Rights Violations in Texas Jails:

A Text Analysis of Noncompliant Jails

Rebecca A. Larsen

Ph.D. Student

Public Policy and Political Economy

University of Texas at Dallas

PPPE 6355

December 10, 2023

Introduction

In August 2023, 22 family members and former inmates sued Harris County Jail in Texas for creating a “place of torment and punishment.” The lawsuit details medical neglect leading to deaths, assaults by corrections officers, refusal to intervene during violence, and significant mental health neglect (Grench, 2023). Since 2005, there have been 273 deaths in the county jail’s custody (Attorney General of Texas, 2023) and at least 52 individual lawsuits (Stuckey, 2023). The Texas Commission on Jail Standards, the state’s regulatory commission, repeatedly found the jail to be out of compliance with jail standards and as of the writing of this report the county remains noncompliant (TCJS, 2023). On July 23, 2023, 180 miles north, in the significantly smaller East Texas county of Rusk, Johnny Bradley was pronounced dead after being found unresponsive in his jail cell. The Texas Commission on Jail Standards soon after found the jail to be out of compliance due to leaving Bradley in his cell for hours without any face-to-face checks which are required by law (KLTV, 2023). Johnny is one of 113 people who have died in a Texas local jail this year.

The increase of people being held in local U.S. jails is one of the leading contributors to increases in incarceration (Sawyer & Wagner, 2023; Wiseman, 2019). Most people held in U.S. jails are pretrial, meaning they have not been convicted of a crime and are awaiting their trial in jail (Leslie & Pope, 2017; Stevenson, 2018). Since 2010, the Bureau of Justice Statistics has reported an increase in people dying in local jails. Deaths in jails are routinely two and three times above the national average of the general population (BJS, 2021). Beyond deaths, reform groups and news reports have called attention to harmful conditions, mental health care neglect, and regular abuse (Texas Jail Project, 2023; Wang, 2021). Conditions and rights violations in state and federal prisons have been extensively studied, but research on violations within jails is limited. Few sources of data exist for county jail conditions given that each state has different standards, and individual counties and sheriff's offices are tasked with running the jails.

This report asks an exploratory research question: How and where are rights being violated in Texas county jails? Two hypotheses are tested, H1: Lower population jails will have more standard's violations, and H2: Most violations will be due to mental health care standards. Descriptive statistics, correlations, and a dictionary based method are implemented to test the hypotheses. The first hypothesis is rejected and preliminary support is found for the second hypothesis. After validating the models, next steps are explored to improve upon the limitations within this report and expand to future work that seeks to both create more data to analyze human rights in jails and better understand how mental health violations are occurring in Texas jails.

Background and Puzzle

The goal of this paper is to use text-as-data to better understand jail violations in county jails. The Texas Commission of Jail Standards has four inspectors that are tasked with examining each of the 246 Texas local jails annually. Though not meant to be a full oversight body, they are the only agency that has access to all jails in Texas and does inspect them for specific state regulations. The commission reviews jails for several safety, mental health, and medical requirements. Further, when incidents are reported, such as deaths, the inspectors will conduct special inspections to determine compliance. These reports provide insight into how jails are violating rights.

This exploratory report asks the research question: How and where are rights being violated in Texas county jails? Despite the high profile and persistent violations occurring in Harris County, I expect most of the violations to occur in smaller counties due to significant growth of rural jails (Vera Institute, 2023) and the lower levels of staff and resources along with higher levels of deaths reported by rights groups (Wang, 2021). Therefore, Hypothesis 1 is: Lower population jails will have more standard's violations. Suicide and severe mental health issues occur at higher rates in jails than the general population (Copp & Bales, 2018; BJS, 2021). Further, the Sandra Bland Act of 2017 was passed after Sandra Bland died of suicide in a highly scrutinized arrest and death in the rural Waller County jail. That act created additional state mental health screening, reporting, and suicide watch requirements that TCJS now inspects for (Silver, 2017). Therefore, Hypothesis 2 is: Most violations will be due to mental health care standards.

Literature Review

Jail incarceration has tripled over the past 30 years, with rural counties predominantly driving this growth (Copp & Bales, 2018). Despite this, there is limited research on jails with most of the literature focusing on prisons.[1] The prison literature has documented human rights abuses (e.g. Bierie, 2011; Ross, 2011; Reiter et al, 2020) and consistently recommended that effective oversight needs to come from independent, third-party bodies (Deitch, 2020; Faithi, 2010; Worsley & Memmer, 2020). What does exist concerning local jails predominantly focuses on frequency of mental health conditions and prevalence of suicide and death.

Sixty four percent of people in jail have a mental health problem according to previous diagnosis or observation by mental health professional during incarceration (Copp & Bales,

[1] In the United States, jails are run by counties and local governments to house people pretrial, after an arrest subsequent to release if charges are dropped, some short county-level sentences (less than a year), and at times as holds for other agencies, such as U.S. Marshalls or Immigration and Customs Enforcement. The vast majority of people are held pretrial (between 75% to 80% on average). Prisons are run by state or federal governments and house people post-conviction. They tend to have more systematized procedures, documentation, and screening and therefore there tends to be more data on them compared to local jails.

