rencì

NC COVID-19 Data Travels

2021 Report

Adaptive Data Utilization During a Pandemic: A Report of COVID-19 Data Processes in North Carolina

Renaissance Computing Institute (RENCI)
University of North Carolina at Chapel Hill (UNC-Chapel Hill)

PROJECT TEAM

Ashok Krishnamurthy, Deputy Director, RENCI (Project Principal Investigator; Interviewer)
Stanley C. Ahalt, Director, RENCI (Interviewer)
Jay Aikat, Chief Operating Officer, RENCI (Interviewer)
Joanna (Asia) Mieczkowska, Associate Operations Officer, RENCI (Organizer, Interviewer, Reporting)
Julie Hayes, Executive Assistant, RENCI (Writer, Administrative Support)
Stefania Knight, Administrative Assistant, RENCI (Coordination, Administrative Support)
Paul Mihas, Assistant Director of Education and Qualitative Research, UNC Odum Institute (Subject Matter Expert, Consultant)
Sharita Thomas, Research Associate, UNC Sheps Center (Subject Matter Expert, Analysis)
Donna Shaw, Qualitative Data Analyst, NCSU graduate student (Analysis)
Jennifer Rees Patterson, Qualitative Data Analyst, consultant (Analysis)
Christopher Daley, Administrative Assistant, RENCI (Research Support)
Marcus Anderson, Graphic and Visual Design Specialist, RENCI (Graphics Support)
Kelsey Urgo, Freelancer (Graphics Support)
Stephanie Suber, Communications and Outreach Manager, RENCI (Formatting)

FUNDING STATEMENT

This report is supported by the University of North Carolina at Chapel Hill's NC Policy Collaboratory through funds from a State appropriation by the North Carolina General Assembly (NCGA) per subdivision (23) of Section 3.3. of Session Law 2020-4 (S.L. 2020-4; H1043) effective May 4, 2020 and per the federal CARES Act.

ACKNOWLEDGEMENT

This report would not have been possible without the participation and valuable input from the interviewees. We extend our sincere thanks to those who took the time to meet with us and answer our questions.

TABLE OF CONTENTS

Executive Summary 3
Introduction 7
Methods 9
Results 12
Discussion 24
References 28
Appendices 31
Appendix A. Timeline of State and Federal COVID-19 Legislation 31
Appendix B. Interview Request 33
Appendix C. Interview Guide 36
Appendix D. Glossary 38

EXECUTIVE SUMMARY

As we face the COVID-19 pandemic caused by the novel coronavirus, SARS-COV-2, medical professionals, technologists, community leaders, and policymakers must leverage data as both an adaptable surveillance tool and as a means of community-level education. Private and public organizations are collecting a variety of COVID-19-related data from a multitude of sources across North Carolina. Simultaneously, these organizations have been called on to use data to build sustainable and robust predictive models while also grappling with the current pandemic. Meanwhile, the average North Carolinian may have wondered how to make sense of the role data plays in managing the COVID-19 pandemic.

This report summarizes the results and observations from key informant interviews with a diverse group of experts from across North Carolina. We provide recommendations that may be useful in the future. Our purpose was to understand and document the evolution of North Carolina's COVID-19 response, with a focus on gaining a better understanding of COVID-19 data sources; data collection and reporting protocols and objectives; data uses and dissemination; data aggregation and centralization; COVID-19 testing; and North Carolina's response in comparison with surrounding states. It is important to note that even as we were conducting interviews and analyzing data over the last several months, the pandemic situation is dynamic and has been continuously evolving. Thus, the processes followed for data collection, analysis, and reporting about COVID-19 were continuously reformed by the various institutions and agencies involved. With this initial report on COVID-19 data processes, we hope to aid readers in understanding how the different pieces fit together and where there exist significant gaps for the purposes of both addressing our current pandemic and preparing for future public health crises.

From June through December 2020, the project team conducted a total of 30 in-depth, semi-structured interviews with 41 key informants via video conference. The final list of interview participants included representation from public, private, university, and government-level health and research organizations across the state of North Carolina.

