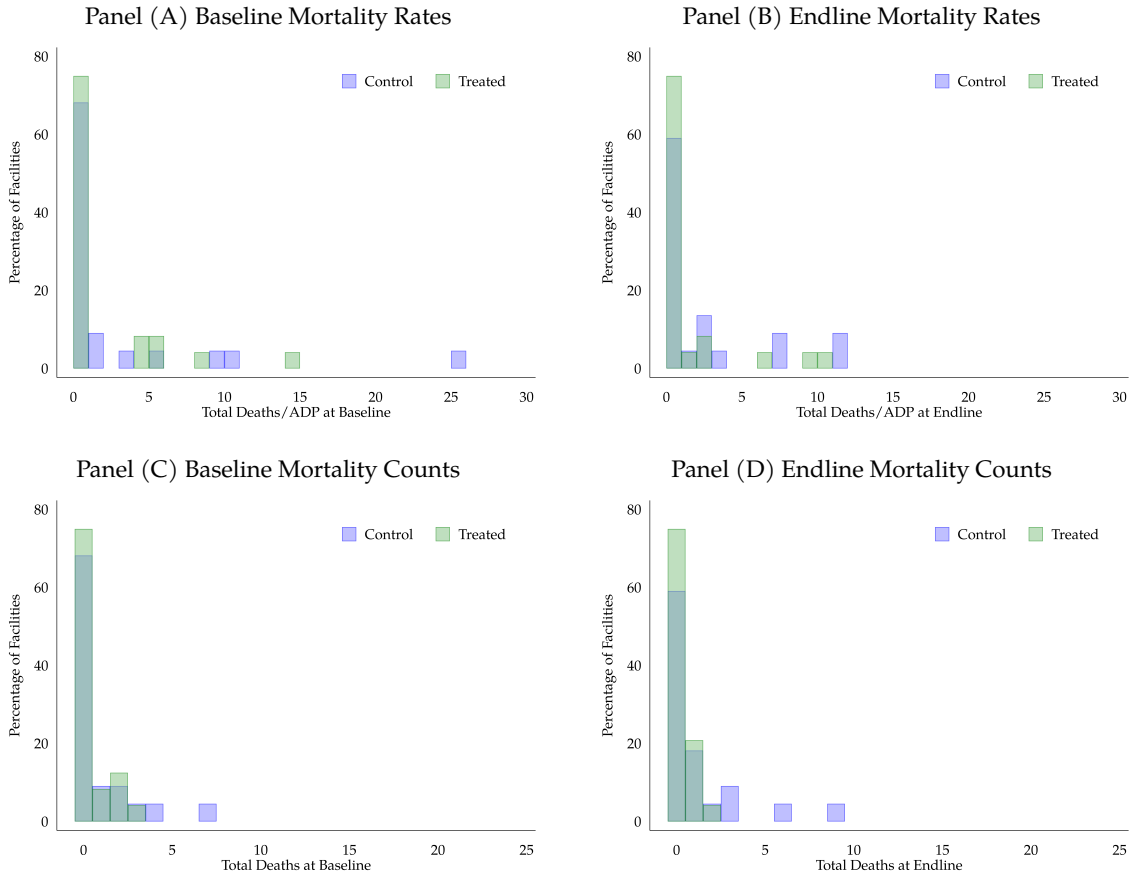
C Additional Study Design and Implementation Details **A.29**
C.1 Recruitment A.29
C.2 Enrollment A.31
C.3 Randomization and Stratification A.32
C.4 Index of Items A.33

D Description of Data Elements **A.34**
D.1 Facility Survey A.34
D.2 Staff Survey A.34
D.3 Medical Audit A.34
D.4 Leadership Interviews A.35
D.5 Inmate Interviews A.35
D.6 Expert Prediction Survey A.35

E Coding and Construction of Indices **A.38**
E.1 Coding Committees and Codebook A.38
E.2 Variable Transformation A.38
E.3 Coding Outcomes A.39
E.4 Coding Facility Survey A.39
E.5 Coding Staff Survey A.40
E.6 Coding Medical Audits A.40
E.7 Coding Media-Reported Deaths A.40
E.8 Construction of Indices A.40

A Additional Figures and Tables

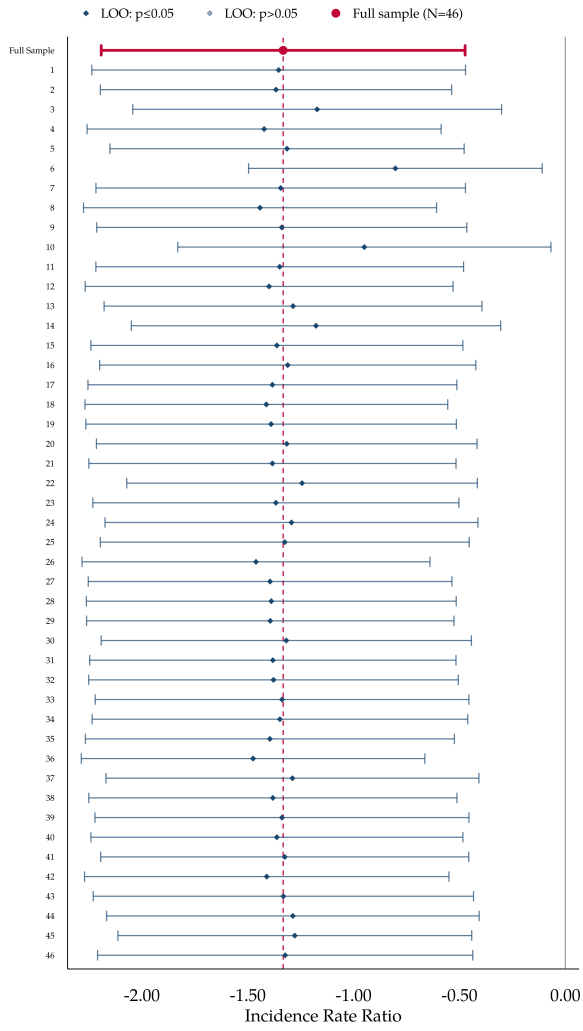
Appendix Figure A1: Distribution of Media Deaths



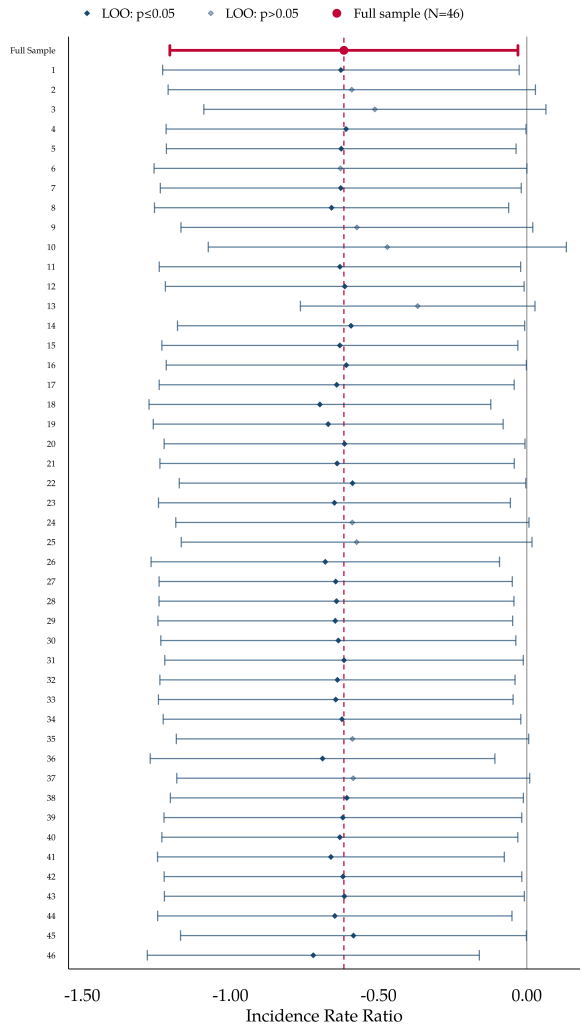
Notes: This figure shows the distribution of deaths according to media reports using mortality rates (Panels (A) and (B)) and counts (Panels (C)-(D)). Panels (A) and (C) report baseline values and Panels (B) and (D) report endline values. Counts are the total deaths in the 12-month period prior to or after accreditation assignment. Mortality rates are computed as the total deaths over the specified period, divided by facility ADP in the corresponding period and multiplied by 1,000.