2018). More people with severe mental health issues are in jails and prisons than in hospitals (Torrey, et al., 2010). Suicide rates are seven times higher in jails than in prisons (BJS, 2021), and six times higher in the smallest jails than largest jails (Meagher & Chammah, 2015). People with mental health diagnoses in a Midwestern jail experienced more threats and were at greater risk of assault (Ellison et al., 2022). Kajeepta et al., (2021) observed that avoidable premature deaths are associated with county jail incarceration, and high staff turnover is associated with higher death rates (Adler & Chen, 2023). Given these limited but compelling research findings concerning mental health and death, there is a need for data development and assessment to better understand rights violations in jails.

Research Design

This study uses a text-as-data approach from noncompliant reports created by the inspectors of the Texas Commission on Jail Standards (TCJS). The data is batch downloaded, preprocessed, and analyzed using descriptive statistics and a dictionary-based method. Both approaches are validated through hand coding a sample.

Data

Noncompliant jail reports are made public by TCJS. However, once a jail has taken the required actions to remedy the violation(s), the reports are no longer available. Through open records requests, the nonprofit group Texas Jail Project has obtained and shared all noncompliant reports from 2013 to November 2023. The documents include the standard being violated, along with why the jail failed to adhere to the standard. There are a total of 426 noncompliant reports that vary significantly in length. Many of the reports include one or two standards violations, while others contain dozens in a single report. Noncompliant reports were batch downloaded using the Chrome plugin Batch Link Downloader. The reports were in PDF format in varying degrees of quality with most of them in a photo format.[2] Using the Action Wizard function in Adobe Acrobat, the reports were OCRed to recognize the text and then converted into plain text files (UTF-8). The data was then read into R for pre-processing and analysis.

The data was converted into a corpus that includes the document ID and all the text included in each report. I then separate out the data by county and year units. This is done by relying on information included in the file names. I then hand checked each county-year unit name and fixed any issues. From there, the document ID, County, Year, and the text of the noncompliant report was incorporated into the original dataset from the corpus.

Hypothesis 1 Methodology: Number of Reports and Length of Reports

H1: Lower population jails will have more standard's violations. For the first hypothesis, the independent variable is county size. This was collected from the U.S. Census Bureau for the year of the violation report. Two dependent variables are considered: number of noncompliant reports and number of violations. For the first dependent variable, I run a Pearson's R correlation to see if the expected negative association exists, and to determine if further testing is justified. Given the lack of relationship, I do not conduct further analysis on number of reports and county size. For the second dependent variable, I use length of report as a proxy for the number of violations. Because of this, I calculate this variable without any further pre-processing of the data. Any commonly used text preprocessing could alter the outcome as the approach relies on the assumption that longer documents contain more violations. Using the quanteda package in R, I create a document feature matrix. The document feature matrix transforms the text into a matrix where the rows are the original text, and the columns are the tokens (in this case words). There are a total of 3,283,880 words. Next, I calculate the total words

[2] When a PDF is in a photo format, the individual words within the text are not recognized.

of each report by summing the rows within the document feature matrix. I then read this back into the dataset to be a variable of total words in each report. Again, county size is incorporated into the data and a Pearson's R correlation is calculated to determine if there is a strong, negative relationship as expected. To control for possible outliers, several different models of correlations are run for both reports and length. Once again, no strong relationship exists and further testing is not justified.

Finally, I validate the approach of number of words as a proxy for number of violations. I do this by organizing the data by total words and then taking three stratified random samples of 20 reports from the lowest third, the highest third, and the middle third. I inspect each document sampled by hand, counting the number of violations in the reports. Then I calculate the average number of violations for each of the 20 reports by grouping to assess whether the expectation of fewest to most violations is realized for each data section. I conduct a series of t-tests to see if there is a significant difference between the document groups.

Hypothesis 2 Methodology: Dictionary Based Method

H2: Most violations will be due to mental health care standards. A dictionary-based approach is used to test hypothesis 2. Dictionary based approaches can be useful in understanding context of documents, propensity of concepts, and variation over time and regions (Grimmer, et al., 2022). I begin by pre-processing the text using the tm package in R. I remove punctuation, convert all letters to lower case, remove white space, and remove stop words. I then create a dictionary of custom stop words that were not captured by the tm package's stop words. These include the always used words: "texas", "standards", "jail", "commission", "facility", "inspect", "shall", "cell," "jail", and "inmate." Once these words are removed, I stem the text, again using the tm package. This process transforms words into their root form which can assist in not missing similar words (i.e. suicide, suicidal) or over-counting the same concepts (Denny & Spirling, 2018). From here, the clean text was extracted from the corpus and added to the original dataset, which now has Document ID, County, Year, Population, Text, Clean Text, and Number of Words.