Interviewees were asked a series of approximately 20 standardized questions covering the following focus areas: Data sources, data reporting, data modeling, and data connections. Transcripts were imported and organized into NVivo versions 11 and 12 for analysis. Three team members knowledgeable and experienced in qualitative research methods independently coded the transcripts. Inductive themes were identified through the techniques of cutting and sorting, repetition, and similarities and differences.

Based on the analysis, the project team compiled the following recommendations to improve pandemic response and better prepare for future public health crises:

1. **Future pandemic response should be centralized through the North Carolina Department of Health and Human Services.**
   Standardized and coordinated information sharing is the foundation of effective pandemic response. Key informants expressed their appreciation for the leadership exemplified by NCDHHS following the COVID-19 outbreak and a desire for an even more streamlined process when preparing for and responding to future pandemics. Key informants expressed frustration over requirements imposed by the federal government that were made without appropriate guidance and with very short timelines for compliance. Even in such cases, the key informants emphatically asserted that the leadership and coordination provided by NCDHHS helped alleviate these difficult circumstances.

2. **The state should work to support and sustain the cross-sector collaborative networks established during the COVID-19 outbreak.**
   A consistent theme mentioned by key informants is that cross-sector collaboration was initially a major facilitator in the collection and use of COVID-19-related data. NC was particularly well-positioned to collect data regarding the pandemic because of prior collaborations, for example, the North Carolina Disease Event Tracking and Epidemiologic Collection Tool (NC DETECT) (NCDHHS & UNC School of Medicine, 2021). Thus, many of the new collaborations developed from prior and existing relationships and a desire to maximize the combined impact of the work being done by colleagues working at different institutions. NC is fortunate to have a number of strong research institutes and will benefit from formalizing many of the collaborative networks that have organically developed since March and encourage even more data synergy and consistency in data collection processes moving forward.

3. **Pandemic-related data should be publicly accessible and available in a format that is easy to download and utilize.**
   Proactive data collection and timely analysis facilitates identification of patterns and the dissemination of timely information. To increase access to such data, NC should publicly release data in an easy to download format to not only inform the public but also to facilitate analysis by the broad community of data scientists. Further, data should be translated in a manner useful to both data scientists and the greater public.

4. **Contact tracing and surveillance methods need to be more robust and systematic.**
   Early detection of cases remains the most important factor in effective management of an infectious outbreak. This effort becomes more challenging for communities that are more remote and/or economically disadvantaged. Development of a standardized protocol that supports extensive testing and reporting will empower communities by providing them with guidance on best practices and appropriate response in the event of an infectious outbreak. Starting a process now to develop and strengthen community-level or city/town-level surveillance methods should function to limit the impact of future infectious outbreaks by increasing the chances for early detection. This, in turn, should minimize economic disruptions at the source of the problem. Robust and practical guidelines for clinical and administrative workforces will need to be created and standardized at the state level, and then effectively communicated and funded.

Further, the project team created the following figure describing the data flow of COVID-19 data across NC based on the interviews.

