Pre-Analysis Plan

Evaluating the Effects of NCCHC Accreditation on Health Care and Health Outcomes in U.S. Jails

Original Version: May 26, 2021

Last Updated: June 5, 2024

Deviations in Red

Table of Contents

1. Introduction

1.1 Background

In the landmark decision *Estelle v. Gamble* (1976), the Supreme Court ruled that reasonably adequate health care to prisoners was a right guaranteed by the Eighth Amendment. In *Estelle*, the Court noted that "[a]n inmate must rely on prison authorities to treat his medical needs; if the authorities fail to do so, those needs will not be met. In the worst cases, such a failure may actually produce physical torture or a lingering death[.] … In less serious cases, denial of medical care may result in pain and suffering which no one suggests would serve any penological purpose." The standard of care in *Estelle*, incorporated into the states through the due process clause of the Fourteenth Amendment, has since been used to challenge inadequate medical, dental, and mental health care in prisons, jails, and juvenile detention facilities.

Since then, the provision of health care (medical, dental, and mental health) has become a major expense for correctional facilities. A 2013 report by the Pew Charitable Trust highlighted this underappreciated state and local budgetary strain: "Despite increasing interest among policymakers and taxpayers in improving outcomes and controlling costs in health care and corrections, the intersection of these two areas…has garnered comparatively little attention" (Bell et al., 2013). The Pew report noted that annual health care spending in correctional facilities had a median growth rate of 50% between 2001 and 2008. A 2014 report by the Pew Charitable Trust noted that correctional health care spending rose in 41 states from fiscal 2007 to 2011, with median growth of 13%, after adjusting for inflation. Especially relevant is that many insurance providers, including Medicaid and Medicare, do not cover the care of incarcerated individuals. This "inmate exclusion policy" has contributed to a significantly under-resourced correctional health care system that is isolated from mainstream medicine and exempted from accreditation and external quality oversight mandates (Fiscella et al., 2017).

Local governments are particularly hard hit by the health care burden in jails (Vera,

2015). In 2017, 10.6 million people were booked into jail (Zeng, 2019). As the entry point for individuals into the criminal justice system, jails provide an environment for an abrupt transition of care.[2] While correctional populations have dropped slightly in recent years, the U.S. still has the highest incarceration rate in the world. Separately, decades into the well-intended but poorly executed de-institutionalization movement and the closure of state mental health facilities, the U.S. criminal justice system has absorbed an increasing number of individuals with mental illness and is now the largest provider of mental health services (Major County Sheriffs of America, 2019).

Faced with high inmate populations and significant medical needs among this group, jails benefit from guidance on how to meet the constitutional mandate established in *Estelle*. In *Estelle,* the Supreme Court held that "deliberate indifference to serious medical needs of prisoners" constitutes an Eighth Amendment violation. This principle applies regardless of how the medical care is provided, e.g., by governmental employees or by contracted medical providers. However, these constitutional standards do not provide the specific guidance on day-to-day operations that jails often seek, with health care improvements often piecemeal and in response to litigation. Because of this haphazard approach, accreditation bodies such as the National Commission on Correctional Health Care (NCCHC) emerged to provide a framework for health care quality in jails.

Yet only a small fraction of jails seek any type of voluntary accreditation, despite high levels of interest and many jails informally asking NCCHC for advice or trying to "follow the standards." Jails may not voluntarily seek out accreditation due to inertia, cost, or lack of perceived benefit. Today, inmate death and morbidity remain high despite constitutional mandates. Common causes of death in jails include suicide, heart disease, and alcohol and drug intoxication (Zeng et al., 2016). Few facilities have protocols for dealing with inmates at risk of suicide and few have staff certified to deal with addiction treatment and provide care for withdrawal symptoms (Fiscella et al., 2017). When these inmates return to their communities without treatment, many die at high rates in the weeks following release (Fiscella et al., 2017) and may spread disease to the community.

Burden on Low-Income Communities: The legal requirement that correctional facilities provide adequate health care to inmates, coupled with myriad challenges in this population, has many counties searching for solutions. Inmates tend to be sicker than age-matched controls, with higher rates of diabetes, hypertension, asthma, and HIV (Wilper et al., 2009, p.669). Many are also from low-income communities. As a result, ensuring that jail inmates return to those fragile neighborhoods in good health yields a public health benefit. Because health issues, particularly mental illness and substance abuse disorders, contribute to recidivism (Pew Charitable Trust, 2014), improving health care can also improve public safety and help low-income communities meet their crime-reduction objectives.

Contribution: To the best of our knowledge, we are proposing the first randomized trial of accreditation in the jail health care context. We are aware of one study that randomized accreditation of hospitals in South Africa (sample size = 20 hospitals) and found that accreditation improved quality metrics (Quality Assurance Project, 2003). A recent meta-analysis by Mumford et al. (2013) assessed around a dozen facilities (non-correctional care) to determine if the benefits of accreditation were greater than the costs but found insufficient evidence.[3]

Theoretically, the effects of accreditation on the performance of service providers are ambiguous. If the accreditation process a) focuses on items that do not improve health or b) serves only as a rubber stamp, accreditation could have an inconclusive or negative effect on health service provision by diverting attention and resources to less useful activities or by fostering complacency with poor care. If, on the other hand, accreditation serves to disseminate best practices and improve accountability and organization of health care services, it may have a positive and measurable effect on health outcomes. Whether improved inputs (i.e., service provision) translate into better outputs (i.e., improved inmate health, linkage to care, and staff satisfaction, or reduced health care costs, recidivism, and litigation) are policy-relevant empirical questions. Indeed, proposals to reform health care in correctional facilities often state that "states and counties should require accreditation of health care services" (Fiscella et al., 2017).