After reviewing a sample of the reports to build an understanding of the different types of violations, I searched for the most used language areas. I did this to determine if a dictionary-based approach for mental health violations was a logical next step in the analysis. Table 1 shows the top language areas used. Suicide and mental health are the top two with 116 and 197 reports utilizing this language, respectively. This preliminary assessment justified building a dictionary of mental health terms to better understand if and how mental health violations are driving noncompliance.

Table 1. Top Violation Language Used by Report

Violation Language	Number of Reports Containing Subject
Suicide	197
Mental Health	116
Training	107
Sanitation	99
Medication	47
Death	31
Detox	56
Restraint	55

To build the mental health dictionary, I read a sample of 20 reports that included the terms “suicide” and “mental health.” This revealed commonly used terminology and violations. Using the stringr package in R, I searched for the stemmed version of each word and created the following dictionary: “ment”, “suicid”, “disab”, and “ccq.” [3] After building the dictionary, I subset the document frame matrix to the dictionary terms and calculated the total mental health dictionary words per report. Then I added this new variable of total mental health words into the dataset. Next, I analyzed the dictionary results, examining the words by total words and by county.

I validate using three different approaches. First, I validate the dictionary created to see if it is properly capturing mental health violations through the counting of mental health words. I take a random sample of 20% of the documents (86 reports), and hand count the number of mental health related topics. Grimmer et al. (2022) recommends hand-coding to create a sample to validate the dictionary. I create a column of my own count and run a Pearson’s R correlation to determine how strong the association is, while qualitatively accounting for the reasons behind any differences. The next approach is a binary one by report, where I use the dictionary to determine whether there was any mental health violations or none (1,0). Using the same sample of reports, I hand-code whether there were mental health violations or not. From these results, I created a Confusion Matrix to give an additional measure of model validation by calculating both precision and recall (Shung, 2018). Finally, I create a Structural Topic Model to understand the main associated themes across the documents to determine if the top topics are capturing mental health violations or if other topics are most frequently recognized across the documents. Structural Topic Models (STM) are an unsupervised classification system that can uncover latent topics across a corpus of texts (Lebryk, 2021). Utilizing the stm package, I run an STM across

[3] CCQ is an acronym for Continuity of Care Query that is required within 72 hours of booking. These queries check to see if an inmate has prior contact with the public mental health system, such as a state hospital.

the documents to identify the top 10 latent topics and interpret the findings qualitatively as a robustness check for the dictionary based approach.

Results and Discussion

There were a total of 426 noncompliant jail reports from 2013 to 2023. As Figure 1 shows, more jails are noncompliant over time, with most noncompliant reports occurring in 2023 (as of November of 2023). Falls and Harris Counties have the most occurrences of noncompliance, 11 and 10 respectively (See Figure 2). The other most frequent violators were Bowie County with nine, and Bosque, Red River, and Coryell Counties each with seven. Figure 3 maps the violations by frequency (1 to 11, with 0 being grayed out). This reveals a pattern of more violations occurring in the eastern part of Texas compared to other regions.

Hypothesis 1

The descriptive visualizations suggest a pattern of violations that is beyond the size of the county, due to the concentration in the eastern region (Figure 3) and the mix of population sizes among top noncompliant counties (Figure 2). However, further consideration is needed before rejecting Hypothesis 1.