Appendix Figure A2: Leave-One-Out Robustness for Mortality

Panel (A) Facility-Reported Deaths

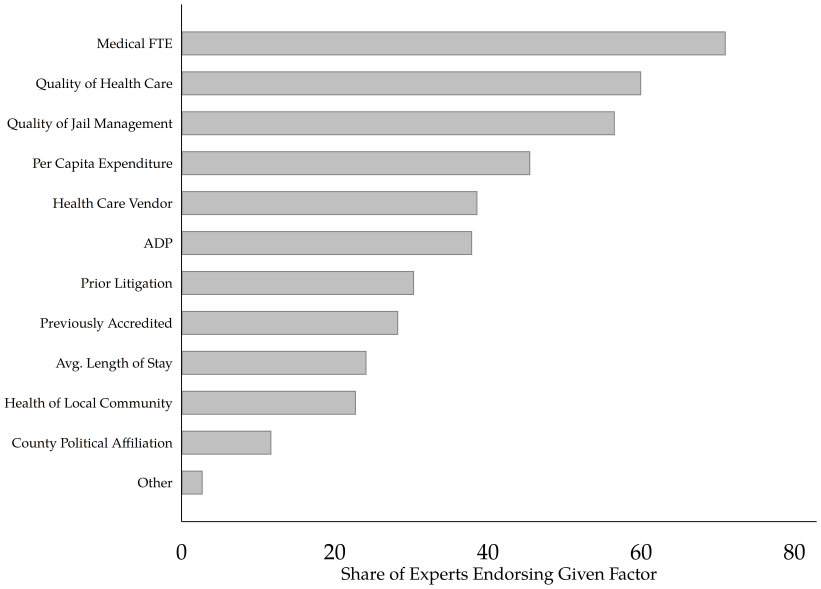


Panel (B) Media-Reported Deaths



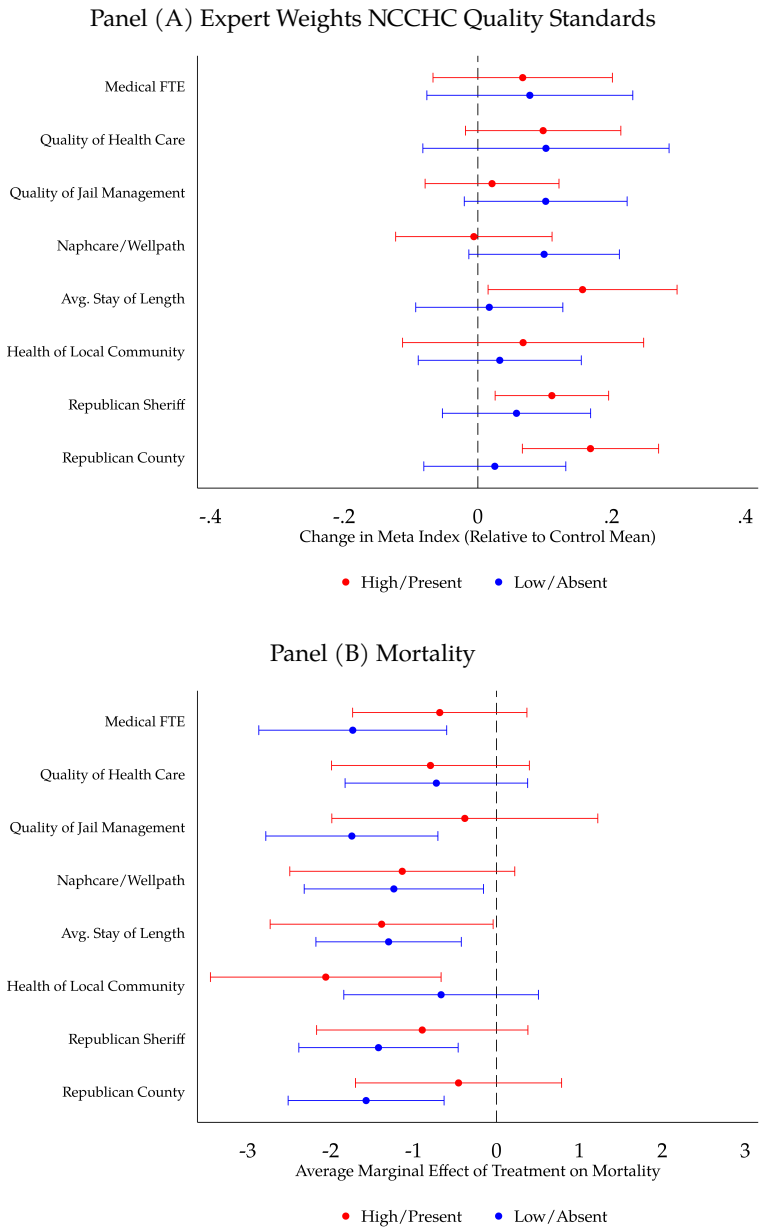
Notes: This figure presents treatment effects of accreditation assignment on mortality from Equation 2 excluding one facility at a time. Data come from facility-reported deaths for Panel (A) and from media-reported deaths for Panel (B). The outcome variable is the total number of deaths that occurred in the twelve months prior to each facility's medical audit (Panel (A)) or after accreditation assignment (Panel (B)). Average marginal effects from Poisson regression and corresponding 95 percent CIs are reported. The red marker and dashed line show the full-sample estimate. All regressions include randomization strata fixed effects.

Appendix Figure A3: Expert Predictions and Heterogeneous Treatment Effects



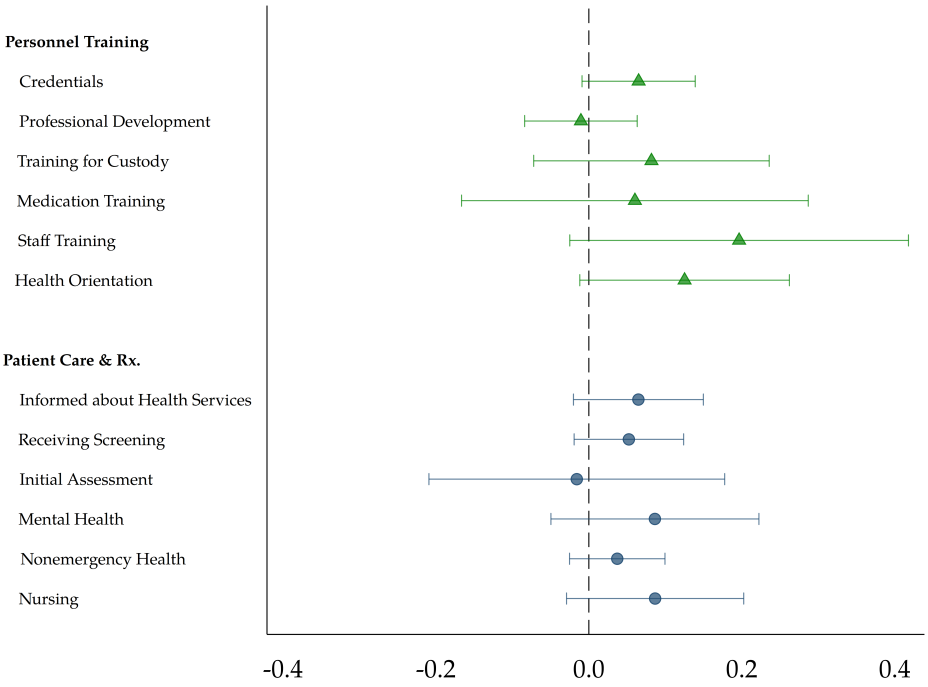
Notes: Data are from the expert survey. Respondents were asked "Which of the following factors do you think will impact whether NCCHC accreditation of jail health care services will affect the outcomes of interest? Recall, our outcomes of interest include health care procedures, health outcomes, staff outcomes, and litigation against jails." The figure plots the share of all respondents choosing a given factor. Respondents could select up to five factors listed in the graph.

Appendix Figure A4: Heterogeneous Treatment Effects



Notes: This figure shows heterogeneity in the treatment effects of accreditation assignment on NCCHC quality standards (Panel (A)) and mortality (Panel (B)). The outcome variable in Panel (A) is the expert-weighted meta-index of quality standards. The outcome variable in Panel (B) is the number of facility-reported deaths in the 12 months prior to the medical audit. For each heterogeneity factor predicted by experts, estimates come from a pooled specification that interacts accreditation assignment with an indicator for whether the factor equals 1, while controlling for the factor main effect, randomization strata fixed effects, and a baseline outcome control for Panel (A). The reported coefficients and 95 percent confidence intervals correspond to the implied treatment effects for facilities with the factor equal to 0 and 1, respectively, from Equation 1 in Panel (A) and Equation 2 in Panel (B). Panel (A) plots OLS coefficient and 95 percent CIs, while Panel (B) plots average marginal effects from Poisson regressions and correspondent 95 percent CIs. An indicator equal to 1 for the heterogeneity factor corresponds to having the respective variable above the sample median for continuous variables, or the respective binary indicator equal to 1 for binary variables. Quality of Health Care is constructed using the staff-survey question "Rate the quality of medical services for inmates at this facility" (asked on a 1-5 Likert scale from "Very Poor" to "Excellent", and averaging responses by facility). Quality of Jail Management is measured through the Vertical Coordination index from the staff survey (see Section IV.1).

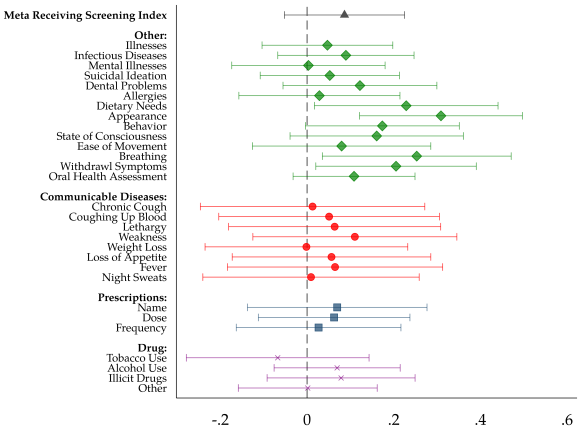
Appendix Figure A5: Effect of Accreditation Assignment on Sub-Standards



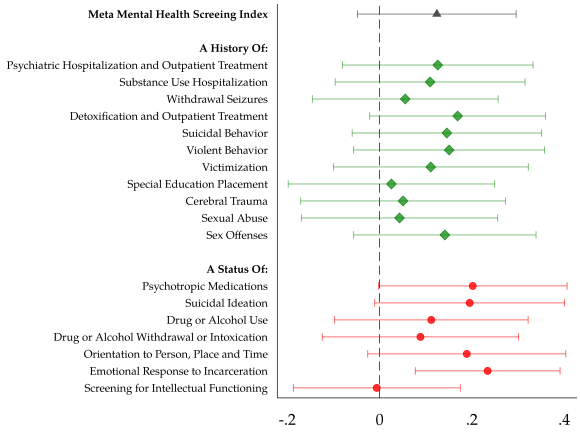
Notes: This figure shows treatment effects of accreditation assignment on each subsection under NCCHC quality standards Personnel Training and Patient Care and Treatment from estimating Equation 1 (for detailed definitions of the standards and their calculation see Appendix Section E.8). Coefficients and 95 percent confidence intervals are plotted. Under Personnel Training Standard, the "Credentials" subsection refers to "The facility's qualified health care professionals are legally eligible to perform their clinical duties". "Professional Development" refers to "The facility's qualified health care professionals maintain current clinical knowledge and skills". "Training in Custody" subsection refers to "Correctional officers are trained to recognize the need to refer an inmate to a qualified health care professional". "Medication Training" subsection refers to "Personnel who administer or deliver prescription medication are appropriately trained". "Staff Training" subsection refers to "Does the provided documentation show that there is a documented plan for custody staff to follow when a health situation arises and health staff are not present". "Health Orientation" subsection refers to "Health staff are properly acclimated to work in the correctional environment and understand their roles and responsibilities". Under Patient Care and Rx. Standard, "Informed about Health Services" refers to "Upon arrival at the facility, inmates are informed of the availability of health care services and how to access them". "Receiving Screening" subsection refers to "Screening is performed on all inmates upon arrival at the intake facility to ensure that emergent and urgent health needs are met". "Initial Assessment" subsection refers to "Inmates receive initial health assessments". "Mental Health" subsection refers to "Mental health screening is performed to ensure that urgent mental health needs are met". "Nonemergency Health" refers to "Inmates' nonemergent health care needs are met". "Nursing" subsection refers to "Nursing assessment protocols and procedures are appropriate to the level of competency and preparation of the nursing personnel and comply with the relevant state practice acts". All regressions include baseline controls and randomization strata fixed effects.