1.2 Intervention

We will partner with NCCHC, the oldest correctional health care accrediting body in the U.S. NCCHC grew out of a program of the American Medical Association in the early 1970s with a goal of understanding and improving health care for inmates (Gibson and Phillips, 2016). Early AMA surveys revealed that health care in jails was highly disorganized, with inmates often serving as front-line providers. A set of standards and a pilot accreditation program followed these surveys with funding from the Law Enforcement Assistance Administration (Modlin, 1979). The pilot program expanded slowly to 23 states. In the 1980s, funding from the Robert Wood Johnson Foundation enabled NCCHC to be established as a stand-alone nonprofit organization. Today, NCCHC develops separate sets of standards for jails, prisons, juvenile detention facilities, mental health services, and opioid treatment programs. In addition, NCCHC supports a peer-reviewed journal (the *Journal of Correctional Health Care*), provides certification for correctional health professionals, and hosts educational conferences and webinars.

Accreditation of jails is based on the NCCHC *Standards for Health Services in Jails*. These standards are widely accepted by the health, legal, and corrections professions. The seven sections of the standards cover the following areas: governance and administration; health promotion, safety, and disease prevention; personnel and training; ancillary health care services; patient care and treatment; special needs and services; and medical-legal issues.

According to NCCHC, accreditation confers many potential benefits. It promotes and documents an efficient and well-managed health care delivery system. It adds to the status of the facility, increases staff morale, aids recruiting efforts, helps to obtain community support, and provides additional justification for budgetary requests. Accreditation can also help protect a facility's assets by minimizing the occurrence of adverse events, reducing liability premiums and protecting facilities from lawsuits related to health care. Importantly, accreditation benefits the health of the public,

staff, and inmates by assuring that those incarcerated and released receive adequate and appropriate health care according to nationally accepted standards.

Approximately 453 jails and 163 prisons were accredited by NCCHC between 2008 and 2014. In a recent observational analysis of which types of facilities were more likely to be initially compliant with NCCHC standards, the authors find that jails were more likely to be noncompliant than prisons and that more admissions led to greater noncompliance (Gibson and Philips, 2016).

1.3 Study Aims

In this study, we plan to rigorously evaluate the impact of health care accreditation of jails by conducting a staggered encouragement RCT, randomizing 22 facilities to treatment ("accredit now") and 22 to control ("accredit later").

2. Sampling and Experimental Protocols

2.1 Recruitment and Sampling

We are recruiting 40+ medium-sized jails (i.e., average daily population approximately of 100+). Because of NCCHC's deep connections with correctional facilities, there are numerous targeted meetings and conferences where the study could be advertised, including those of the Committee of State Sheriffs (where NCCHC is a paid sponsor), the National Sheriffs' Association, the American Jail Association, and NCCHC itself. We have found that many jails are interested in accreditation but have not been accredited due to cost, lack of information and institutional inertia. Our informational sessions with facilities and financial support for study participation helps reduce some of these barriers. We plan to recruit at least 40 jails across the U.S. This would give us a broader geographic scope than focusing on a single state and minimize the possibility of spillover effects.

The NCCHC accreditation process begins with the applicant completing a voluntary self-survey questionnaire (SSQ) to evaluate their current practices and help them prepare for the accreditation survey. The in-person visit to the facility is scheduled when NCCHC

deems the applicant ready (usually 3 to 6 months after completing the SSQ). A survey team of trained correctional health care experts then collects data on-site to measure compliance with each of the standards. The team tours the facility, reviews health records and other documents, and gathers data through observation and interviews with staff and patients. The survey also records facility characteristics such as average daily population (ADP), capacity, and health care staffing. The lead surveyor then submits a report to NCCHC's Accreditation and Standards Committee, which meets 6 times a year to evaluate reports and determine compliance. The facility name remains anonymous to the committee.

We developed a baseline facility survey which includes many of the same elements of the survey NCCHC conducts. The survey includes information such as jail capacity, ADP, inmate demographics, size and composition of health staff, and health care policies and procedures. We will use this baseline survey to assess randomization balance. The endline survey will include a facility component, a review of medical charts, inmate interviews, and exit interviews with key leadership staff. In addition, Harvard will also conduct online staff surveys intended to capture the views of custody and health care staff at both baseline and endline – these surveys are not typically done by NCCHC and are a novel component we have added. The endline medical survey will be conducted by Harvard-employed physicians or physicians-in-training surveyors so that they can be completely similar in treatment and control locations and independent.[1] Indeed, after randomization, Harvard will have minimal contact with the facilities until the endline, so as to not interfere with the accreditation process.

2.2 Randomization

After orientation, facilities will sign consent (co-signed by custody and health) and data use agreements per Harvard policy. Following baseline survey completion, we will randomly assign facilities to treatment (accredit now) and control (accredit later) groups.

[1] In addition, we will deploy a survey to experts in health care, policy, and corrections, as well as previously incarcerated individuals. The purpose will be to forecast the results after a brief explanation of the trial as well as to provide input on potentially important heterogeneity (See 5.4).