Figure 1. Noncompliant Reports Over Time

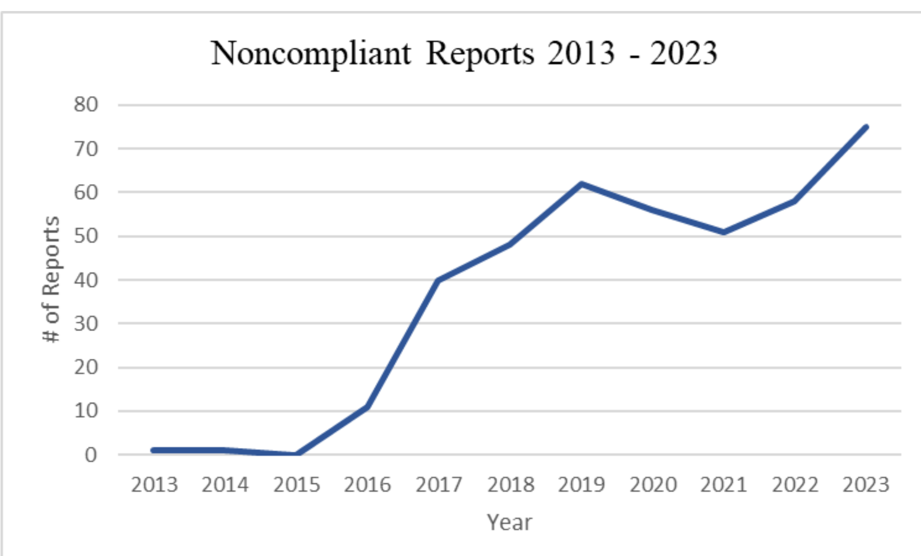


Figure 2. Top Noncompliant Counties

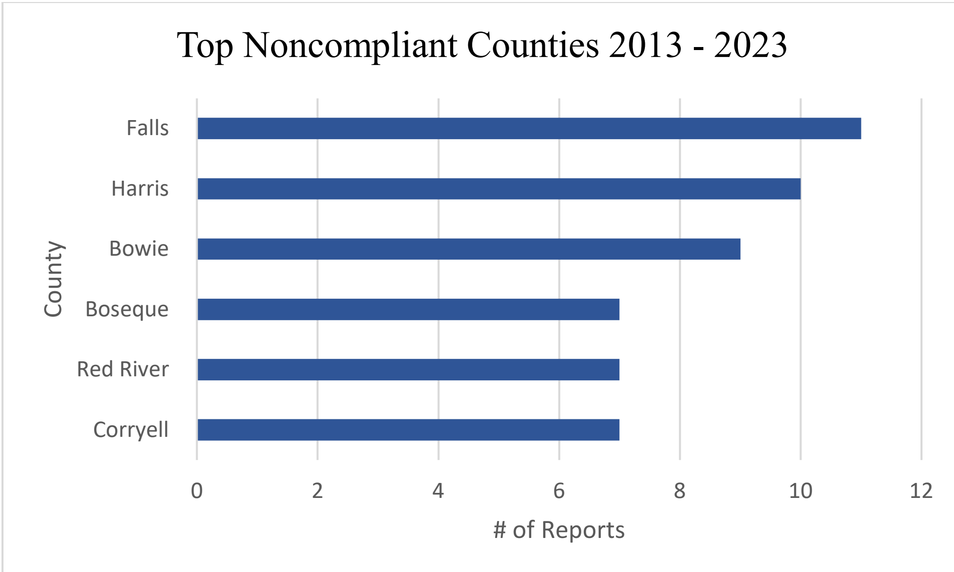
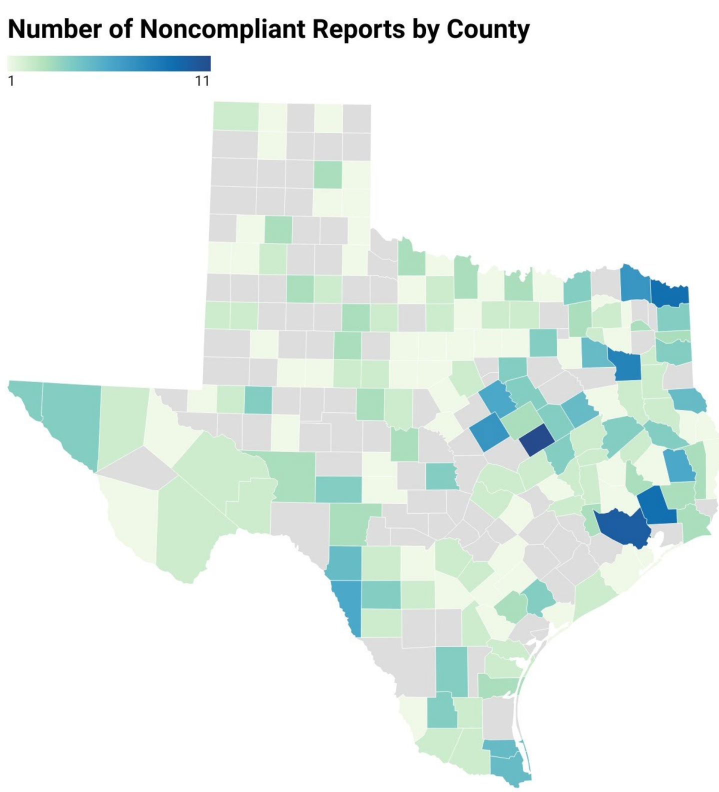


Figure 3. Number of Noncompliant Reports by County [4]



[4] Created using Datawrapper: https://app.datawrapper.de/select/map

Figure 4 presents the counties with reports over 1,000 words as a proxy for the counties with the most violations across reports. Again, Falls and Harris County are the top with Falls County having the longest report by a significant amount (>30,000 words). However, not all the counties with the highest number of noncompliant reports have the most violations across the dataset (comparing Figure 2 and Figure 4), and there is again variation in county size represented. Table 2 presents a series of correlations to further explore the potential relationship between county size and number of noncompliant reports, and county size and the number of violations (proxied by length of report). H1 expects there to be a negative moderate to strong relationship and that is not realized in any of the Pearson's R correlations. This was explored across all counties, by removing Falls County, removing Harris County, and both Falls and Harris County. Both Falls and Harris County are potential outliers in either direction due to county size and high number of violations (Falls County population being <17,000 and Harris County's being > 4.7 million). However, removing these counties does not reveal a relationship outside of their influence. The only negative correlations are number of violations by county size when removing Harris County ($r=-0.03$) and removing both Harris and Falls Counties ($r=-0.06$), both being too weak to warrant further exploration. The strongest correlations are positive for number of reports by county size for all counties ($r=0.24$) and without Falls County ($r=0.26$). Yet these are still too weak to warrant support for a relationship in the opposite direction of H1. Overall, there is not evidence to support Hypothesis 1. Number of noncompliant reports and number of violations are not driven by smaller counties, nor by county size more generally.