Appendix Figure A6: Effect of Accreditation Assignment on Screening

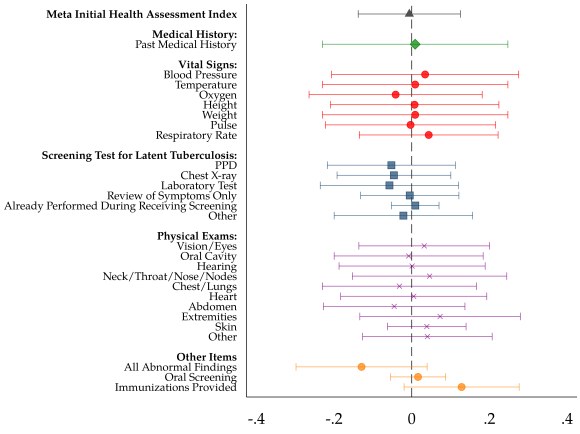
Panel (A) Receiving Screening



Panel (B) Mental Health Screening



Panel (C) Initial Health Assessment



Notes: The figure shows treatment effects of accreditation assignment on each component under subsections "Receiving Screening" (Panel (A)), "Mental Health Screening" (Panel (B)), and "Initial Health Assessment" (Panel (C)) from estimating Equation 1 at the patient level. Two medical records per facility were audited in the endline medical audit, each of which refers to a distinct patient. Coefficients and 95 percent confidence intervals are plotted. In each figure, the black triangle represents the estimate for meta index by taking the average of all the components within the subsection. All regressions include baseline controls and randomization strata fixed effects, and use standard errors clustered at the facility level.

Appendix Figure A7: Word Cloud of Jail Leadership Interviews Regarding Experience with Accreditation

Panel (A) Experience Undergoing Accreditation



Panel (B) Challenges Faced During Accreditation



Panel (C) Biggest Change Implemented



Panel (D) Relationship between Custody and Medical Staff



Notes: Data are from qualitative interviews. This figure shows word clouds of interviews with jail leadership of accredited facilities who were asked 4 questions about their experience undergoing accreditation. The specific questions/stems included: Panel (A) "Please describe your facility's experience undergoing NCCHC accreditation process."; Panel (B) "Were there any challenges your facility faced during the NCCHC accreditation process? If so, what were they?"; Panel (C) "In your opinion, what is the biggest change your facility has implemented as a result of the NCCHC accreditation process?"; Panel (D) "Do you think that the relationship and coordination between correctional and health care staff has changed due to the NCCHC accreditation process? If so, how has it changed?". All word clouds exclude common filler words (i.e., yes, any, thing, every, really, though). Additionally, question-specific words are filtered from each panel to avoid repetition of parts of the question such as "accreditation" in Panel (A), "challenge" in Panel (B), "changes" in Panel (C), and "relationship" in Panel (D).

Appendix Table A1: Essential and Important Subsections

Category	Subsections
Essential Standards	A01 - Access to Care
	A02* - Responsible Health Authority
	A03 - Medical Autonomy
	A04* - Administrative Meetings and Reports
	A05* - Policies and Procedures
	A06* - Continuous Quality Improvement Program
	A08* - Health Records
	B02* - Infectious Disease Prevention and Control
	B03* - Clinical Preventive Services
	B05* - Suicide Prevention and Intervention
	B07 - Communication on Patients' Health Needs
	C01* - Credentials
	C03* - Professional Development
	C04* - Health Training for Correctional Officers
	C05* - Medication Administration Training
	C06 - Inmate Worker
	C08 - Health care Liaison
	D01 - Pharmaceutical Operations
	D02* - Medication Services
	D05* - Medical Diets
	D07* - Emergency Services and Response Plan
	D08 - Hospital and Specialty Care
	E01* - Information on Health Services
	E02* - Receiving Screening
	E04* - Initial Health Assessment
	E05* - Mental Health Screening and Evaluation
	E07* - Nonemergency health care Requests and Services
	E09 - Continuity, Coordination, and Quality of Care During Incarceration
	E10 - Discharge Planning
	F01* - Patients with Chronic Disease and Other Special Needs
	F02 - Infirmary-Level Care
	F03* - Mental Health Services
	F04* - Medically Supervised Withdrawal and Treatment
	F05* - Counseling and Care of the Pregnant Inmate
	F06* - Response to Sexual Abuse
	G01* - Restraint and Seclusion
	G02* - Segregated Inmates
	G03* - Emergency Psychotopic Medication
Important Standards	A07* - Privacy of Care
	A09* - Procedure in the Event of an Inmate Death
	A10* - Grievance Process for health care Complaints
	B01* - Healthy Lifestyle Promotion
	B04 - Medical Surveillance of Inmate Workers
	B06* - Contraception
	B08 - Patient Safety
	B09 - Staff Safety
	C02 - Clinical Performance Enhancement
	C07* - Staffing
	C09* - Orientation for Health Staff
	D03* - Clinic Space, Equipment, and Supplies
	D04* - On-Site Diagnostic Services
	D06 - Patient Escort
	E03 - Transfer Screening
	E06 - Oral Care
	E08* - Nursing Assessment Protocols and Procedures
	F07* - Care for the Terminally Ill
	G04* - Therapeutic Relationship, Forensic Information, and Disciplinary Actions
	G05 - Informed Consent and Right to Refuse
	G06 - Medical and Other Research

Notes: This table shows the NCCHC classifications of subsections into Essential or Important Standards. The asterisk indicates whether the subsection is measured using the Harvard study metrics.

Appendix Table A2: Comparison of Jails that Enrolled vs. Not Enrolled

	Mean	Coefficient	SE
	(1)	(2)	(3)
Mortality	0.431	-0.507	(0.423)
Mortality/ADP	1.106	-0.208	(0.762)
2019 ADP	636.817	-403.919	(249.617)
Has Health Vendor	0.776	0.138	(0.151)
NaphCare/Wellpath	0.172	0.112	(0.101)
For-Profit Health Vendor	0.655	0.301*	(0.160)
Sheriff Up For Election	0.483	-0.022	(0.165)
Anticipated Vendor Change	0.140	-0.220	(0.202)
Enrolled Facilities		46	
Excluded Facilities		12	

Notes: This table compares enrolled facilities with those that declined participation to the study. Column (1) reports the sample mean. Column (2) reports the coefficient on an indicator equal to one if the jail enrolled in the experiment. Column (3) reports the associated robust standard error. Mortality refers to media-reported deaths in 2020. Mortality/ADP is calculated as the number of deaths in 2020 divided by facility-reported ADP in 2020 and multiplied by 1,000. For-profit Health Vendor is an indicator equal to 1 if the health vendor is a private company. Sheriff Up For Election is an indicator equal to 1 if the facility Sheriff was up for reelection in 2021 or 2022. Anticipated Vendor Change is an indicator equal to 1 if the facility anticipated any changes to its healthcare services provision at the time of screening. $*$, $**$, and $***$ refer to the statistical significance at 10, 5, and 1 percent level, respectively.

Appendix Table A3: Subsidies for Accreditation by Treatment Status

	Treatment	Control
Subsidy for Accreditation	$7,028.24	$2,181.82
Survey Incentives	$804.35	$1,104.55
Total Amount	$7,832.59	$3,286.36

Notes: This table reports subsidies and incentives provided to facilities by treatment status. "Subsidy for Accreditation" includes the total contribution towards accreditation. For both groups, "Survey Incentives" includes incentives for finishing facility survey and staff survey. For facility surveys, facilities can opt-in to receive $500 in direct payment via check or direct deposit, or can credit the payment towards their accreditation or re-accreditation at endline. For staff surveys, staff who complete the survey are entered into a lottery for a $100 gift card. Additional details can be found in Appendix Section E.5. "Total Amount" is the sum of "Subsidy for Accreditation" and "Survey Incentives". The difference between Treatment and Control for total amount is statistically significant (p-value<0.001).

Appendix Table A4: Effect of Accreditation Assignment on Completion of Endline Instruments

	Endline Study Components				
	Facility Survey	Medical Audits	Death Logs	Staff Survey	Qualitative Interviews
	(1)	(2)	(3)	(4)	(5)
Treat	–	0.165	–	-0.009	0.083
	–	(0.115)	–	(0.087)	(0.138)
Control Mean	1.000	0.727	1.000	0.909	0.682
Observations	46	46	46	46	46
Strata FE	✓	✓	✓	✓	✓

Notes: This table shows the treatment effects of accreditation assignment on indicators for successful completion of each specific study component across columns (1) – (5). The indicators for successful completion of the facility survey and successful completion of death logs are 1 for all facilities in the sample as all facilities completed these study components. The indicator for successful completion of interviews is equal to 1 if the facility completed either the inmate or the jail leadership interview. All regressions include randomization strata fixed effects. Robust standard errors are reported in parentheses. $*$, $**$, and $***$ refer to the statistical significance at 10, 5, and 1 percent level, respectively.

Appendix Table A5: Effect of Accreditation Assignment on NCCHC Quality Standards

	Control	ITT	p-value	
	Mean	(T−C)	Classical	RI
Panel (A) Individual Standards				
Government Administration	0.550	0.014	0.521	0.526
		(0.021)		
Safety and Prevention	0.522	0.052	0.155	0.107
		(0.036)		
Personnel Training	0.538	0.083*	0.071	0.083
		(0.044)		
Ancillary Services	0.520	0.024	0.252	0.235
		(0.021)		
Patient Care and Rx	0.535	0.066**	0.032	0.035
		(0.030)		
Special Needs	0.593	0.011	0.815	0.794
		(0.048)		
Medical–Legal	0.568	0.065	0.219	0.204
		(0.052)		
Panel (B) Composite Indices				
Simple Mean	0.547	0.045*	0.093	0.079
		(0.026)		
Principal Component	0.556	0.074*	0.076	0.067
		(0.040)		
Expert Weights	0.544	0.049*	0.072	0.066
		(0.027)		
Observations		46		
Control group		22		
Treatment group		24		

Notes: This table shows the treatment effect of accreditation assignment on the seven NCCHC quality standards from estimating Equation 1, both individually (Panel (A)) and aggregated into composite indices (Panel (B)). Quality standards are created using data from the baseline/endline facility survey and look-back medical audit. For detailed definitions of the standards and their calculation, see Appendix Section E.8. The composite indices are constructed by combining the seven quality standards in three pre-specified ways as described in the main text: equal weights, the first component using principal component analysis, and reciprocal rank weights using our expert survey responses. Treatment effects (ITT) are obtained from estimating Equation 1, including controls for the baseline mean of the respective outcome variable at the facility level and randomization strata fixed effects. Robust standard errors are reported in parentheses. p-values based on robust standard errors and randomization inference (RI) are shown. $*$, $**$, and $***$ refer to the statistical significance at 10, 5, and 1 percent level, respectively.