When deciding on a randomization method, we considered pair-wise matching, stratification, and re-randomization. Pair-wise matching would have been challenging with our small sample size if we experience any attrition; if one facility drops out, we would have to drop the entire pair since imbalance is suboptimal for variance estimation (Athey and Imbens, 2017). This would reduce our sample size by two facilities for each attrition. Re-randomization makes it more challenging to conduct valid inference as the exact nature of the re-randomization strategy would have to be taken into account considered. In addition, in experiments with randomization at the unit-level, stratification is superior to complete randomization, and pairing is superior to stratification in terms of precision of estimation (Athey and Imbens, 2017). Bruhn and McKenzie (2009) show that stratification and pair-wise matching perform best in smaller samples like ours and dominate re-randomization methods. They explain that variables that are likely to be correlated with outcomes of interest are the most promising candidates for stratification.

In each batch, we use a stratified randomized design based on facility size as measured through the average daily population (ADP) of participating facilities. We stratify on ADP due to preliminary analysis from baseline survey responses suggesting that ADP is significantly correlated with over 30% of the other outcome variables.

Randomization occurred in batches to avoid asking our early survey completers to wait before finding out their randomization status, hence reducing the possibility of attrition and congestion for the accrediting agency. Therefore, we performed stratified randomization based on size and batch.

2.3 Hypotheses Tested

We hypothesize that accreditation will improve health care and health outcomes. Below, we divide our hypotheses into two groups: outcomes for medical audience and outcomes for an economic audience.

Hypothesis Group A: Outcomes for the Medical Audience

To test these hypotheses, we plan on measuring medical care and health outcomes. Examples include which medications are provided to a patient (if any), whether medical diets are provided per prescriber order, and what acute medical problems are identified and treated.

Specifically, our primary outcome measures include:

- Screening for STIs
- Suicide assessment screening and monitoring and treatment plans
- Receipt and documentation of MAT medication and treatment history
- Management and treatment of hypertension
- Healthy lifestyle promotion
- Administration of medical diets
- Timeliness of care

Our secondary outcome measures include:

- The presence of required fields in the receiving screening
- Initial health assessment
- Mental health assessment
- Documentation of oral care

We will measure these outcomes using two survey instruments: our baseline and endline facility survey as well as the endline medical audit.

Hypothesis Group B: Outcomes for the Economic Audience

We hypothesize that accreditation will have wider-reaching effects than just those directly related to medical care and health. Therefore, in addition to the aforementioned outcomes, we will assess if accreditation affects facilities' adoption of policies and procedures, which would create more efficient and effective administrative outcomes as well as better medical outcomes.

Accreditation may also affect staff and inmate morale, as measured by retention of high-quality staff, positive attitudes among staff regarding health, job satisfaction, and positive inmate attitudes.

To test this set of hypotheses, we test the effect of accreditation on indicators of various procedural health outcomes and procedural administrative outcomes which could reasonably be expected to change within a 16-month period. An example of procedural administrative outcomes is the existence of a quality improvement committee (CQI) committee. An example of a procedural health outcome is the steps taken (if any) when an acutely suicidal inmate is identified.

The data for assessing these hypotheses would come from the baseline and endline facility survey, which asks about health policies, protocols, and assessments. We will also use the staff surveys to understand staff attitudes, coordination, and job satisfaction. The staff surveys query both custody and health staff and occur at baseline and endline. Lastly, we plan to conduct inmate interviews and interviews with leadership to ascertain their views.

2.4 Primary Outcomes for the Economic Audience

To consolidate outcomes, we use indices. We consider two main approaches for the determination of index categories:

Option A: Use a Modified Categorization of NCCHC Standards

The NCCHC groups their standards into seven main categories (e.g., "Governance and administration" and "Health promotion, safety, and disease prevention".) We include all outcomes that relate to a given standard in that category into an index. Each index is described in Table 1 with examples. Under this approach, our outcome indices would be closely aligned with the aspects of jail health care directly targeted by the intervention.

Table 1. NCCHC Indices

Index	Definition	Examples
NCCHC A	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.	Health record content, staff FTE, continuous quality improvement, grievance process
NCCHC B	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
NCCHC C	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.	Staff credentials, custody staff training, health staff orientation
NCCHC D	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
NCCHC E	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, oral care, discharge planning, health care requests
NCCHC F	Standards in this section address patients with special health care needs and establish compliance requirements specific to each health care issue.	Mental health services, MAT, sexual abuse protocol, pregnancy care
NCCHC G	The standards in this section ensure that health services comply with legal requirements.	Use of restraints, segregation, informed consent/refusal

Option B: Using a Categorization of Outcomes Based on a Production Function Approach [2]

[2] 6/5/24: We focus on the NCCHC inspired quality standards (Option A) in the main analysis. In addition, we are able to assess the effect of accreditation on mortality as a primary outcome measure. Initially, we expected to use the provision of death logs as a quality measure, but realized that death log data also contain information on deaths (number and cause), allowing us to assess whether mortality is affected directly.

Our outcomes measure various aspects of jail health care. Most of these can be thought of as either inputs (e.g., formalized policies, personnel qualifications, and available equipment) or outputs (e.g., patients with a certain disease or if a screening is conducted in time). We can further disaggregate outputs into health outputs, such as disease burden, and procedural outputs, such as patient screening. Each group of outputs is described in Table 2 with examples. This approach allows us to understand which aspects of health care are impacted by NCCHC accreditation.