Figure 4. Most Violations by Report

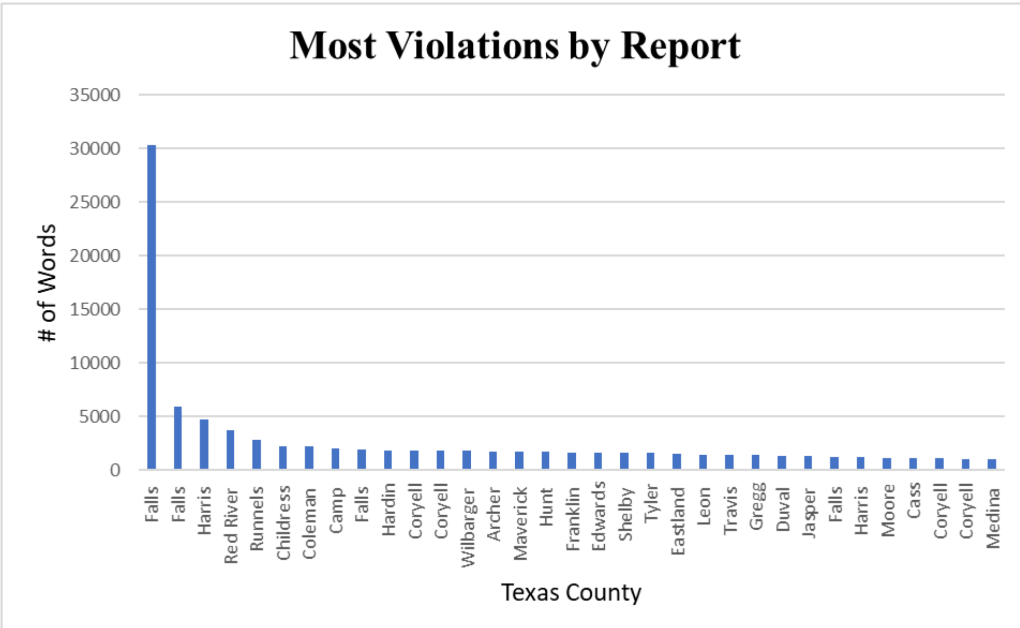


Table 2. Correlations Between County Size and Report Number and Length

No. of Reports	*r*	Violations	*r*
All Counties	0.24	All Counties	0.097
Without Falls Co	0.26	Without Falls Co	0.084
Without Harris Co	0.009	Without Harris Co	-0.03
Without Falls/Harris Co	0.012	Without Falls/Harris Co	-0.06

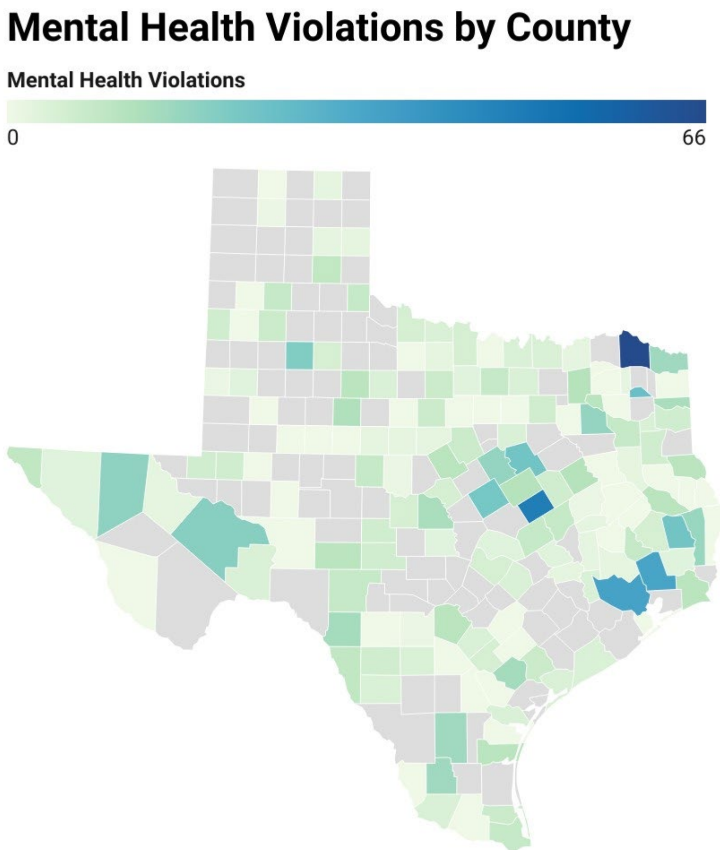
Hypothesis 2

The second hypothesis considers mental health violations through the Dictionary Based Method detailed above. There was a total of 993 mental health words across reports. Figure 5 reveals that Red River County had the most instances of mental health violations, while once again Falls and Harris Counties were among the top violators. The eastern region appears to have the most mental health violations, in correspondence with the overall violations, though the spatial concentration is less pronounced among mental health violations specifically (See Figure

5). The consistency of top violators with overall reports and mental health reports offers some preliminary support for Hypothesis 2. Yet 53% of the reports contained mental health violations (227 reports), while 47% did not (199 reports).