Appendix Table A6: Effect of Accreditation Assignment on Meta-Index by Expert Type

	Control	ITT	p-value	
	Mean	(T−C)	Classical	RI
Formerly Incarcerated	0.545	0.051*	0.066	0.063
		(0.027)		
All Other Experts	0.544	0.049*	0.074	0.066
		(0.027)		
Observations		46		
Control group		22		
Treatment group		24		

Notes: This table shows the treatment effect of accreditation on the meta-index of quality standards from estimating Equation 1 using weights separately calculated for the formerly incarcerated and all other experts. Detailed definitions of the standards and expert weights can be found in Appendix E.8. Treatment effects (ITT) are obtained from estimating Equation 1, including controls for the baseline mean of the respective outcome variable at the facility level and randomization strata fixed effects. Robust standard errors are reported in parentheses. p-values based on robust standard errors and randomization inference (RI) are shown. $*$, $**$, and $***$ refer to the statistical significance at 10, 5, and 1 percent level, respectively.

Appendix Table A7: Effect of Accreditation Assignment on Production Function Indices

	Control	ITT	p-value	
	Mean	(T−C)	Classical	RI
Panel (A) Individual Outcomes				
Inputs	0.520	-0.003	0.886	0.874
		(0.021)		
Formal Procedures	0.610	0.041	0.143	0.122
		(0.027)		
Procedural Outcomes (Administrative)	0.465	0.031	0.331	0.341
		(0.031)		
Procedural Outcomes (Inmate Health)	0.486	0.060**	0.056	0.046
		(0.030)		
Medical Outcomes	0.341	0.064	0.267	0.239
		(0.056)		
Panel (B) Composite Indices				
Simple Mean	0.484	0.040	0.123	0.112
		(0.025)		
Principal Component	0.552	0.071*	0.088	0.079
		(0.041)		
Expert Weights	0.475	0.041	0.138	0.120
		(0.027)		
Observations		46		
Control group		22		
Treatment group		24		

Notes: This table shows the treatment effect of accreditation assignment on the five production function indices from estimating Equation 1, both individually (Panel (A)) and aggregated into composite indices (Panel (B)). Production function indices are created using data from the baseline/endline facility survey and look-back medical audit. For detailed definitions of the standards and their calculation, see Appendix Section E.8. The composite indices are constructed by combining the five production function indices in three pre-specified ways as described in the main text: equal weights, the first component using principal component analysis, and reciprocal rank weights using our expert survey responses. Treatment effects (ITT) are obtained from estimating Equation 1, including controls for the baseline mean of the respective outcome variable at the facility level and randomization strata fixed effects. Robust standard errors are reported in parentheses. p-values based on robust standard errors and randomization inference (RI) are shown. $*$, $**$, and $***$ refer to the statistical significance at 10, 5, and 1 percent level, respectively.

Appendix Table A8: Effect of Accreditation Assignment on Mortality (Media-Reported Deaths)

	Count Model		Linear Model					
	Poisson		LPM		OLS		WLS	
	Total Deaths	Total Deaths	≤ 1 Deaths	≤ 1 Deaths	Deaths/ ADP	Deaths/ ADP	Deaths/ ADP	Deaths/ ADP
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treat	-0.617**	-0.496**	0.193*	0.182*	-0.805	-0.809	-1.675*	-1.523**
	(0.300)	(0.228)	(0.107)	(0.104)	(1.111)	(1.103)	(0.912)	(0.737)
RI p-value	.087	.144	.086	.095	.48	.486	.097	.079
Control Mean	1.227	1.227	0.773	0.773	2.232	2.232	2.440	2.440
Observations	46	46	46	46	46	46	46	46
Baseline Control		✓		✓		✓		✓

Notes: This table reports estimates of the treatment effect of accreditation assignment on mortality obtained from estimating Equations 1 and 2. Data are from media-reported deaths. The outcome variable for columns (1) and (2) is the total number of deaths that occurred in the twelve months after accreditation assignment (Panel (B)). The outcome variable for columns (3) and (4) is an indicator equal to 1 if the number of deaths that occurred in the facility in the twelve months after accreditation assignment is less than or equal to 1. The outcome variable for columns (5)-(8) is the total number of deaths that occurred in the twelve months after accreditation assignment, divided by facility ADP at endline and multiplied by 1,000. Columns (2), (4), (6), and (8) control for baseline deaths. For columns (2), (6), and (8), the baseline is the total number of deaths in the twelve months before accreditation assignment, divided by facility ADP at baseline and multiplied by 1,000. For column (4), the baseline measure is an indicator equal to 1 if the number of deaths that occurred in the facility in the twelve months before accreditation assignment is less than or equal to 1. Columns (7)-(8) weight the observations by ADP at baseline. Columns (1) and (2) report average marginal effects from Poisson regression and corresponding standard errors, and use the logarithm of endline ADP as offset variable. All specifications include randomization strata fixed effects. Robust standard errors are reported in parentheses. $*$, $**$, and $***$ refer to the statistical significance at 10, 5, and 1 percent level, respectively.

Appendix Table A9: Effect of Accreditation Assignment on Mortality by Cause

	Illness	Suicide & Homicide	Undetermined	Drug Overdose
	(1)	(2)	(3)	(4)
Treat	-0.583***	-0.108	-0.604*	-0.085
	(0.155)	(0.203)	(0.314)	(0.103)
Control Mean	0.682	0.500	1.045	0.227
Observations	46	46	46	46

Notes: This table reports Poisson estimates of the treatment effect of accreditation assignment on mortality, by cause of death. Data are from administrative death logs. The outcome variable is the total number of deaths that occurred for each cause in the twelve months prior to each facility's medical audit. Average marginal effects from Poisson regression and corresponding standard errors are reported. All specifications use the logarithm of endline ADP as an offset variable, and include randomization strata fixed effects. Robust standard errors are reported in parentheses. $*$, $**$, and $***$ refer to the statistical significance at 10, 5, and 1 percent level, respectively.

Appendix Table A10: Effect of Accreditation Assignment on Mortality by Cause

	Illness	Suicide & Homicide	Undetermined	Drug Overdose
	(1)	(2)	(3)	(4)
Treat	-1.555**	-0.281	-1.847*	-0.224
	(0.655)	(0.617)	(1.040)	(0.365)
Control Mean	1.327	1.044	2.251	0.437
Observations	46	46	46	46

Notes: This table reports OLS estimates of the treatment effect of accreditation assignment on mortality, by cause of death. Data are from administrative death logs. The outcome variable is the number of deaths that occurred for each cause in the twelve months prior to each facility's medical audit, divided by facility ADP at endline and multiplied by 1,000. All specifications control for baseline deaths, i.e. the number of deaths that occurred for each cause in the twelve months before accreditation assignment, divided by facility ADP at baseline and multiplied by 1,000. All specifications weight by ADP at baseline and include randomization strata fixed effects. Robust standard errors are reported in parentheses. $*$, $**$, and $***$ refer to the statistical significance at 10, 5, and 1 percent level, respectively.

Appendix Table A11: Heterogeneity of Treatment Effects on Mortality

	NCCHC All Experts	
	(1)	(2)
Quality Standards (Expert Weights)		-0.246
		(0.770)
Treat	-1.333***	0.871
	(0.438)	(1.257)
Treat * Quality Standards		-2.833**
		(1.258)
Control Mean	2.455	2.455
Observations	46	46
Baseline Control		✓

Notes: This table reports estimates for heterogeneity in treatment effects of accreditation assignment on mortality. Mortality data are from administrative death logs. The outcome variable is the number of facility-reported deaths in the 12 months prior to the medical audit. Column (1) reports the main effect of accreditation on mortality from a Poisson model with an exposure adjustment for endline ADP. Column (2) reports heterogeneity in the effect of accreditation from a Poisson model that interacts treatment with an indicator for the endline expert-weighted meta index of NCCHC quality standards being above median. In Column (2), the reported coefficient for Treat is the subgroup-specific marginal effect of accreditation for facilities below the median, while the interaction coefficient reports the difference in the treatment effect for facilities above the median relative to facilities below the median. Standard errors for the marginal treatment effect are obtained from the margins computation. The baseline control in Column (2) is a binary indicator for the baseline expert-weighted meta index of NCCHC quality standards being above median. All specifications include randomization strata fixed effects. Robust standard errors are reported in parentheses. $*$, $**$, and $***$ refer to the statistical significance at 10, 5, and 1 percent level, respectively.

Appendix Table A12: Effect of Accreditation Assignment on NCCHC Quality Standards Constructed Using Facility Survey Only, No Medical Audit

	Control	ITT	p-value	
	Mean	(T−C)	Classical	RI
Panel (A) Individual Standards				
Government Administration	0.548	0.009	0.648	0.660
		(0.020)		
Safety and Prevention	0.529	0.059	0.114	0.078
		(0.036)		
Personnel Training	0.538	0.083*	0.071	0.083
		(0.044)		
Ancillary Services	0.556	0.025	0.154	0.160
		(0.017)		
Patient Care and Rx	0.575	0.041*	0.063	0.095
		(0.021)		
Special Needs	0.660	0.063	0.271	0.252
		(0.056)		
Medical–Legal	0.568	0.065	0.219	0.204
		(0.052)		
Panel (B) Composite Indices				
Simple Mean	0.568	0.050*	0.062	0.052
		(0.026)		
Principal Component	0.532	0.063*	0.100	0.099
		(0.037)		
Expert Weights	0.567	0.052**	0.053	0.048
		(0.026)		
Observations		46		
Control group		22		
Treatment group		24		

Notes: This table shows the treatment effect of accreditation assignment on the seven NCCHC quality standards from estimating Equation 1, both individually (Panel (A)) and aggregated into composite indices (Panel (B)). Quality standards are created using only data from the baseline/endline facility survey. For detailed definitions of the standards and their calculation, see Appendix Section E.8. The composite indices are constructed by combining the seven quality standards in three pre-specified ways as described in the main text: equal weights, the first component using principal component analysis, and reciprocal rank weights using our expert survey responses. Treatment effects (ITT) are obtained from estimating Equation 1, including controls for the baseline mean of the respective outcome variable at the facility level and randomization strata fixed effects. Robust standard errors are reported in parentheses. p-values based on robust standard errors and randomization inference (RI) are shown. $*$, $**$, and $***$ refer to the statistical significance at 10, 5, and 1 percent level, respectively.