Table 2. Production Function Indices

Index	Definition	Examples
Inputs	Variables in this index include outcomes related to the availability of staff and equipment at the facility.	Staff FTE, staff qualifications, equipment
Formal Procedures	Variables in this index include processes that are explicitly laid out in the policies and procedures or detailed in other paperwork gathered from the facility.	Grievance policy, emergency response plan, screening forms, health record forms
Procedural Outcomes (Administrative)	Variables in this index include outcomes related to management and administrative processes that do not involve the management of or interactions with inmate patients.	Administrative meetings, continuous quality improvement, annual reviews of policies, custody staff trainings
Procedural Outcomes (Health)	Variables in this index include outcomes related to patient-facing procedures.	Screening schedule, medication administration, point of care tests, 24-hour emergency services, sick call schedule
Medical Outcomes	Variables in this index include outcomes related to the receipt of medical care.	Vaccinations, emergency room admissions, CD4+ T-Cell count, diabetes control, medications continued

Weights for Index Calculation

To construct these indices, we must aggregate the variables that fall into them. There are several ways to make this aggregation. For all options, we must first transform the outcomes so that higher values are more desirable and then standardize them using the control group mean and standard deviation. Next, the simplest approach would be to give all variables the same weight and calculate the simple average over all variables that are to be summarized in the specific index. An alternative approach would be to use principal component analysis with the first principal component as an outcome for every index. One caveat is that the weights with which the outcomes are aggregated to components are not necessarily positive. Anderson (2008) proposes an algorithm to aggregate outcomes to an index using the inverse covariance matrix for weighting. Our preliminary power analysis found that indices created with that algorithm perform poorly when a large number of outcomes are to be aggregated.

The same approach could also be used to create an index summarizing all potential outcomes we collect. An additional option we consider is calculating a weighted average where weights are determined by asking experts (e.g., health care, corrections, policy experts, and previously incarcerated individuals, detailed in section 5.4) the importance of different outcome categories. Specifically, we ask the experts to rank the importance of the index categories described under Primary Outcomes Options A and B. We can then calculate one overall index by weighting the indices according to the average expert ranking. If, for example, experts rank Medical Outputs higher than Administrative Processes, variables falling in this index category will get a higher weight in the overall index. This calculation of an overall index will be done separately for the NCCHC and Production Function categories.

3. Harvard Study Timeline

May 20-May 21	Fundraising, exploratory research, and hiring
Jan 21 - Jul 21	Study Preparation (Recruitment, Surveys, IRB, Registration, PAP, Study Manual)
Aug 21 – Mar 23	Harvard administers baseline survey for all 44 facilities
Oct 21 – Apr 23	Harvard randomly assigns facilities to treatment and control group and refers accredit-now facilities to NCCHC
Jun 22 – Nov 23	NCCHC begins accrediting 22 accredit-now facilities on 12-month timeline
May 23 - Apr 24	Harvard administers endline survey for all 46 facilities
Dec 23 - Aug 24	NCCHC begins accrediting 22 accredit-later facilities on 12-month timeline

4. Data Collection

4.1 Baseline

Instead of surveyors touring facilities in-person or online, which could potentially induce the Hawthorne effect, we ask facilities to complete an online baseline facility survey through Qualtrics before the initiation of the accreditation process. The survey asks for various details about the delivery of health care services in the facility, and asks facilities to upload documents, policies, and protocols related to health care. Harvard coders will independently review these documents to assess compliance with each of the standards of NCCHC accreditation.

We also administer a staff survey through Qualtrics and encourage responses from as many staff members in each facility as possible. This survey targets both health care and custody staff, and measures staff morale and job satisfaction as well as the presence of COVID safety measures and other compliance indicators.

4.2 Endline

6/5/24: Originally, we had planned for the endline survey to occur 14 months after baseline. However, this was not possible due to delayed scheduling of the on-site NCCHC visits. Instead, facilities are broken into three main batches for endline: (1) Spring-Summer, (2) Summer-Fall, and (3) Late Enroll. The endline process began in May 2023 for the Spring-Summer batch, in July 2023 for the Summer-Fall batch, and in March 2024 for the Late Enroll batch. These windows are created by taking the minimum and maximum of the latest and earliest dates of endline for Accredit Now facilities within each strata.

Surveyors will review health staff information, inmate health records, policies and procedures, and other pertinent documents. Surveyors will also survey inmates, correctional officers, and health staff to assess how they view the coordination and delivery of health services in their facilities. We will also administer the same staff survey at endline to assess any changes in staff satisfaction and compliance with health care standards from staff members' perspectives. Lastly, we will interview inmates as well as health care and correctional leadership in the facilities to assess their perceptions of health care delivery and institutional norms in their jails. We will include these qualitative observations in our analysis.

4.3 Data Quality Checks

To ensure the reliability of outcome measures, we aim to only obtain data that is objective, verifiable, and coherent with appropriate documentation. Survey responses will be double-coded by checking against documentation provided by facilities. Response categories will be clearly defined in advance, and responses to survey items will be coded based on a pre-established coding scheme. Data analysts will train survey coders and work with them to address any errors and poorly understood questions. Prior to beginning to code documents, all survey coders must attend a training session, pass a practice test, and complete CITI and data safety trainings. Data analysts will regularly monitor coder output to account for outliers, inconsistencies, and missing values, and resolve any ambiguities when coders identify areas for amendment or follow-up. Two

coders will code each facility survey response to ensure consistency and avoid subjective interpretations. A third tiebreaker will then systematically resolve any ambiguous coding decisions. The survey coders are blinded to the treatment status.