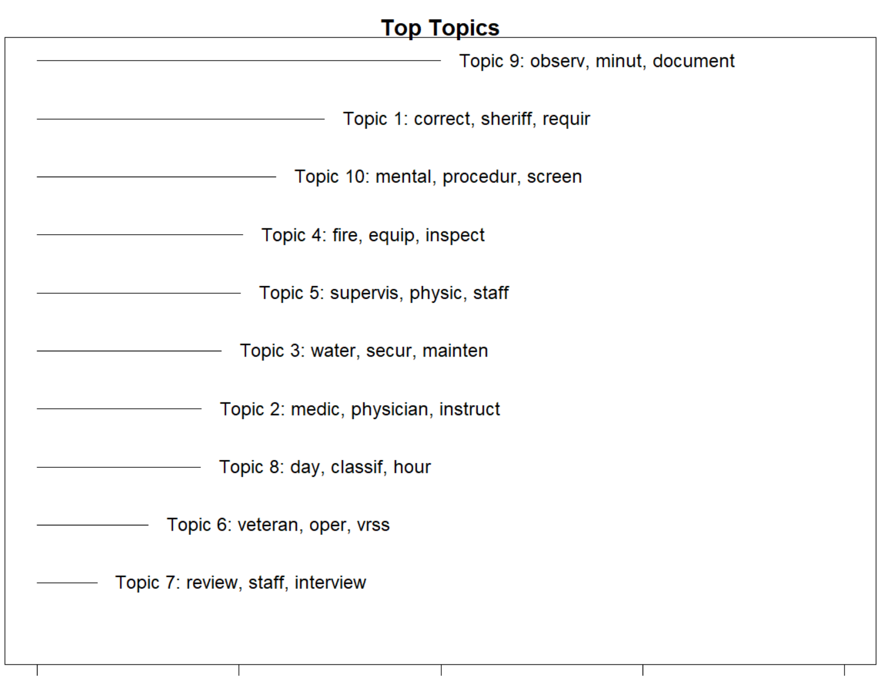
A Structural Topic Model was next considered to see if the top topics captured mental health violations or if other topics emerged. Figure 6 reveals the results of this unsupervised learning method. Though it creates the topics without human coding, it does require some subjectivity in interpretation and familiarity with the documents. The most frequent topic across documents does capture a common mental health violation, which involves not observing a jailed person who is suicidal or considered to have a severe mental health condition. Jail standards require observation every 30 to 90 minutes and often jails are in violation for falsifying these documents compared to video footage. The third topic captures violations around proper screening procedures for mental health. However, the remaining top topics are more general and deal with things like corrective actions to be taken by the sheriff or violations around proper medication, which may or may not be mental health related, and vary based on the context of each specific violation. This STM offers some additional support for Hypothesis 2 but more research is necessary to further understand if mental health violations are the leading cause of rights violations in Texas county jails.

Figure 5. Mental Health Violations by County



County	Words
Red River	66
Falls	48
Harris	34
Liberty	34
Camp	25
Hill	23
Coryell	22
Crosby	20
Pecos	19
Culberson	18

Figure 5. Structural Topic Model



Validation

Length as Proxy for Violations

To validate using the length of reports as a proxy for number of violations, I took three samples of 20 from the longest group of reports, shortest group of reports, and middle group of reports. I intentionally removed the longest report from Falls County, given it is a significant outlier and would bias results at over 30,000 words. After taking this stratified sample, I hand counted the number of violations within each report. Table 3 shows the results from this validation process, with the longest group of reports having an average of six violations, the middle having an average of 2.6 violations, and the shortest having an average of one violation. A difference of means test finds the difference between the longest and shortest group to be statistically significant, $t(19) = 4.34$, $p = 0.000176$. The difference between the middle and shortest samples is also significant, $t(19)=6.02$, $p=0.000004$. And finally, the difference between the longest and middle group is significant, $t(21)=2.82$, $p=0.005$. Though this validation approach suggests that length of document can be used as a proxy for number of violations, I found this to be the most accurate for the shorter and middle documents. Some of the longest documents included attachments and had follow-up or repeat language that did not reflect additional violations. This was a helpful discovery to inform next steps, as it might make sense to section out the documents to only cover content about violations and details about the violations.

Table 3. Average Number of Violations from Sampled Report Length Groups

Sampled Group	*M* =
Longest	6
Shortest	1
Middle	2.6

Mental Health Dictionary

After taking a random sample of 86 reports, I hand counted the number of mental health concepts and violations discussed in each report. The dictionary model did well on rarely under-counting, with only four instances of undercounting mental health words across the 86 reports. However, it overcounted 27 times. The undercounting was due to a few instances of poor document quality and the OCR process not capturing words correctly. The overcounting was due to patterns in some of the report language. For example, some rely on standard language and state things like, "failed to complete the screening form for Suicide and Medical/ Mental/ Developmental Impairments." I hand coded this as one instance of a mental health concept and violation, while the dictionary picked up on both suicide and mental and counted it as two separate concepts. Despite these instances of under and overcounting, a correlation of the dictionary words and my count was *r*=.993.