Appendix Table A13: Balance Table on Booked Inmates

	Control Mean	Treatment Mean	p-value
	(1)	(2)	(3)
Panel (A)* Charges and Length of Stay			
Number of Charges	4.195	4.446	0.133
Length of Stay (Days)	92.408	83.594	0.507
Panel (B) Race			
White	0.439	0.493	0.105
Black	0.279	0.290	0.976
Hispanic	0.179	0.189	0.055
Other	0.029	0.018	0.758
Unknown	0.074	0.010	0.095
Panel (C) Sex			
Male	0.886	0.841	0.531
Female	0.113	0.157	0.580
Panel (D) Age			
Age	36.237	36.804	0.520
Number of Inmates	6498	3447	

Notes: This table shows the balance test on the inmates booked into facilities at endline between control and treatment facilities in the 34 facilities with booking records. Due to data availability, information in Panel (A) is available for 4 facilities. Column 3 reports the p-value, which is obtained by regressing the outcome variable on a treatment indicator, controlling for randomization strata fixed effects and clustering standard errors at the facility level. $*$, $**$, and $***$ refer to the statistical significance at 10, 5, and 1 percent level, respectively.

Appendix Table A14: Effect of Accreditation Assignment on Mortality Within and Outside of Jails

	Jail-Only		Non-Jail		All Deaths	
	Total Deaths	Total Deaths	Total Deaths	Total Deaths	Total Deaths	Total Deaths
	(1)	(2)	(3)	(4)	(5)	(6)
Treat	-0.617**	-0.496**	-0.625**	-0.514*	-1.241***	-0.935**
	(0.300)	(0.228)	(0.277)	(0.288)	(0.456)	(0.390)
RI p-value	0.087	0.144	0.070	0.117	0.035	0.071
Control Mean	1.227	1.227	1.091	1.091	2.318	2.318
Observations	46	46	46	46	46	46
Baseline Control		✓		✓		✓

Notes: This table reports estimates of the treatment effect of accreditation assignment on mortality within and outside of jails using the estimating Equation 2. Data are from media-reported deaths. The outcome variable is the total number of deaths in the twelve months following accreditation assignment, disaggregated by location. Columns (1) and (2) report deaths occurring exclusively in jail. Columns (3) and (4) report non-jail deaths (those occurring in hospitals or with location undetermined). Columns (5) and (6) report the total number of deaths occurring in either location over the same period. Columns (2), (4), and (6) control for the baseline death rate, which is the total number of deaths in the twelve months before accreditation assignment, divided by facility Average Daily Population (ADP) at baseline and multiplied by 1000. All columns report average marginal effects from Poisson regressions and corresponding standard errors, using the logarithm of endline ADP as an offset variable. All specifications include randomization strata fixed effects. Robust standard errors are reported in parentheses. $*$, $**$, and $***$ refer to the statistical significance at 10, 5, and 1 percent level, respectively.

Appendix Table A15: Effect of Accreditation Assignment on Staff Sentiment Indices

	Control	ITT	p-value	
	Mean	(T−C)	Classical	RI
Job Satisfaction	0.671	0.028**	0.048	0.097
		(0.014)		
Horizontal Coordination	0.643	0.047**	0.035	0.087
		(0.022)		
Vertical Coordination	0.649	0.029	0.248	0.243
		(0.025)		
Respect Inmates	0.754	-0.008	0.669	0.683
		(0.018)		
Observations		674		
Control group		370		
Treatment group		304		

Notes: This table shows the treatment effect of accreditation assignment on staff sentiment indices from estimating Equation 1. Staff sentiment indices are created using data from the baseline/endline staff survey. Endline responses were obtained from 41 jails and missing endline responses were imputed with baseline values for the other 5 facilities in our sample. The Job Satisfaction index includes: (1) "I received the training I need to do my job well"; (2) "I have the tools and resources I need to do my job well"; (3) "I am satisfied with my job"; (4) "I find my job to be meaningful"; (5) "I feel valued at my job"; (6) "I find my job to be challenging"; (7) "I would recommend my job to others." The Horizontal Coordination index includes: (1) "Custody staff support the implementation of clinical decisions and partner with clinical staff"; (2) "Custody and clinical team staff work collaboratively to give good health care to inmates"; (3) "There is open communication between custody and health staff." The Vertical Coordination index includes: (1) "Staff feel comfortable speaking up when they see something that may negatively affect inmate care"; (2) "When staff speak up about inmate safety and care decisions, those with greater authority are open to their concerns"; (3) "My supervisor considers staff suggestions for improving inmate health and safety"; (4) "My supervisor considers staff suggestions for improving staff health and safety"; (5) "My supervisor takes direct actions to address staff concerns for improving inmate health and safety"; (6) "My supervisor takes direct actions to address staff concerns for improving staff health and safety." Questions (1) and (2) are only asked to health staff. The Respect Inmates index includes: (1) "Inmates at this facility are treated with respect." These questions are answered on a Likert scale and renormalized on a scale from 0 to 1 with higher values representing stronger agreement. Treatment effects (ITT) are obtained from estimating Equation 1, including controls for the baseline mean of the respective outcome variable at the facility level and randomization strata fixed effects. Standard errors are clustered at the facility level and reported in parentheses. p-values based on cluster standard errors and randomization inference (RI) are shown. $*$, $**$, and $***$ refer to the statistical significance at 10, 5, and 1 percent level, respectively.

Appendix Table A16: Effect of Accreditation Assignment on Components of Staff Sentiment Indices

	Control	ITT	p-value	
	Mean	(T−C)	Classical	RI
Panel (A) Horizontal Coordination Index				
Custody and Health Communicate	0.631	0.051**	0.041	0.109
		(0.024)		
Custody and Health Coordinate	0.657	0.039*	0.088	0.187
		(0.022)		
Custody Support Health's Decisions	0.641	0.053**	0.030	0.049
		(0.024)		
Panel (B) Vertical Coordination Index				
Staff Comfortable Speaking Up on Inmate Care	0.696	0.067**	0.055	0.070
		(0.034)		
Authority Open to Concerns on Inmate Care	0.648	0.089***	0.004	0.033
		(0.029)		
Supr. Open to Concerns on Inmate	0.656	0.011	0.710	0.725
		(0.029)		
Supr. Open to Concerns on Staff	0.629	0.028	0.314	0.314
		(0.027)		
Supr. Consider Suggestions on Inmate	0.670	0.028	0.386	0.317
		(0.032)		
Supr. Consider Suggestions on Staff	0.632	0.049*	0.057	0.088
		(0.025)		
Panel (C) Satisfaction Index				
Properly Trained	0.725	0.022	0.322	0.421
		(0.022)		
Sufficient Tools	0.664	0.043*	0.062	0.095
		(0.023)		
Satisfied	0.666	0.030	0.158	0.191
		(0.021)		
Meaningful	0.733	0.012	0.586	0.692
		(0.022)		
Valued	0.570	0.054*	0.067	0.120
		(0.029)		
Challenging	0.745	0.004	0.863	0.883
		(0.021)		
Recommend to Others	0.596	0.039*	0.095	0.158
		(0.023)		
Observations		674		
Control group		370		
Treatment group		304		

Notes: This table shows the treatment effect of accreditation on the components of staff sentiment indices from estimating Equation 1. Endline responses were obtained for 41 jails and missing endline responses were imputed with baseline values for the other 5 facilities in our sample. Treatment effects (ITT) are obtained from estimating Equation 1, including controls for the baseline mean of the respective outcome variable at the facility level and randomization strata fixed effects. Standard errors are clustered at the facility level and reported in parentheses. p-values based on cluster standard errors and randomization inference (RI) are shown. $*$, $**$, and $***$ refer to the statistical significance at 10, 5, and 1 percent level, respectively.

Appendix Table A17: Effect of Accreditation Assignment on Morbidity Screening and Treatment

	Control	ITT	p-value	
	Mean	(T−C)	Classical	RI
Panel (A) Screening Index				
SI	0.432	0.142*	0.064	0.067
		(0.075)		
OUD	0.347	0.119	0.120	0.123
		(0.075)		
HTN	0.386	0.006	0.959	0.967
		(0.114)		
Panel (B) Treatment Index				
SI	0.557	0.049	0.442	0.420
		(0.063)		
OUD	0.576	-0.042	0.497	0.551
		(0.061)		
HTN	0.269	0.043	0.525	0.413
		(0.067)		
Observations (N. of Clusters)	92 (46)			
Control group	44 (22)			
Treatment group	48 (24)			

Notes: This table shows the treatment effect of accreditation assignment on morbidity screening and treatment for three specific conditions: Suicidal Ideation, Opioid Use Disorder, and Hypertension from estimating Equation 1 at the patient level. For each condition, two medical records were audited per facility in the endline medical audit, each of which refers to a distinct patient. Facilities that did not provide medical records are imputed records where standards are coded as 0s for non-compliance. Panel (A)'s outcome variable is a screening index specific for each illness. Items checked in the screening are matched with each type of illness. SI screening index includes "history of or current suicidal ideation," "a history of suicidal behavior," and "the status of suicidal ideation." OUD screening index includes "current or prior withdrawal symptoms," "illicit drugs," "a history of substance use hospitalization," "a history of withdrawal seizures," "a history of detoxification and outpatient treatment," "the status of drug or alcohol use," and "the status of drug of alcohol withdrawal intoxication." HTN screening index includes "blood pressure," "pulse," and "respiratory rate." Panel (B)'s outcome variable is a treatment index specific for each illness. SI treatment index includes (1) "Do health records document that they received education (related to suicidal ideation)?"; (2) "For patients with suicidal ideation, is there evidence of a treatment plan?"; (3) "For patients with suicidal ideation, is there evidence that the inmates were monitored for suicide risk?"; (4) "Whether custody staff or health staff are monitoring patients with suicidal ideation?" OUD treatment index includes (1) "Do health records document that they received education on MAT?"; (2) "Are protocols used for the management of withdrawal?"; (3) "Is follow-up plan and treatment documented?" HTN treatment index includes (1) "Do health records document that they received education on the following? Salt intake/Weight loss/Medication use and frequency"; (2) "Do orders include low sodium diet/duration/specific instructions?"; (3) "Does blood pressure appear to be controlled based on most recent reading?"; (4) "Is follow-up plan and treatment documented?" All specifications include controls for the baseline measure and randomization strata fixed effects. Standard errors are clustered at the facility level and reported in parentheses. p-values based on cluster standard errors and randomization inference (RI) are shown. $*$, $**$, and $***$ refer to the statistical significance at 10, 5, and 1 percent level, respectively.