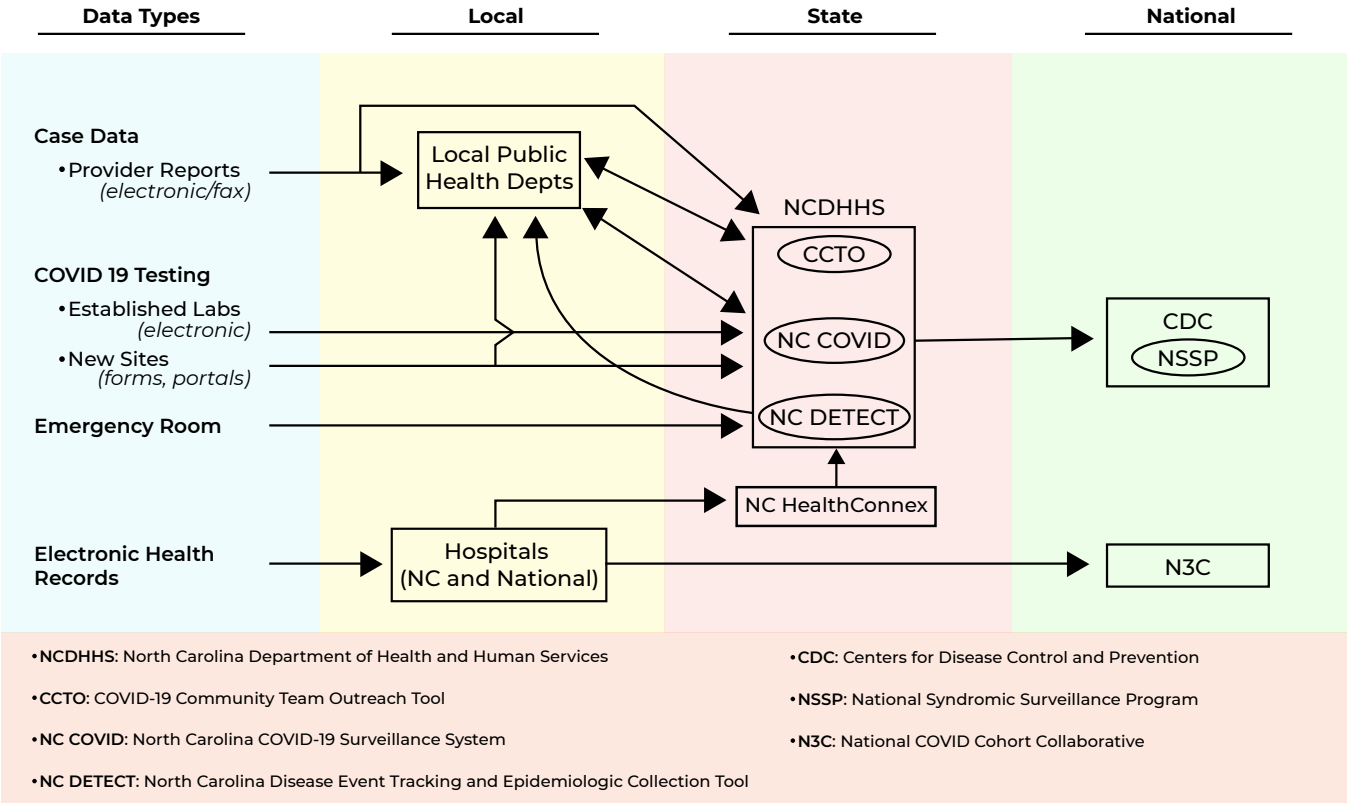


Figure 1. COVID-19 Data Flow in NC

In NC, the sources of COVID-19 data include case data, COVID-19 testing data, emergency room data, and data from electronic health records (EHRs).

Case data, or data from COVID-19 case investigations, are in the form of medical provider reports, sent both electronically and via fax to local health departments and NCDHHS. The local health departments then have a two-way flow of case data with NCDHHS' **COVID-19 Community Team Outreach (CCTO) Tool** for tracing efforts, and NCDHHS' **COVID-19 Surveillance System (NC COVID).**

COVID-19 testing data are gathered from established labs electronically and from new testing sites via forms and newly developed portals; the data are sent to local health departments and NC COVID.

Emergency room data are sent directly to the **North Carolina Disease Event Tracking and Epidemiologic Collection Tool, NC DETECT,** NCDHHS' long-established platform for intaking emergency room data. NC DETECT then communicates the emergency room data to local health departments.