To further to validate for both precision and recall, I created a Confusion Matrix. I coded each of the sampled documents by whether the dictionary identified any mental health violations and then coded whether I did, for a binary coding of either 1 for yes or 0 for no. Table 4 shows the results of the Confusion Matrix from this process, revealing 46 true positives, 3 false positives, 2 false negatives, and 35 false positives. The precision score of the mental health dictionary = .939.[5] The recall score of the mental health dictionary = .958.[6] Finally, an F1 score was calculated to understand the balance between precision and recall (Shung, 2018). The F1 Score = .948.[7] Therefore, despite some issues of overcounting because of violation type phrasing and document quality issues, the model has a high level of both precision and recall.

Table 4. Confusion Matrix

[5] Precision = (True Positive / True Positive + False Positive).
[6] Recall = (True Positive / True Positive + False Negative).
[7] F1 = 2 x (Precision*Recall / Precision + Recall)

		Predicted	
		1	0
	1	46	3
Actual	0	2	35

Next Steps and Limitations

Both the Dictionary Based Method and length of documents as a proxy for number of violations helped to reveal the pattern of violations occurring in Texas county jails. County size was ruled out as driver of violations, but more analysis in needed to fully understand the nature of mental health violations in Texas county jails. The validation revealed the need to improve the OCRed version of some of the documents. Some have been scanned in and were folded when they were scanned. Others had writing over the reports or stamp imprints. These data quality issues are what drove the false negatives in the dictionary approach. Further, the assumption that longer documents had more violations held in the aggregate, but the validation process revealed additional contextual documents that are sometimes included within reports. Future work will need to standardize the documents, pulling out the violation type and explanation within each report and removing letters to judges or repetitive texts. Though the Structural Topic Model was used as a validation method for the mental health dictionary, future work should validate this model and determine which model works best for understanding the themes of rights violations in the documents.

Future work will also incorporate case studies to better understand the pattern of violations occurring in Falls County, Harris County, and Red River County. Each of these were present in the counties with the most reports, the most violations, and the most mental health violations. Reading each of the reports fully for the counties, and exploring these counties' issues qualitatively, could add important context to the quantitative text-as-data project.

Finally, future work could connect these noncompliant reports with other important indicators of human rights violations in jails. A next step will be to create a dataset from the Texas Commission on Jail Standards annual reports to track things like assaults, complaints, and staff turnover in the aggregate to see if noncompliance reports are capturing and helping to reduce the instances of harm and death in jail. Given that the literature on prison oversight recommends third-party oversight, a state regulatory body may be limited in what it is tasked with overseeing, and the efficacy of change the specific standards impact.

Conclusion

There is a need for more research to understand human rights abuses in local jails. A lack of data, along with great variation across jails makes studying these institutions challenging. This text-as-data report presents a promising approach to transform oversight reports into a useful dataset. Noncompliant jail reports created by inspectors for the Texas Commission on Jail Standards showed that violations have increased over time, and that similar counties are committing the most violations, are violating most frequently, and are committing the most mental health violations. Size of county was not found to drive the number or frequency of violations. Instead, it appears that the eastern region of the state has more violations occurring. A dictionary-based method informed the frequency and context of mental health violations, lending preliminary support that these violations are driving noncompliance. Validation efforts confirmed using the length of documents as a proxy for number of violations, but this process highlighted issues with additional information in the documents that needs to be addressed in the next steps of this research. The dictionary based approached was also validated with a high level of both precision and recall, but document quality impacted a small number of false negatives and violation type phrasing contributed to over-counting. Future research will work to address the limitations found through the validation tests and expand to both qualitative assessments and connecting to larger text-as-data projects. Given the substantial percentage of people in jails experiencing serious mental health issues, along with high suicide rates in jails, it is essential to understand how and where these mental health violations are occurring in order to inform interventions.

References

Adler, J.L. & Chen, W. (2023). Jail conditions and mortality: Death rates associated with turnover, jail size, and population characteristics. *Health Affairs, 42*(6). https://doi.org/10.1377/hlthaff.2022.01229

Attorney General of Texas. (2023). Custodial Death Report. https://oag.my.site.com/cdr/cdrreportdeaths

Bierie, D.M. (2011). Is tougher better? The impact of physical prison conditions on inmate violence. *International Journal of Offender Therapy and Comparative Criminology, 56*(3). https://doi.org/10.1177/0306624X11405157

Bureau of Justice Statistics (2021). Mortality in Local Jails, 2000-2019 – Statistical Tables. *U.S. Department of Justice, Office of Justice Programs.* https://bjs.ojp.gov/content/pub/pdf/mlj0019st.pdf

Copp, J.E., & Bales, W.D. (2018). Jails and local justice system reform: Overview and recommendations. *The Future of Children, 28*(1). https://www.jstor.org/stable/26641549