Appendix Table A18: Effect of Accreditation Assignment on Medical Inputs

	Control	ITT	p-value	
	Mean	(T−C)	Classical	RI
Total FTE (Health Staff)	7.204	-0.708	0.249	0.244
		(0.606)		
Vendor Change	0.227	-0.037	0.776	1.000
		(0.128)		
Refer Out (Health Staff)	0.821	-0.030	0.369	0.439
		(0.033)		
Equipment Index	0.876	0.010	0.634	0.669
		(0.021)		
Service Index	0.273	0.073**	0.039	0.047
		(0.034)		
Observations		46		
Control group		22		
Treatment group		24		

Notes: This table reports OLS estimates of the effect of accreditation assignment on medical inputs from estimating Equation 1. Total FTE is medical FTE as reported in the facility survey. Vendor Change is an indicator for whether the facility changed their vendor between baseline and endline as reported in the facility survey. Refer Out is the average of a rescaled Likert scale for whether health care staff can readily refer out. Equipment Index covers medical equipment including: handwashing facilities/alternate means of hand sanitation, examination table, scale, thermometers, a light capable of providing direct illumination, blood pressure monitoring equipment, stethoscope, ophthalmoscope, otoscope, transportation equipment, trash containers for biohazardous materials and sharps, sterilizer for non-disposable medical or dental equipment, equipment and supplies for pelvic examinations if female inmates are housed in the facility, oxygen, AEDs, pulse oximeter, PPE, multiple-test dipstick urinalysis, finger-stick blook glucose tests, peak flow meters, stool blood-testing material and pregnancy test knits. Medical equipment also includes dental equipment, hand washing facilities or alternate means of hand sanitation, dental examination chair, examination light for dental procedures, proper instruments for dental examination, trash containers for biohazardous materials and sharps from dental procedures, dentist's tools and PPE. Service Index covers medical services including: pharmacy, records, radiology, physical therapy unit, emergency, labs, hospice and medical orthoses or prostheses or other aids to impairment. Each component of equipment and services is given equal weight, and the indices are calculated as the simple average share of all included items. All specifications except for Vendor Change include controls for the baseline measure and randomization strata fixed effects. Robust standard errors are reported in parentheses. $*$, $**$, and $***$ refer to the statistical significance at 10, 5, and 1 percent level, respectively.

B Model of Accreditation

Consider a principal (the sheriff), denoted by S, and two agents indexed by $i \in \{C, M\}$: custody staff (C) and medical staff (M). Each agent chooses either high effort $(e_i = h)$ or low effort $(e_i = l)$, with $0 < l < h < 1$. Effort is costly, and costs are linear:

$$C(e_i) = e_i.$$

Health outcomes depend on the minimum of the two effort levels:

$$d = 1 - \lambda \min\{e_M, e_C\}, \tag{3}$$

where d is the mortality rate and $\lambda > 0$ is the marginal product of effort in reducing mortality. This formulation captures the idea that both agents must work together to reduce the "aggregate output" of inmate deaths. If either side exerts low effort, quality deteriorates.

The sheriff incurs liability when mortality exceeds a benchmark corresponding to high effort:

$$\hat{d} \equiv 1 - \lambda h.$$

Let $B > 0$ denote the effective penalty associated with excess mortality. In the absence of accreditation, the sheriff cannot perfectly attribute poor outcomes to individual agents, so penalties are imperfectly targeted.

B.1 Accreditation as Public Signal and Monitoring

Accreditation affects behavior through two channels:

1. **Public Signal:** Accreditation shifts each agent's beliefs about whether the other agent will exert high effort, facilitating coordination across agents.
2. **Monitoring:** Accreditation improves attribution of shirking, increasing agent-specific penalties for unilateral low effort.

Let $s \in \{0, 1\}$ denote accreditation status, observed by both agents, where $s = 1$ means the facility is accredited or engaged in the accreditation process.

Beliefs. Let $p(s) = \Pr(e_{-i} = h \mid s)$, with $p(1) > p(0)$, such that accreditation makes each agent more optimistic that the other agent will comply with standards and exert high effort.

Monitoring. Under accreditation, penalties become more informative. If both agents choose high effort, no penalty is imposed. If both choose low effort, each incurs the baseline penalty ηB, where $\eta \in (0, 1)$ captures imperfect attribution. If one agent chooses low effort while the other chooses high effort, accreditation allows the sheriff to better identify shirking. In that case, the low-effort agent incurs a higher penalty $\hat{\eta} B$, with $\hat{\eta} > \eta$, while the high-effort agent incurs the penalty ηB. We assume $(h - l) < \hat{\eta} B$, implying that the relative cost of effort is bounded by the penalty for shirking under certainty.

B.2 Expected Utility

Given accreditation status s, the expected payoff to agent i from choosing high effort is

$$\mathbb{E}[U_i(h, e_{-i}) \mid s] = p(s)(-h) + (1 - p(s))(-\eta B - h),$$

since if the other agent also chooses high effort, the agent pays only the effort cost, whereas if the other agent chooses low effort, the high-effort agent still pays the effort cost and faces the consequences of failed coordination.

The expected payoff to agent i from choosing low effort is

$$\mathbb{E}[U_i(l, e_{-i}) \mid s] = p(s)(-\hat{\eta}B - l) + (1 - p(s))(-\eta B - l).$$

Agent i chooses high effort if

$$\mathbb{E}[U_i(h, e_{-i}) \mid s] > \mathbb{E}[U_i(l, e_{-i}) \mid s].$$

Substituting the two expressions gives

$$p(s)(-h) + (1 - p(s))(-\eta B - h) > p(s)(-\hat{\eta}B - l) + (1 - p(s))(-\eta B - l).$$

Rearranging yields a threshold condition:

$$p(s) > \tilde{p} \equiv \frac{h - l}{\hat{\eta}B}. \tag{4}$$

Thus, an agent chooses high effort when the probability that the other agent also exerts high effort is sufficiently large.

B.3 Equilibrium with Accreditation

Without accreditation ($s = 0$). If

$$p(0) < \tilde{p},$$

then each agent prefers low effort, and the low-effort outcome (l, l) obtains. This captures a setting with weak coordination, poor monitoring, and limited ability of the principal to distinguish individual shirking.

With accreditation ($s = 1$). If

$$p(1) > \tilde{p},$$

then each agent prefers high effort, and the high-effort outcome (h, h) obtains. Accreditation thus allows the organization to move from a low-effort equilibrium to a high-effort equilibrium through two reinforcing mechanisms. First, by providing a public signal, it aligns expectations across custody and medical staff, making high effort more likely when agents believe others will also comply, facilitating coordination across agents. Second, by improving monitoring and attribution, it increases the cost of unilateral shirking, reducing the attractiveness of low effort even when coordination is imperfect. Together, these channels allow accreditation to both select and potentially enforce the high-effort equilibrium.

C Additional Study Design and Implementation Details

C.1 Recruitment

From July 2021 to April 2022, medium-sized jails across the U.S. were recruited to participate. The study was promoted at targeted meetings and conferences, including those of the Committee of State Sheriffs, the National Sheriffs' Association, the American Jail Association, the National Association of Counties, the National Association of Counties, and NCCHC itself.

In addition, we used the Safety Source Directory 2021 to make random cold calls to recruit additional facilities. The cold call script is available here. Interested facilities could also directly reach out to the Harvard research team. See Figure C8 below for our recruitment email.

If a facility expressed interest in the study, the Project Coordinator scheduled a half-hour informational meeting on Zoom with the facility and their medical leadership. During this session, the Project Coordinator provided a short presentation on the study background and components, as well as the potential benefits of NCCHC accreditation. Facilities also had the opportunity to ask questions about the study and accreditation process during these informational meetings.

Appendix Figure C8: Recruitment Email

Subject Line: Healthcare Accreditation Opportunity

Dear _____,

Congratulations! Your facility has been selected as an excellent candidate to participate in a study at Harvard University. The National Commission of Correctional Healthcare (NCCHC) has partnered with Harvard University to study the effects of accreditation on health care delivery in jails, and we are currently recruiting facilities **that are not currently accredited with the NCCHC to undergo healthcare accreditation.**

This is a terrific opportunity to have the health care provided in your facility accredited by the National Commission on Correctional Health Care (NCCHC) with a substantial discount. Participation involves completing surveys at two timepoints for which facilities will be compensated up to $1,000, an *optional* confidential assessment of the health care delivery system from Harvard University researchers, and healthcare accreditation by the NCCHC at an expedited timeline and substantially discounted rate.

Our study has been supported by the American Jail Association, approved by the Harvard Research Ethics Board, and granted a Certificate of Confidentiality from the National Institutes of Health. We attach: (1) a flyer with more study information; (2) our letter of support from the American Jail Association; (3) our institutional approval from the Harvard research ethics board; as well as (4) our Certificate of Confidentiality from the National Institute of Health, which guarantees complete protection of privacy of all participating facilities in our study, including during civil/criminal/administrative/legislative proceedings. Please feel free to forward these materials to any interested parties.

Would a ½-hour informational meeting to discuss this study work for you next week? Our schedules are extremely flexible, and we are currently recruiting sites, so contact us at your earliest convenience at accredstudy@hks.harvard.edu or by calling [NAME] at [PHONE].

Thank you for your commitment to correctional healthcare.