4.4 Retention

We aim to minimize attrition by building personal relationships when screening jails and highlighting the benefits of health care accreditation (i.e., a well-managed, efficient correctional health care delivery system; increased staff satisfaction, morale, and retention; minimized liability; improved health outcomes). Attrition will be assessed in real-time and we will check that attrition is not differential across study arms. We will work vigorously to avoid attrition and continue to increase financial incentives for participation in the study if necessary. We would seek additional funding opportunities to provide facilities with an additional $1,000-$5,000 for survey completion and continue lowering the out-of-pocket cost of accreditation for facilities. Additionally, we will randomize in multiple "batches" of 10-20 facilities (as described in section 2.2) to reduce drop-outs due to long wait times. If our sample size exceeds 40, latecomers will be placed on a waitlist; they will be separately randomized into the treatment or control group and placed "on standby" as a back-up in case of attrition within their assigned group. We will continue to seek to recruit additional facilities. We will use these backup facilities in place of the attrited facilities in main analyses if attrition is not differential across groups.

In addition to the mechanisms highlighted above, we will also identify facilities that are at risk of attriting and establish a plan for retention. For example, a turnover in correctional leadership such as a new sheriff election could jeopardize the facility's ongoing participation. We will track these risk factor for all facilities (both accredit later and accredit now) on an ongoing basis. For any facility that we believe is high risk, we will conduct a special meeting with leadership that will focus on (1) sharing insights from their baseline survey, (2) providing additional transparency on the NCCHC accreditation process, and (3) emphasizing the importance of completing the endline survey. This meeting will begin with sharing facility-specific findings from the baseline and staff survey that illustrate their initial compliance with the NCCHC standards and staff

satisfaction. Second, we will review the NCCHC process. To provide additional transparency into the accreditation process, we will share the NCCHC survey training material. In addition, we will inform them of flexibility in the timing of accreditation and offer to increase our subsidization of their NCCHC accreditation fee. To help answer any additional questions and provide a perspective from someone familiar with the process, we will put the leadership in contact with a former Jail Commander with expertise in accreditation. Finally, we will conclude the presentation by sharing the difference between the NCCHC accreditation process and the Harvard study, emphasizing the importance of completing the endline survey even if they elect to withdraw from the accreditation process.

4.5 Attrition from Sample

If a facility indicates they want to drop from the accreditation process, we will schedule a call at the facility's earliest convenience to emphasize the importance of completing the Harvard endline survey and offer additional financial compensation for the staff's time. If a facility attrits from the process, we will test whether assignment status predicts attrition.[3] We will then proceed as follows:

(1) **Accredit Now facility completes endline**: Ideally, if a facility drops out of the intervention, they will still be willing to complete the endline survey. In this scenario, we will use instrumental variables following Angrist et al. (1996) to capture the causal effect of the non-random decision of the facility exiting the study. If the decision does not affect our estimators, the facility will be included in our study. Otherwise, the facility will be excluded.

(2) **Accredit Now facility does not complete endline**: In the worst case, if a facility drops out of the study and is unwilling to complete the endline survey, we will use the baseline values to fill in the missing endline values.

(3) **Accredit Later facility does not complete endline:** If a facility in the control group drops from the study and is therefore unwilling to complete the endline survey, we will use the baseline values to fill in the missing endline values.

[3] 6/5/24: We clarified this section after finding errors in the original version.

6/5/24: We anticipated, given the length of the study, disruptions of sheriff elections, and changes in health care vendors, that facilities would potentially attrit during the project, so we continued to recruit and built up a waitlist of facilities to randomize in the event of such attrition. Over the course of the study period, two "Accredit Now" facilities did exit immediately after receiving their treatment assignment. In total, six facilities were recruited later in the study period due to attrition concerns. This resulted in a total number of 46 jails.

4.6 Non-Reporting Bias

One concern is that accredited facilities might look "worse" because they carefully monitor outcomes that others do not. For example, accredited facilities may be more likely to diagnose illnesses at booking because of diligence, making these facilities' health outcomes appear "worse" than non-reporting facilities. To address this potential non-reporting bias, we will penalize facilities that fail to meet a given standard. For example, a facility might not meet the criteria for a screening in 7 hours, but another facility might not report anything on screening. Those that report they are over the time limit might be penalized for the length of the time they are behind on average, whereas those that don't keep report at all will be punished harder.[4] If we cannot verify if a policy is implemented as required (e.g., because the documentation needed is not obtained from the facility), we will assume the process is not properly implemented.

4.7 Demand Bias

There are several ways that our process reduces the possibility of demand bias. Demand bias occurs when study participants anticipate or are aware of what the researcher is investigating, and hence may alter their behavior or survey responses to include what they think the researcher or surveyor wants to see. One key characteristic of our study is that the intervention and evaluation are separate; Harvard researchers are independently

[4] 6/5/24: Given the way we constructed the indices by dichotomizing variables at the median, penalization is equivalent to assigning the minimum.

assessing the impact of the NCCHC accreditation process and are not involved in the accreditation process itself.

Second, we use administrative data which is objective and verifiable, such as checking the contents of medical records. The chances of facilities creating documents they do not already have are low given staff time constraints. We also test for both policy presence and implementation; the baseline survey assesses whether facilities have policies and protocols to meet NCCHC standards, and the medical survey assesses whether these policies and protocols are implemented in practice. This addresses demand bias by allowing the Harvard research team to independently verify compliance through policy implementation in the rare case that facilities do alter their written policies to meet what they think the researchers want to see.