COVID-19 data from EHRs reside in hospitals, both locally and nationally, and the NC data are sent from local hospitals to the NC HealthConnex platform, NC's Health Information Exchange (HIE) system, and the National COVID Cohort Collaborative (N3C); N3C is a partnership between the National Center for Advancing Translational Sciences (NCATS) and the National Center for Data to Health (CD2H) to create a centralized national data resource, the NCATS N3C Data Enclave, which

allows rapid collection and analysis of COVID-19 data (National Center for Advancing Translational Sciences [NCATS], 2021). NC HealthConnex also communicates this information back to NCDHHS. All of the COVID-19 data received by NCDHHS then gets communicated at the federal level to the Centers for Disease Control and Prevention (CDC) and the National Syndromic Surveillance Program (NSSP).

The project team also set out to analyze NC's COVID-19 response in comparison to the neighboring states of South Carolina, Georgia, Tennessee, and Virginia by comparing the reporting metrics available on the COVID-19 dashboards of each state.

This work provides insight into the data collection, reporting, dissemination, and collaboration efforts among the key stakeholders in North Carolina. We hope that these observations and recommendations will inform COVID-19 policy and legislation to improve processes and, ultimately, the many lives impacted by COVID-19 in North Carolina and beyond.

INTRODUCTION

On January 30, 2020, the World Health Organization (WHO) declared the outbreak of the novel coronavirus disease, known then as "2019-nCoV," a public health emergency of international concern (World Health Organization [WHO], 2020a). First identified in Wuhan, China, the virus has become a global pandemic, with over 107 million recorded cases and 2.36 million deaths reported worldwide as of January 2021 (Johns Hopkins University [JHU], 2021). In the U.S., there are currently over 27.4 million cases and 473,000 deaths (Centers for Disease Control and Prevention [CDC], 2020).

Similarly, in 2003, the WHO issued a global alert concerning the outbreak caused by the SARS-Coronavirus (WHO Media Centre, 2003). At that time, the Centers for Disease Control and Prevention (CDC), in collaboration with national and international organizations, produced guidance for an appropriate pandemic response at the local public health level. The CDC identified the following as key components of preparedness and response as learned from the 2003 SARS outbreak (Centers for Disease Control and Prevention [CDC], 2004):

- Up-to-date information on local, national, and global SARS activity;
- Rapid and effective institution of control measures;
- The appropriate resources and organizational and decision-making structure; and
- Trained staff vital to swift and decisive implementation.

While the WHO previously established recommendations for how best to prepare and handle a pandemic, few were truly prepared for the exceptionally rapid and widespread impact of the COVID-19 virus. As COVID-19 spread throughout the world, the WHO officially declared COVID-19 a global pandemic on March 11, 2020 (WHO, 2020b), and policymakers at every level were ultimately forced to recognize the severity of the virus and take action to mitigate the spread.

At the federal level, the first U.S. piece of legislation was the "Coronavirus Preparedness and Response Supplemental Appropriations Act, 2020," which was enacted on March 6, 2020; this functioned to provide emergency funding to the Food and Drug Administration (FDA), the Small Business Administration's Disaster Loans Program Account, the Centers for Disease Control and Prevention (CDC), and other key government divisions and programs.

The next major piece of federal legislation was the "Families First Coronavirus Response Act" enacted on March 18, 2020, which focused on expanding family and medical leave, unemployment benefits, and other government assistance programs; it also required insurance companies to cover the cost of COVID-19 testing.

The third piece of federal legislation was the "Coronavirus Aid, Relief, and Economic Security Act (CARES Act)" enacted on March 27, 2020--this was the most notable of the federal legislation passed up to that point, as it suspended student loan payments and sent a $1,200 stimulus check to all qualifying citizens, along with other mitigation strategies.

These first three pieces of federal COVID-19 legislation passed in March set the stage for subsequent legislation passed over the following months, which have involved new and evolved strategies to mitigate the now ubiquitous impact of the virus. Most recently, on December 21, 2020, the "Consolidated Appropriations Act, 2021" was passed by Congress, providing $600 stimulus checks to qualifying citizens and increased funding to various government offices (see Appendix A).