C.2 Enrollment

Following the informational meeting, the Project Coordinator sent the facility an email summarizing the session and a link to a screener survey used to confirm their eligibility. Criteria for eligibility included: (a) an average daily population (ADP) of 100 to 3,000, and (b) no prior accreditation with the NCCHC, although they could have received a different accreditation process in the past. Facilities that met this criteria received enrollment documents including: (1) a consent form (co-signed by medical and custody); and, (2) a Data Use Agreement (DUA). Facilities were also able to opt in to receive a $500 payment for completing each of the baseline and endline surveys. The study team also applied and was granted a Certificate of Confidentiality (COC) from the Department of Health and Human Services which guaranteed protection of information obtained from all participating facilities in our study, including during federal, state, and local civil/criminal/administrative/legislative proceedings. The American Jail Association also wrote a letter endorsing participation in the study.[31]

After signing the Consent and DUA, the facility was randomly assigned a date to attend an Orientation Meeting. During this meeting, the team answered remaining questions, obtained facility contact information, confirmed prior accreditation status and current health care vendor, and walked through the next steps for the study. The presentation is linked here.

For enrollment into sub-facility surveys, including the inmate interview, leader interview, and staff surveys, individuals within the enrolled facilities were recruited to participate in three different ways. For the inmate interviews, we asked each facility to randomly select three inmates to participate. For the interviews with facility leadership, we asked the facility to select one person on the custody side and one person on the health care side to participate. For the staff survey, facilities provided us with a list of health and custody email addresses to invite to participate in the survey.

[31] Appendix Table A2 demonstrates no significant differences across facilities that enrolled vs. those that did not.

C.3 Randomization and Stratification

Facilities were randomized into the treatment and control group after completing the baseline facility survey to minimize attrition. Facilities randomized to the treatment group (Accredit Now) began the accreditation process immediately. Because many correctional facilities cited monetary costs as a major barrier to seeking accreditation, the Accredit Now facilities accreditation costs were highly subsidized. Facilities randomized to the control group (Accredit Later) were invited to undergo the accreditation process after completing the endline survey (see Figure 2 for details of the study and accreditation process timeline).

Randomization occurred in three batches to avoid asking the facilities that were recruited earlier to wait before receiving their randomization status, thus minimizing the likelihood of attrition. Within each batch, we also stratified facilities into two groups based on average daily population (ADP), given the correlation between facility size and health outcomes in the literature (Adler and Chen 2023). Facilities whose ADP was larger than the median ADP of the batch were considered "large" while facilities whose ADP was smaller than the median were considered "small."

C.4 Index of Items

Appendix Table C1: Index of Items

Hyperlink to Item	Description of Item
Pre-Analysis Plan	Pre-Analysis Plan.
Baseline Facility Survey	The Qualtrics survey facilities completed at baseline.
Endline Facility Survey	The Qualtrics survey facilities completed at endline.
Facility Survey Codebook	This is the codebook that research assistants used to code-up facility survey responses.
Facility Survey Clarification Email	This email was sent to facilities after study team members reviewed the facility survey to request additional documents or ask questions related to their answers.
Staff Survey	The Qualtrics staff survey that facility staff members were invited to complete at baseline and endline.
Staff Survey Email	This email was sent to health and custody staff members at each facility to invite them to take the staff survey.
Facility Leadership Interview Questionnaire and Script	The questions asked to health and custody leaders as part of the leadership interviews at endline.
Inmate Interview Questionnaire and Script	The questions asked to inmates as part of the inmates interviews at endline.
Expert Survey Prolific Screener Survey	The screening survey posted to Prolific to find individuals who had previously been incarcerated and invite them to complete the full expert survey.
Expert Survey	The Qualtrics expert survey that individuals were invited to complete.
Medical Audit Workbook	The workbook that medical scribes completed during the virtual medical audit.
Certificate of Confidentiality	Certificate of Confidentiality obtained for the study from the National Institute of Health.
Screening Survey	This survey was completed by each interested facility to confirm study eligibility.
Recruitment Email	This email was sent to jails across the U.S. to recruit them for the study.

D Description of Data Elements

D.1 Facility Survey

After the orientation meeting, the facility received the link to the baseline facility survey. Each facility had 2 weeks to complete it and received email and phone reminders 5 days, 3 days, and 1 day before the survey deadlines.

At endline, the facility was asked to update their answers to the survey. The survey they received was prepopulated with their answers from baseline and they only had to update the answers to the questions that had changed, in addition to answering the three new questions that were added to the endline survey. The three new questions added were as follows: (1) A copy of inmate booking records from 2021 to present (including the name of individual, DOB, gender, race, and arrest charges); (2) Number of health care related lawsuits and/or consent decrees filed against the facility by year from 2017 to present; and (3) Estimate of the amount spent on health care each year from 2017 to present. The endline facility survey is available here.

After the facility submitted the online facility survey at both baseline and endline, the research team reviewed their responses and followed up with any clarifying questions regarding information provided or missing documents. Facilities had 14 days to provide missing information or clarify questions. If the information was not provided within the 14 days, the clarification request was closed and research team used the facility's original survey response. See the clarification request email linked here.

The material uploaded by the facility to the online survey was then downloaded to a secure server in a facility-specific folder. The research team then performed initial cleaning of the data and prepared documents for in-depth coding.

D.2 Staff Survey

After the facility survey was sent, the research team obtained staff email lists from each facility at baseline. This list included both custody and medical staff emails. The template email is available here and the staff survey questions are linked here. These questions are related to staff satisfaction, coordination, and perceptions of care. In some cases, the facility was unable to provide the list of emails, and we then worked directly with our facility contacts who distributed these surveys on our behalf. The deadline to complete the staff survey was two weeks.

The study team sent reminder emails 5 days, 3 days, and 1 day before the survey deadline to increase the number of responses. If at least 5 individuals completed the survey and a gift card lottery was permitted by the facility, one staff member was randomly selected to receive a $100 gift card.

At endline, we requested an updated staff email list and sent the survey to both the baseline and endline email list.

D.3 Medical Audit

Facilities were asked to complete a virtual medical audit at endline. To be able to measure the effect of accreditation and changes over time, medical charts were requested for both the endline period and period of time before baseline. Each facility received a checklist of material to prepare in advance of the audit available here which included two health records per condition per period for those with suicidal ideation, hypertension, and opioid use disorder, if available, as well as two inmates at random. We also collected the 12-month death log, sick log and any administrative reviews of deaths or autopsies.

Each facility was able to select a three-hour timeslot for the audit to occur. The medical audit was conducted by a medical expert who was blinded to the treatment status of a facility and led the completion of a catalog of standardized questions available here. If the facility used an electronic medical record system, the facility screen shared the records live on the call with the team. If the facility used paper records, the facility received a document camera from the Harvard study team to use to display the records virtually. In all cases, the identifying information of the patient was obscured and not recorded by the study team.

D.4 Leadership Interviews

At endline, we conducted qualitative interviews with one health leader (*e.g.*, nurse or health services administrator) and one custody leader (*e.g.*, warden or sheriff) at each participating facility to understand their perceptions of the health care delivery. These interviews were conducted on Zoom, took approximately 45 minutes each, and were conducted in private. The interview script, including questions, is available here.

D.5 Inmate Interviews

At endline, qualitative interviews were also conducted with three incarcerated individuals at each facility. The facility was asked to randomly select three incarcerated individuals to participate, of whom at least two should have received health care in the last month. The goal of these interviews was to understand their perceptions of the health care. These interviews were conducted on Zoom, took approximately 45 minutes each, and were conducted in private. The interview script, including questions, is available here.

D.6 Expert Prediction Survey

Definition of Expertise: We recruited respondents with different *types* of expertise to a prediction survey. Given that our focus was on health services in jails, the three domains of expertise were considered: (1) Jail administrators and healthcare providers, (2) Previously incarcerated individuals, and (3) Policy analysts and academics.

Recruitment to Prediction Survey: From October 2022 to March 2024, we collected predictions on our study from 145 experts across different domains of expertise. Jail Administrators and jail health care providers were recruited to complete the survey at national corrections conferences. In addition, we used Safety Source (2021) to cold call various directors of facilities that were not associated with the study. Previously incarcerated individuals were recruited to participate through a screener survey posted on Prolific. To be eligible for the survey, the screener asked them to select all statements that apply to them from a list of 13 options with one of the options being that they have personal experience with the criminal justice system. If they select this option, they then receive another questions asking to select which experience align to their experience. All respondents that selected that they have previously spent time in jail, were then invited to take the expert survey. Finally, policy analysts and academics were recruited at a variety of conferences.

Expert Prediction Survey Aims: The purpose of the expert surveys was to: (1) assess whether accreditation was predicted to affect outcomes and the directionality of that effect; (2) assess whether the predicted effect differed by type of expertise; (3) ascertain baseline features of facilities might be important for heterogeneity; and (4) compare expert predictions to the study findings.

Deployment: The survey was deployed through Qualtrics (available here).

Incentives: All participants received a $20 gift card for participation. To incentivize survey accuracy, three experts and two previously incarcerated individuals with survey responses most accurately aligned with the results of the study are to receive an additional $100 gift card.

Survey Flow and Demographics: Following the collection of basic individual background characteristics and a brief description of the study, participants were asked what baseline factors might modify the effects of accreditation. Summary statistics on the experts can be found in Appendix Table D1. Next, respondents were asked *whether* accreditation would affect patient care and treatment, inmate deaths, litigation and health care related spending. We incentivized responses to such answers noting that those who were closest to actual results would receive a bonus payment. In addition, we asked participants to rank the importance of the quality standards and define what is included in those categories. For many of these questions, we also asked the degree of confidence the respondents had in their answers. We use these data to understand whether experts from different domains ranked outcomes differently, suggesting they value or weight things differently, and to take a disciplined approach to assessing for heterogeneity.

Appendix Table D1: Expert Survey Demographics

	Number of Respondents	Percent of Respondents
Age		
18-24	3	2.07%
25-34	23	15.86%
35-44	41	28.28%
45-54	51	35.17%
55-64	20	13.79%
65-74	5	3.45%
75+	2	1.38%
Gender		
Male	66	45.52%
Female	78	53.79%
Transgender	1	0.69%
Race/Ethnicity		
Asian	12	8.28%
Black or African	19	13.10%
Hispanic	4	2.76%
Native American	1	0.69%
Mixed race	6	4.14%
White/Caucasian	101	69.66%
Prefer not to say	2	1.38%
Schooling		
Some college but no degree	22	15.17%
Associate's degree (2 years)	22	15.17%
Bachelor's degree (4 years)	29	20.00%
Post-graduate degree	64	44.14%
High school degree or less	8	5.52%

Notes: This table shows the demographics of those who participated in the expert survey. The number of recruited experts is 145. The first column lists the key demographic information, the second is the number of participants that fall into that group, and the third column calculates the share of respondents within that group (adding to 100 for each).