Denny, M.J & Spirling, A. (2018). Text preprocessing for unsupervised learning: Why it matters, when it misleads, and what to do about it. *Political Analysis, 26*(2). https://dx.doi.org/10.2139/ssrn.2849145

Deitch, M. (2020). But who oversees the overseers?: The status of prison and jail oversight in the United States. *American Journal of Criminal Law, 47*(2). https://heinonline.org/HOL/LandingPage?handle=hein.journals/ajcl47&div=13&id=&page=

Dye, M.H. (2010). Deprivation, importation, and prison suicide: Combined effects of institutional conditions and inmate composition. *Journal of Criminal Justice, 38*(4). https://doi.org/10.1016/j.jcrimjus.2010.05.007

Ellison, J.M., Cain, C.M., Baker, B., & Paige, B. (2022). The dangers of short-term confinement: Indicators of Safety Risks Among Individuals in Jails. *International Journal of Offender Therapy and Comparative Criminology*. https://doi.org/10.1177/0306624X221110

Fathi, D.C. (2010). The challenge of prison oversight. *American Criminal Law Review, 47*(4). https://heinonline.org/HOL/LandingPage?handle=hein.journals/amcrimlr47&div=50&id=&page

Grench, E. (2023, August 7). "A place of torment": 22 families, former inmates sue Harris County over jail conditions. *The Texas Tribune.* https://www.texastribune.org/2023/08/07/harris-county-jail-deaths-injuries-lawsuit/

Grimmer, J., Roberts, M.E., & Stewart, B.M. (2022). *Text as Data: A New Framework for Machine Learning and the Social Sciences.* Chapter 16: Word Counting. Princeton University Press.

Haugen, M. (2018). In rural areas, jail populations are skyrocketing – including pretrial detainees. *Texas Public Policy Foundation*. https://www.texaspolicy.com/in-rural-areas-jail-populations-are-skyrocketing-including-pretrial-detainees/

Kajeepta S., Mauro, P.M., Keyes, K.M., El-Sayed, A.M., Rutherford, C.G., & Prins, S.J. (2021). Association between county jail incarceration and cause-specific county mortality in the USA, 1987-2017: a retrospective, longitudinal study. *Lancet Public Health*, *6*(4): 240-248. doi: 10.1016/S2468-2667(20)30283-8.

Lebryk, T. (2021). Introduction to the Structural Topic Model (STM): A unique way to use topic modelling for social science research. *Towards Data Science. Medium.* https://towardsdatascience.com/introduction-to-the-structural-topic-model-stm-34ec4bd5383

Meagher & Chammah, 2015 Meager, T. & Chammah, M. (2015). Why jails have more suicides than prisons. *The Marshall Project.* https://www.themarshallproject.org/2015/08/04/why-jails-have-more-suicides-than-prisons

Reiter, K., Ventura, J., Lovell, D., Augustine, D., Barragan, M., Blair, T., Chesnut, K., Dashtgard, P., Gonzalez, G., Pifer, N. & Strong, J. (2019). Psychological distress in solitary confinement: Symptons, Severity, and Prevalence in the United States, 2017-2018. *American Journal of Public Health, 110*. https://doi.org/10.2105/AJPH.2019.305375

Ross, J.I. (2011). Moving beyond Soering: U.S. prison conditions as an argument against extradition to the United States. *International Criminal Justice Review, 21*(2). https://doi.org/10.1177/1057567711408083

Shung, K.P. (2018). Accuracy, precision, recall or F1? *Towards Data Science. Medium.* https://towardsdatascience.com/accuracy-precision-recall-or-f1-331fb37c5cb9

Silver, J. (2017, June 15). Texas Gov. Abbott signs "Sandra Bland Act" into law. *The Texas Tribune.* https://www.texastribune.org/2017/06/15/texas-gov-greg-abbott-signs-sandra-bland-act-law/

Stuckey, A. (2023, May 9). Search our data: 52 lawsuits filed by Harris County Jail inmates, families over treatment and conditions. *Houston Landing.* https://houstonlanding.org/search-our-data-harris-county-jail-inmates-families-sue-over-treatment-and-conditions/

Torrey, F.E., Kennard, A.D., Eslinger, D., Lamb, R. & Pavie, J. (2010). More mentally ill persons are in jails and prisons than hospitals: A survey of the states. *National Sheriffs' Association, Treatment Advocacy Center.* https://www.ojp.gov/ncjrs/virtual-library/abstracts/more-mentally-ill-persons-are-jails-and-prisons-hospitals-survey

Wang, L. (2021). Rise in jail deaths is especially troubling as jail populations become more rural and more female. *Prison Policy Initiative.* https://www.prisonpolicy.org/blog/2021/06/23/jail_mortality/

Worsley, M.K. & Memmer, A. (2017). Transparency behind bars: A history of Kansas jail inspections, current practices, and possible reform. *Journal of Criminal Justice and Law, 1*(2). https://jcjl.pubpub.org/pub/v1-i2-worsley-memmer-jails-oversight-enforcement/release/2