In North Carolina, Governor Cooper declared a State of Emergency on March 10, 2020 in Executive Order No. 116. Notably, this Executive Order established the Governor's Novel Coronavirus Task Force on COVID-19 and authorized the hiring of temporary employees to support the NC Department of Health and Human Services (NCDHHS) and local health departments. Further, the Executive Order authorized and ordered the State Health Director to provide recommendations on travel restrictions for state employees, work with local health departments to develop COVID-19 surveillance and control measures, and maximize availability of COVID-19 testing (NC Exec. Order No. 116, 2020).

Governor Cooper then passed a series of Executive Orders limiting mass gatherings (NC Exec. Order No. 117, 2020), requiring the closure of statewide public schools (NC Exec. Order No. 117, 2020), limiting the operations of restaurants and bars (NC Exec. Order No. 118, 2020), and expanding unemployment benefits (NC Exec. Order No. 118, 2020). On March 27, 2020, Governor Cooper enacted Executive Order No. 121, which was the noteworthy "Stay at Home Order," requiring North Carolinians to only leave home for essential purposes; this Executive Order also forced non-essential business operations to cease.

This report examines North Carolina's response to the COVID-19 pandemic through an analysis of:

- COVID-19 data sources
- Data collection and reporting protocol and objectives
- Data uses and dissemination
- Data aggregation and centralization
- COVID-19 testing
- North Carolina's pandemic response in comparison with surrounding states

In May, Phase One of the reopening plan began, easing restrictions on travel, gatherings, and business operations (NC Exec. Order No. 138, 2020). While NC attempted to progressively "reopen" over the next few months, Governor Cooper declared a modified Stay at Home Order and a curfew for certain businesses and activities on December 8, 2020 as COVID-19 cases continued to increase in the state (NC Exec. Order No. 181, 2020). The most recent executive order, No. 190, lifted restrictions on restaurant take-out of mixed beverages (see Appendix A).

In the following sections, we have outlined our methods for conducting the study, the results of the analysis, lessons learned across the study, as well as a meaningful discussion of the results. Based on the results, we have provided key recommendations for potential policy and legislative changes at various levels to improve NC's pandemic response, both for the present circumstances and future public health crises. Further, we have included appendices containing supporting figures and materials. Our hope is that this report will aid readers in understanding the full picture of this complex public health crisis to best address current gaps in data collection, reporting, and dissemination as well as improve the interconnectivity and intercommunication among the relevant stakeholders across NC.

METHODS

From June through December 2020, the project team conducted a total of 30 in-depth interviews with 41 key informants via the Zoom video cloud platform. The purpose of the in-depth interviews was to understand and document the evolution of North Carolina's COVID-19 response, with a focus on gaining a better understanding of COVID-19 data sources; data collection and reporting protocol and objectives; data uses and dissemination; data aggregation and centralization; COVID-19 testing; and North Carolina's response in comparison with surrounding states.

INTERVIEWERS

The semi-structured interviews were conducted by RENCI project team members: Stan Ahalt, Director; Ashok Krishnamurthy, Deputy Director; Jay Aikat, Chief Operating Officer; Asia Mieczkowska, Associate Operations Officer; Sharita R. Thomas, Qualitative Research Consultant; and Donna Shaw, Qualitative Data Analyst. Real-time transcriptions of the interviews were captured by Stefania Knight and Julie Hayes.

KEY INFORMANT INTERVIEW RECRUITMENT & PARTICIPANTS

Key informants were identified as experts in their fields who were known to be involved with COVID-19-related data. Potential interviewees were identified through a series of steps, including project team discussions, external peer consultations, and web-based searches. Prior to conducting interviews, the project team met to prioritize the list of potential interviewees based on their involvement in and proximity to COVID-19 data. A snowball sampling approach was utilized to recruit key informants beyond the initially identified expert group. The interviews were not intended to be statistically representative of the state, and the number of interviewees does not affect the integrity of data collected. However, we attempted to obtain coverage from all regions of North Carolina to account for geographic and demographic differences. Out of the final interview participants, 40% of key informants were women.