Appendix Figure D9: Facility Photos

Private Space



Sharps Removal



Notes: Figure shows two examples of facility photo documentation as required by the facility survey. Left panel shows a private patient examination space. Right panel shows sharps removal container.

E Coding and Construction of Indices

Coding for this project required two major tasks: (1) Coding up raw material (*e.g.,* responses and forms from facilities); and (2) Data cleaning and analysis. We also developed processes and procedures to make sure coding was performed in a systematic, unbiased way across various members of the team.

E.1 Coding Committees and Codebook

One of the more laborious tasks was translating the NCCHC questionnaires into analyzable data elements. This required determining which questions were most objective and thus should be included in our Harvard study. We also had to develop a mapping between the questions and binary fields for whether that particular component of the standard was met or unmet (Yes/No, 1/0).

Coding Committees were assembled and worked for several months examining each individual NCCHC question and developing the most objective version of that question for Harvard surveyors as well as specific guidance as to whether the particular criteria was met. The coding committee for the Facility Survey consisted of Harvard undergraduates working with the study authors and Coordinator, alongside predoctoral fellows. The Coding Committee for the medical audits included medical students, the Project Coordinator and was overseen by one of the PIs with expertise in medicine. The Facility Survey Committee created a codebook for the coding by research assistants, who typically were Harvard college or Harvard law students. The codebook described how each answer should be coded as to ensure consistency across RAs. Importantly, the RAs were trained to code *whether the service/process* was done as indicative of meeting the standard, not the *outcome of that test/process itself*. For example, *testing* for sexually transmitted infections (STI) was considered meeting the standard even if the patient was ultimately diagnosed with a STI. Each row of the codebook defined the variable name, format (*e.g.,* binary, text), reasons for skipping, question, instructions for coding, valid entries, and the variable description. A link to the codebook is available here.[32] See Appendix Section E.4 for further details.

E.2 Variable Transformation

The last step was the more familiar data cleaning and analysis by predoctoral fellows from the Health Inequality Lab and the Malcolm Wiener Center. This required resolving disagreements between two independent manual coding RAs with a third tie-breaker if needed. Transformation of variables was also required. Variables were recoded so that higher values were preferred. In addition, we performed the following transformations:

1.) For binary questions – we maintained the dichotomous structure and applied the penalty rule described below.

2.) For continuous outcomes – we dichotomized them for purposes of aggregation using the NCCHC standard rule whenever it was specified or the median across the baseline sample.

   a.) For example, for the question how many days after admission was a mental health assessment performed, NCCHC provides a clear standard (within 14 days) and any facility not meeting that standard was given a 0.

   b.) As an example of the second approach, for the question, "how many FTE does a mental health specialist contribute?" - NCCHC does not have a clear standard. Thus, responses higher than the median at baseline were given the value of 1 while those lower than or equal to the median were given a value of 0.

3.) For grouped outcomes, we use the number of outcomes in the group as the denominator and compute the share of all items in the group that were performed as the numerator. For example, one of the grouped questions was whether health records documented patient education on salt intake, weight management or medication.

[32] At endline, only the answers that changed from the baseline survey were coded.

4.) For some text variables, we encoded to a binary variable. For example, to the question "What is the name of the Responsible Health Authority (RHA)," we coded as 0 if there was no name given as that indicated there was no RHA.

In addition, non-response was often indicative that standards were not being met and we developed specific procedures for coding missings (see Appendix Section E.2.2).

E.2.1 Training of the Manual Coding Research Assistants

In order to be hired as a manual coding Research Assistant for the project, students had to undergo several steps. First, the applicants had to sign a form stating they would maintain strict confidentiality if they happened upon identifying information about a facility in the process of coding up their responses. Second, they did CITI training – which was required by the IRB. Third, they had to complete a training class with the Project Coordinator. The training guided new coders through the process and showed them where all the documents were located and what information they contained. Lastly, they completed a coding test which was supervised by a senior coder consisting of a spreadsheet with each of the 202 variables across the top. The coding was checked for accuracy and feedback was provided. A particular question was manipulated to assess whether the coder would detect the ambiguous response and reach out to study leadership. If the coder passed these checks, they were used in the actual manual coding exercise.

E.2.2 Penalization for Missing Data

If a facility was unable or unwilling to provide information, we assume the worst case scenario.

1. **Refused participation in survey**: If a facility refused to conduct a crucial quantitative part of the Harvard study – such as the medical audit – all responses to that specific survey instrument are coded as unmet. As noted in the main text, the baseline and endline facility surveys were completed by all 46 facilities.
2. **Refused to answer a specific question**: If a facility participated in the survey but refused to answer a specific question, the response is coded as not meeting the specific criteria.
3. **Gateway questions**: Whether some questions are asked depends on how a previous question is answered. For example, during the medical audit, if a facility answers Yes to "Is the inmate tested for STIs," they are then asked "Of the sample tested for STIs, which diseases were tested: Chlamydia, Syphilis, HIV.”" However, facilities that answer no to "Is the inmate tested for STIs," are not asked the following questions, and were penalized as not having met any of the subsequent criteria.

E.3 Coding Outcomes

An important coding decision was how to reduce the dimensionality of the outcomes in a principled manner. This was done by categorizing all the various questions in the facility survey and medical audit into one of the seven quality standards. Given the importance of mortality as an outcome, it was coded separately using both administrative death logs and media reports. See Appendix Section E.8 for further details.

E.4 Coding Facility Survey

The facility survey consisted of 62 questions that were asked directly to a jail administrator and 43 requested documents. Based on the work of the *Coding Committee* the responses were divided two types of variables: (1) coder to determine (CTD) in which RAs determined whether the each answer provided by the facility met the corresponding quality metric and manually coded this determination, and (2) facility to answer (FTA) in which the facility administrator answered directly through the online portal. Each of the FTA questions were then mapped to the seven quality standards and to 40 substandards. Of the 328 CTD questions, the codebook specifically provided instructions on how to ascertain whether the criteria was being met. The coding process for the Facility Survey was as follows:

1. For each facility, two RAs independently coded the baseline facility survey by inspecting the uploaded documentation and following the codebook.

2. To ensure independence and to maintain security, each RA received a personal folder only they could access on the server – and with only the one facility's information they were coding at the time. This was done to prevent RAs from checking their work against others and to preserve security by keeping data access strictly limited.
3. The two manual coding RAs were blinded as to whether the facility was in the treatment or control groups.
4. After completion, a third team member resolved any flagged discrepancies by re-reviewing the entire section the discrepancy was found (not just the question item). The third team member was also blinded to the first two reviewers' coding.

E.5 Coding Staff Survey

These data are individual-level data with most of the answers being multiple choice. For Likert outcomes, we rescaled so the outcomes were between 0 and 1 but kept continuous.

E.6 Coding Medical Audits

For the medical audits, we received raw workbooks from the medical scribes, which we then cleaned and adjusted certain variables based on NCCHC standards. For example, if an incarcerated individual had not been in the facility for at least 14 days, it would not be applicable to have an initial health assessment. Additionally, when specific medical items (*e.g.*, STI testing, receiving screening) are not recorded, we penalize them by coding them as zeros, indicating that the facility did not meet the standard. These data are inmate-patient level, with two records requested per condition per facility and period, including general patient records. The records are collapsed to the facility level for the construction of quality metrics in Section VI and Appendix A. However, for analyses focused on mechanisms in Section VII, we kept data at the patient-inmate level to assess screening and intake.

E.7 Coding Media-Reported Deaths

Consistent with applicable federal law on deaths in custody reporting, we exclude deaths occurring prior to booking (during arrest or transport to jail). Coding was done in accordance with a structured protocol involving two independent coders who were blinded to treatment status. Coders were provided a protocol specifying search terms, source types, and the set of variables to be recorded. For each identified death, we record the date and location of death to the extent allowed by the data sources. One coder, remaining blinded to treatment status, conducted a formal adjudication process in which each source was re-verified and discrepancies in death details were resolved by prioritizing sheriff records when available or sources with the most complete and current information.

To maximize coverage across the variation in media-related reporting, we systematically searched three categories of sources: (i) traditional media outlets and public releases from Sheriff's offices and medical examiners, (ii) reports from oversight bodies, including District Attorney investigations and Department of Corrections publications, and (iii) independently-maintained public databases, including Reuters and Incarceration Transparency.

We determine the location of death based on the reported place of official pronouncement. Deaths are classified as in-jail when reported as occurring within the facility with no conflicting report that the individual was pronounced deceased elsewhere. Cases in which individuals are transported and subsequently pronounced deceased at a hospital are classified as hospital deaths.

E.8 Construction of Indices

As per our pre-analysis plan, we created indices corresponding to each of the quality standards. The indices were created using the questions within each of the data sources (facility survey, both CTD and FTA as well as medical audit) that corresponded to that standard. The NCCHC groups their standards into seven main categories, listed below and described in further detail in Table 1 with typical examples:

- Governance and Administration
- Health Promotion, Safety, and Disease Prevention
- Personnel and Training
- Ancillary health care Services
- Patient Care and Treatment
- Special Needs and Services
- Medical-Legal Issues

To construct the meta indices of quality we aggregate questions that pertain to a given standard. We aggregate them using the pre-specified approaches described below:

Equal Weights: Each question was equal weight by computing the simple average across all variables in a given index.

Principal Component Analysis: We also use principal component analysis (PCA) to aggregate individual components (Bai 2003). We include all 224 outcomes and use the first component to create a meta-index, rescaling to ensure the principal component is positive and lies in the unit interval (Lise and Postel-Vinay 2020).

Expert Weights: We created a weighted overall average with the weights corresponding to the importance of each standard as ranked by experts (See Appendix Section D.6 for further details).[33]

We follow an analogous construction of indices using the production function approach described in Section IV.2.

[33]This method is also inspired by work from Pew Charitable Trusts (2017). We computed weights from reciprocal ranks as described in Section IV.3.