In addition, multiple coders are coding each facility's survey response, so that we can verify that the two entries are consistent and there is no subjective interpretation of facility documents or other survey data. We will also incorporate the qualitative interviews with inmates into the analysis, and inmates have no incentive to speak more favorably about the delivery of health care services or misrepresent the quality of these services. Lastly, we test for outcomes that are not included in NCCHC standards, such as staff morale (through the staff survey). There is a lower chance of research subjects anticipating these outcomes.

5. Empirical Analysis

5.1. Economic Specification

We will use the following empirical specification to estimate the effect of the treatment:

$$\boldsymbol{Y_i = \alpha + \beta Treat_i + \gamma' X_i + \varepsilon_i}$$

Where $\boldsymbol{Y_i}$ is the outcome of interest for jail i at *endline*, $\boldsymbol{Treat_i}$ is a dummy which equals 1 if a jail is randomized into the accredit-early group, $\boldsymbol{\beta}$ is the estimate of the average

treatment effect, and $\boldsymbol{X_i}$ is a vector of control variables including baseline survey variables and strata fixed effects as recommended by Bruhn and McKenzie (2009). Outcomes of interest are listed in Appendix B.

6/5/24: We did not anticipate that many facilities in our study would not be able to obtain accreditation. Therefore, we also use 2sls to estimate the above with assignment to treatment as an instrument for receipt of accreditation.

5.2. Balance Checks

We will assess differences across groups of facilities on key characteristics to ensure balance between facilities in the treatment and control groups. These characteristics may include ADP, prior accreditation, vendor characteristics, whether jails have a private health care vendor, region or state, and baseline outcomes.

5.3. Heterogeneous Effects

Secondary analyses will include assessing heterogeneity by jail ADP, whether the jails have a private health care vendor, and the number of health staff-to-inmate ADP ratio. Factors of interest for our secondary analyses will also be determined based on our findings from a survey of experts (see section 5.4).

For jail ADP, it may be possible that results differ by facility size due to economies of scale and a high fixed cost for establishing effective health care policies and procedures (part of the motivation for stratifying by this variable). In addition, private health care vendors often have specific policies and protocols followed in each of their contracted facilities, which could lead to faster take-up of NCCHC policies and protocols during the accreditation process. Lastly, the staff-to-inmate ratio might affect a facility's ability to comply with standards regarding the turnaround time for health care processes.

5.4. Expert Survey

An expert survey will be deployed to previously incarcerated individuals and experts in corrections, policy, and health care. The purpose of this survey is to 1) ascertain what different aspects of heterogeneity might be important, 2) obtain weights (as discussed in section 2.4), and 3) compare expert predictions to the studying findings.

The survey will be deployed through Qualtrics and contains three main sections.

The structure of the survey is as follows. First, participants will be asked to select what factors they think will affect whether NCCHC accreditation has an impact on health care procedures, health outcomes, staff outcomes, and litigation. The goal of this question is to understand what factors participants think are important to whether NCCHC accreditation has an impact. We will then perform heterogeneity analysis on the variables participants believe will affect the response to treatment.

In the second section, we ask participants to rank the importance of both the NCCHC standards and production function outcomes. We provide them with the 7 NCCHC standards and 5 production function outcomes, and define what is included in those categories. The categories of outcomes that the participants deem more important will be given higher weights that will be used to collapse all indices into a single index.

Finally, we ask participants to predict whether and to what extent accreditation will impact health care procedures, health outcomes, staff outcomes, and litigation. Once our study is completed, we will compare the expert predictions to the actual findings. Correct answers will be incentivized.

5.5. Spillover Effects

We will investigate potential spillover effects of jail health care accreditation on the overall preparedness, prevention practices, and management of COVID-19 by examining screening practices for new and current inmates and staff; hygiene practices, isolation

practices; modifications to intake and release policies; provision of onsite treatment; and referral for offsite treatment. Please see Table 1 for further details.

The distance between counties participating in our study is generally large, so we do not anticipate cross-county spillover effects. In cases wherein multiple jails from the same county are participating in our study, we ensure that these jails can be treated as different units. First, we hold meetings with them to discuss their organizational setup and ensure that there is no staff overlap or information-sharing across jails within the same county. Second, we ask these jails to sign commitment letters indicating that they will treat each jail as an independent entity and not update policies and protocols across jails, should some be randomized into the treatment group and others in the control group from the same county.

5.6. Standard Error Adjustments

Robust standard errors will be used for facility-level outcomes. For staff-level outcomes, standard errors will be clustered at the facility level.

6. Power

From March 2020 through the beginning of June 2020, Alsan and Yang administered a weekly COVID-19 survey to correctional facilities across the United States in partnership with NCCHC. The survey asked respondents to identify how their facility was coping with the pandemic, including questions on PPE, testing, screening, and incidence rates. We use the means and standard deviations across six binary outcomes from these survey results as a basis for our power calculations, as our experiment will focus on a similar group of respondents and outcomes. Table 1 provides an overview of the Minimum Detectable Effect (MDE) across eight sample sizes. See Appendix Exhibit 1 for additional information.

Table 1.