After identifying potential interview participants, the project team prioritized and randomly assigned interviews among the project team interviewers. The interviewers contacted their assigned interview participants via email to provide consistent information of the project aims and the interview request (see Appendix B). However, there was concern with email fatigue and response rate, so the standardized email was edited to be more concise mid-way through the interview participant contact process; the original standardized email was included as a supplemental attachment of additional information (see Appendix A). Key informant recruitment ended once saturation was reached in response data. Participation was voluntary, and consent to record and transcribe responses was obtained prior to the start of the interview. The final list of interview participants included 41 key informants across 30 interviews.

STRUCTURED INTERVIEW GUIDE

The project team met with qualitative research consultants to develop the interview guide (see Appendix C for both the initial and revised interview guides). The interview guide was developed through guidance from the funder, pilot interviews of external key informants, and from grey literature and environmental scans. The questions were designed around the following domains: Data sources; data collection and reporting protocol and objectives; data uses and dissemination; data aggregation and centralization; and COVID-19 testing.

INTERVIEWS

Key informant interviews were conducted in an open-ended, semi-structured format. The interviewer had flexibility to omit and/or rephrase questions based on context and expertise of the key informant. Interviews of 30 to 45 minutes were conducted by lead project researchers over the online video conferencing platform, Zoom. Interviews were automatically recorded and stored in the cloud via the Zoom platform. Additionally, project associates attended each interview to transcribe and take notes in real-time.The first four interviews were conducted as pilot interviews before the official interview guide was established. The subsequent 11 interviews were conducted using the initial interview guide (see Appendix C). After reformatting and trimming the questions, the following 18 interviews were conducted using the newly condensed question format of 17 high priority questions (see Appendix C).

CONFIDENTIALITY

Interview participants were informed of the voluntary nature of this project and that recordings would be transcribed for analyses by the project team. Interview participants were informed that the recordings would be deleted after the conclusion of the study and would not be shared outside of the project team. Further, interview participants were informed that the content of the interview would be de-identified, and any information used in the report would not cite an interviewee by name unless permission was given voluntarily. The method of reporting data in aggregate encourages candid responses and positive rapport.

ANALYSIS

Four pilot interviews were conducted to inform the development of the interview guide. Subsequent interviews were recorded and transcribed verbatim. Scribes attended each interview to transcribe in real time and subsequently reviewed and edited transcripts for accuracy using the recordings.

Transcribed data was imported and analyzed using the qualitative data analysis software NVivo, versions 11 and 12. Transcripts were read and coded using an a priori approach. Three team members knowledgeable and experienced in qualitative research methods independently coded transcripts. Analysis team members set regular meetings to compare, review, and refine codes. Discrepancies in codes were resolved through discussion. Emerging themes and coding memo notes were also shared and discussed as a group.

The method of content analysis utilized aligns with directed content analysis. Directed content analysis allows for themes to arise directly from the data, known as the inductive approach, but initial coding is guided by previous theory or previous research. As such, some of our codes were previously defined from the interview guide and were established prior to analysis, while the remaining codes were defined during analysis. Inductive themes were identified through the techniques of cutting and sorting, repetition, and similarities and differences.

Rigor was ensured through: (1) Triangulating different sources of data (key informant interviews, literature and grey literature review, and notes); (2) employing independent coding of transcripts and inter-coder agreement; and (3) utilizing an iterative data collection and analysis process.

COMPARING NORTH CAROLINA'S REPORTING METRICS

Regarding North Carolina's reporting efforts and metrics, the COVID-19 dashboards of the neighboring states of South Carolina, Georgia, Tennessee, and Virginia were analyzed and compared with the reporting metrics and data available on the NCDHHS website. This includes information available on the homepage of each dashboard, as well as subsections (tabs) that are easier for users to navigate and access via the homepage. The number of new cases and deaths related to COVID-19 were compared to the numbers reported by the New York Times and Johns Hopkins University, as both of these national institutions report the number of new cases and deaths in all 50 states. The purpose of this particular analysis was to gauge how accurate NCDHHS reporting is compared to large international institutions that utilize a wide variety of resources for their reporting.

RESULTS

Key informants indicated that their initial COVID-19 data processes were primarily ad hoc, driven simultaneously by curiosity and responsibility to patient and community welfare. Later, these processes evolved in response to state and federal reporting mandates as well as via cross-sector collaboration. Responses from key informants are summarized across interviews to present common themes. These themes can be used to inform communities, researchers, and policymakers in North Carolina with the goal of improving the current COVID-19 response and increasing preparedness for future pandemic and disaster response planning.

COVID-19 DATA SOURCES

Data collection and reporting is vital for accurate and effective planning during a pandemic. Key informants indicated several sources for the COVID-19 data collected.

PRIMARY DATA SOURCES

Primary data are data that have been generated by the interview participant or their organization. It can be qualitative, like surveys or interviews; observational, like experiments; or individual patient encounter data. Most primary data are private and usually used internally. The main source of primary COVID-19 data mentioned by interview participants was their healthcare facilities' Electronic Health Records (EHR) systems. EPIC EHR systems were the most frequently mentioned. EPIC Clarity, a Third Normal Form (3NF) data warehouse and relational database (Oracle or Microsoft SQL database), was also mentioned by one participant as being used to create data-intensive internal reports. Further, one interview participant collected qualitative primary data through surveys and interviews of local government officials.

SECONDARY DATA SOURCES

Almost all interview participants mentioned secondary data sources, which are data that exists or was previously generated by organizations such as government institutions, healthcare facilities, or institutions of education. Secondary data can be "public," meaning it is accessible to anyone in the general public, or it can be "private," meaning one needs special permissions to access the data. Secondary news media data are one such publicly accessible source mentioned by many of the interview participants. News media sources mentioned include The New York Times COVID-19-data on GitHub (The New York Times [NYT], 2021), The New York Times mask wearing survey data (Katz et al., 2020), The News & Observer (The News & Observer [N&O], 2021), and WRAL.com (WRAL, 2021). The data from The New York Times was cited by multiple interview participants who expressed its importance in understanding regional differences apparent in the county-level data as well as regional time trends.

One interview participant mentioned that insurance claims data from BlueCross BlueShield or Medicaid was not a good source because of "data lag" or data latency. Data lag, a common data problem for researchers, refers to the difference in time from when an event happens or is reported to when the relevant data becomes available for use. Insurance claim data, which can provide insight on individual level interactions with health systems, often lags by three to six months (Majumder & Rose, 2020). Other secondary data sources mentioned include the North Carolina Electronic Disease Surveillance System (EDSS) data feed of the NCDHHS Division of Public Health (North Carolina Department of Health and Human Services [NCDHHS], 2021c), and the North Carolina Disease Event Tracking and Epidemiologic Collection Tool (NC DETECT) (NCDHHS & UNC School of Medicine, 2021). NC DETECT is the state syndromic surveillance system that has long been used by hospitals to report emergency department data.

Secondary COVID-19 Data Sources Used by Interviewees:

- News media
- Online mapping tools
- State Surveillance systems
- Geolocation software
- Review of scientific literature
- Annual demographic poll data
- Press releases from the private sector

Additional secondary sources utilized for COVID-19 data activities included: SafeGraph (SafeGraph, 2021), scientific literature, annual demographic poll data, PolicyMap (PolicyMap, 2021), and mobility and weather data through web scraping. One interviewee mentioned scanning for manufacturer press releases to remain informed on ventilators and other personal protective equipment (PPE).

OBJECTIVES & INFLUENCES FOR COVID-19-RELATED DATA COLLECTION

TIMEFRAME OF DATA COLLECTION

While the data collection start dates ranged from early February through July, most interview participants indicated that COVID-19 data collection for them or their companies began in mid- to late-March. No interview participant mentioned having a pre-established protocol to guide the data collection process or objective, but all mentioned feeling compelled to do something, either for their patients or to be proactive. Of those interviewed, the general trend was that they started assessing the availability of personal protective equipment (PPE) in February through mid-March 2020. This evolved so that in March, more complex systems were in place to focus on hospitalizations and modeling capacity. COVID-19 tests were not available in hospitals in March, but they were still able to report any "patient under investigation" (PUI) to NCDHHS. Interviewees who used surveys to collect COVID-19-related data were able to do so starting in April.