	Total No. Facilities							
	30	**40**	**50**	**60**	**70**	**80**	**90**	**100**
Outcome								
Inmate COVID Testing	0.530	0.455	0.404	0.368	0.340	0.317	0.299	0.283
Inmate COVID Screening	0.244	0.211	0.189	0.172	0.159	0.149	0.140	0.133
Health Staff COVID Screening	0.403	0.345	0.307	0.280	0.258	0.241	0.227	0.215
Meds. on Discharge	0.424	0.364	0.323	0.294	0.272	0.254	0.239	0.226
Any Lawsuit	0.403	0.345	0.307	0.280	0.258	0.241	0.227	0.215
Any Death	0.350	0.300	0.267	0.243	0.224	0.209	0.197	0.187

7. Funding and Human Subjects Review

This study is funded by the Abdul Lateef Jamil Poverty Action Lab (J-PAL) at the Massachusetts Institute of Technology (MIT), the William T Grant Foundation, the Preparedness and Treatment Equity Coalition, and Arnold Ventures. Study approval has been granted by the Harvard Institutional Review Board (IRB).

8. References

Athey, Susan, and Guido W. Imbens. "The econometrics of randomized experiments." In *Handbook of economic field experiments*, vol. 1, pp. 73-140. North-Holland, 2017.

Angrist, J.D., Imbens, G.W. and Rubin, D.B., 1996. Identification of causal effects using instrumental variables. *Journal of the American statistical Association*, 91(434), pp.444-455.

Bell, Ellyon, Kavita Choudhry, Kil Huh, Matt McKillop, Matt Mulkey, and Maria Schiff. 2013. “Managing Prison Health Care Spending.” The Pew Charitable Trusts. https://www.pewtrusts.org/-/media/legacy/uploadedfiles/pcs_assets/2014/pctcorrectionshealthcarebrief050814pdf.pdf

Bruhn, Miriam, and David McKenzie. 2009. "In Pursuit of Balance: Randomization in Practice in Development Field Experiments." *American Economic Journal: Applied Economics*, 1 (4): 200-232.

Douds, Anne S., and Eileen M. Ahlin. 2016. "Do NCCHC dental standards have any teeth?." *Journal of Correctional Health Care* 22(3): 180-188.

Dumont, Dora M., Annie Gjelsvik, Nicole Redmond, and Josiah D. Rich. 2013. "Jails as public health partners: Incarceration and disparities among medically underserved men." *International journal of men's health* 12(3): 213.

Estelle v. Gamble, 429 U.S. 97 (1976)

Fiscella, Kevin, Leo Beletsky, and Sarah E. Wakeman. 2017. "The inmate exception and reform of correctional health care." *American journal of public health*. 107(3): 384.

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.

Lee, D.S., 2009. Training, wages, and sample selection: Estimating sharp bounds on treatment effects. *The Review of Economic Studies*, 76(3), pp.1071-1102.

Ludwig J, Mullainathan S, Spiess J. Machine-Learning Tests for Effects on Multiple Outcomes. Published online 2017. doi:10.48550/ARXIV.1707.01473

Major County Sheriffs of America. 2019. *Sheriffs Addressing the Mental Health Crisis in the Community and in the Jails*. Washington, DC: Office of Community Oriented Policing Services.

Modlin, Herbert C. 1979. "Medical care in correctional institutions: The AMA project." *Journal of the American Academy of Psychiatry and the Law*, 7: 118-124.

Mumford, Virginia, Forde, Kevin, Greenfield, David, Hinghcliff Reece, and Jeffrey Braithwaite. 2013 Health services accreditation: what is the evidence that the

benefits justify the costs? *International Journal for Quality in Health Care* 25: 5: 606–620

"NCCHC Historical Time Line." National Commission on Correctional Health Care. Retrieved March 10, 2020. https://www.ncchc.org/time-line.

Quality Assurance Project; 2003 *The impact of accreditation on the quality of hospital care: KwaZulu-Natal Province Republic of South Africa*. Bethesda, MD: (http://www.hadassah-med.com/media/2021788/TheImpactofAccreditationonQofHospitalCare.pdf).

Salmon, J. Warren, John Heavens, Carl Lombard, and Paula Tavrow with foreword by James R. Heiby, and commentaries by Stuart Whittaker, Marie Muller, and Marilyn Keegan, and Anne L. Rooney. 2003. "The impact of accreditation on the quality of hospital care: KwaZulu-Natal Province, Republic of South Africa." *Operations Research Results* 2(17). Bethesda, MD: Published for the U.S. Agency for International Development (USAID) by the Quality Assurance Project, University Research Co., LLC.

The Pew Charitable Trusts. 2016. "How and When Medicaid Covers People Under Correctional Supervision." Retrieved March 6, 2020, from www.pewtrusts.org/en/research-and-analysis/issue-briefs/2016/08/how-and-when-medicaid-covers-people-under-correctional-supervision.

U.S. Department of Justice. Office of Justice Programs. 2013. "Census of Jails." Washington, DC: Bureau of Justice Statistics. Retrieved January 14, 2020, from, https://doi.org/10.3886/ICPSR26602.v1

Vera Institute 2005 The Price of Jails Measuring the Taxpayer Cost of Local Incarceration https://www.vera.org/publications/the-price-of-jails-measuring-the-taxpayer-cost-of-local-incarceration

Wilper, Andrew P., Steffie Woolhandler, J. Wesley Boyd, Karen E. Lasser, Danny McCormick, David H. Bor, and David U. Himmelstein. 2009. "The health and health care of US prisoners: results of a nationwide survey." *American journal of public health* 99(4): 669.

Zeng, Zhen, Margaret Noonan, E. Ann Carson, Ingrid Binswanger, Patrick Blatchford, Hope Smiley-McDonald, Chris Ellis et al. (2016). "Assessing inmate cause of death: Deaths in custody reporting program and National Death Index." *World Health* 46: 50.

9. Appendix A

Exhibit 1. Power Calculations

The research team originally conducted power calculations for this experiment in April, 2020 and concluded that a sample size of 40 total facilities would be sufficient to detect a significant treatment effect. This was based on the assumption that the Stata *power twomeans* command used in these power calculations, implements the following equation:

$$d = \frac{(m_A - m_B)}{(\sigma)} \qquad (1)$$

where d is the *standardized* effect size, m_A, m_B are the population means expressed in raw (original) units, and σ is the standard deviation of either population (since they are assumed equal) (Cohen 1988).

However, only the Stata *power onemean* command implements this formula. Stata defines the effect size for a two-sample means test as the difference between the experimental-group mean and the control-group mean[5]:

$$d = m_A - m_B \qquad (2)$$

Consequently, the correct way to interpret an MDE of 0.53 in Table 1 is that there would need to be a 53-percentage point difference between the control and treatment group means to detect a significant effect.

10. Appendix B: Data Collection and Outcomes

Key outcomes we plan to measure in baseline, and endline surveys include:

[5] See Stata 16 power manual for more information.

A. Meeting NCCHC's compliance and protocol standards (indexed for each major category). These include, for example,
   a. **<u>Governance and administration:</u>**
      i. **Health Records:** A confidential health record is created and maintained using a standardized format.
      ii. **Procedure in the event of an Inmate Death:** The responsible health authority conducts a thorough review of all deaths in custody in an effort to improve care and prevent future deaths.
      iii. **Responsible Health Authority:** Facility has qualified staffing and sufficient number of staff regularly available.
      iv. **Administrative Meetings and Reports:** Documentation of quarterly administrative meetings.
      v. **Continuous Quality Improvement (CQI):** Documentation of CQI meetings and reports.
      vi. **Privacy:** Patients receive health care in private settings.
      vii. **Grievance process:** Existence of a formal grievance process.
   b. **<u>Health promotion, safety and disease prevention:</u>**
      i. **COVID* and infectious disease prevention:** Appropriate measures are taken to limit the spread of COVID and infectious diseases.
      ii. **Healthy lifestyle promotion:** Health care policies procedures, and practices emphasize health promotion, wellness, and recovery.
      iii. **Suicide Prevention and intervention:** Suicides are prevented when possible by implementing prevention efforts and intervention.
      iv. **Contraception:** Contraception is made available as clinically indicated.
      v. **Clinical Preventative Services:** Testing and treatment available for STIs.
   c. **Personnel and Training:**
      i. **Health Training for Correctional Officers:** Correctional officers receive training on how to assist in emergencies.
      ii. **Orientation:** All new health staff receive in depth training/orientation.
   d. **<u>Ancillary Health Services:</u>**
      i. **Medical diets:** Medical *diets* are provided that enhance patients' health.

ii. **Medication:** Patients receive necessary medications and are able to continue taking medications prescribed prior to incarceration.
iii. **Emergency Services:** Planning for emergency health care ensures that all staff are prepared during emergencies.
iv. **Clinic space and supplies:** Facility has necessary clinic space and supplies to provide adequate health care.
v. **On-site diagnostics:** Facility has equipment to administer various clinical tests on-site.

e. **Patient Care and treatment:**
i. **Receiving Screening:** Screening is performed on all inmates upon arrival at the intake facility to ensure that emergent and urgent health needs are met.
ii. **Initial Health assessment:** Inmates receive initial *health assessments*.
iii. **Mental Health screening and evaluation:** Inmates receive initial *health assessments*.
iv. **Non-emergency Health Care Requests and Services:** Inmates' nonemergent health care needs are met.
v. **Nursing Assessment Protocols and Procedures:** 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
vi. **Continuity, Coordination, and Quality of Care During Incarceration:** Patient medical, dental, and mental health care I coordinated and monitored from admission to discharge.
vii. **Information on Health Services:** Inmates receive information on how to request medical services.

f. **Special Needs and Services:**
i. **Patients with chronic issues:** Patients with chronic disease, other significant health conditions, and disabilities receive ongoing multidisciplinary care aligned with evidence-based standards.
ii. **Infirmary level care:** Infirmary-level care, when provided, is appropriate to meet the health care needs of patients.

iii. **Mental Health Services**

iv. **Medically Supervised Withdrawal and Treatment:** Inmates who are intoxicated or undergoing withdrawal are appropriately managed and treated.

v. **Pregnant inmate:** Pregnant inmates are given comprehensive counseling and care in accordance with national standards and their expressed desires regarding their pregnancy.

vi. **Sexual Assault Response:** In the case of a report of sexual assault, the facility enacts a response plan including providing emergency contraception and separating victim and abuser.

g. **Medical Legal Issues:**

i. **Use of restraints:** Compliance indicators related to the use of restraints.

ii. **Segregation:** Appropriate monitoring of inmates in segregation.

iii. **Informed consent/refusal**: Patients must indicate they provide consent prior to receiving medical care and also have the ability to refuse care they do not want.