INITIAL DATA COLLECTION OBJECTIVES

Interview participants were asked about their initial objectives for COVID-19 data collection. A common initial objective for collecting COVID-19 data was the need to monitor hospital resource supply and utilization, including tracking intensive care unit (ICU) volumes, negative pressure rooms, positive COVID-19 patients, and PPE supply burn rates. Other interview participants collecting primary qualitative data sought to observe regional workforce issues and then disseminate information on best practices for topics like utilizing and billing telehealth, safely preparing facilities for patient care, and acquiring and managing PPE. Additionally, some interviewees noted the initial objective to protect students and employees on college campuses, while another objective was advising regional organizations. Other objectives mentioned centered on understanding the impact of COVID-19 on local governments, namely, reductions in revenue and workforce disruptions from telecommuting. One group had the initial objective of collecting data on daycares that were still open at the beginning of the pandemic to assist frontline workers and others in NC with identifying places they could take their children.

DRIVING INFLUENCES FOR COVID-19 DATA COLLECTION

Some interview participants commented on how patient and public need influenced their objectives. For instance, participants remarked that they wanted to be among the first providers of a vaccine and, therefore, needed to understand or mitigate vaccine hesitancy. One interview participant identified peer influence as the initial driving factor to start COVID-19 data collection and development of a dashboard-if a large hospital started a dashboard, the assumption for him was that the outbreak had the potential to be serious. And from that perspective, there would be a need to understand how to prepare for an outbreak, and the dashboard could be utilized to inform that process.

HOW DATA COLLECTION OBJECTIVES EVOLVED

Interviewees were asked questions to understand how data collection efforts shifted or evolved over the months. Interestingly, many interviewees noted that their main objective remained unchanged since the beginning of the pandemic, and it was more that they made small adjustments and iterated on their data collection efforts as more data came in about the virus and the broader community's response. For instance, a group conducting surveys noted that they added questions to gather information on the racial disparities around COVID-19 impact.

The most prominent shifts in data collection were a direct result of state and federal mandates for COVID-19 data. A few ways in which data requests evolved included a departure from solely reporting the percentage of positive tests to now also requiring negatives as well as comparing asymptomatic and symptomatic positivity rates. This was an important development, as up until that time, data from hospitals and labs were only based on individuals testing positive, meaning when a patient tested negative, they would no longer be a part of hospital-based reporting. Further, state mandates in the summer added "order-based" questions to reporting, which included indicating race and ethnicity and whether patients were symptomatic or pregnant. Others noted a shift in requirements for patient types and counts; i.e., a shift from overall inpatient counts to COVID-19-related deaths. As a result of these changes, some interview participants mentioned the need to retrospectively look at data not initially reported in order to understand time trends and control needs.

USES OF COVID-19-RELATED DATA

DASHBOARDS

The most common use of COVID-19-related data, mentioned by 14 interview participants, was the creation of "dashboards." Some interview participants were tasked by NCDHHS to help predict cases to keep the public and other health professionals informed. Others took it upon themselves to inform the public and improve on the data that was already available.

While no previous protocol for data collection of this type existed, interview participants mentioned existing processes which could be adapted and applied to the COVID-19 pandemic's data needs. One interviewee mentioned that the creation of an operational dashboard was facilitated through the pre-established practice of capacity tracking for isolation rooms, negative pressure rooms, and ventilators through their EPIC EHRs. At that time, most external reporting went through NCDHHS or other means to reach the CDC. Other dashboards utilized standardized weekly reporting to keep regional organizations informed on current state resources and utilization.

OPERATIONAL DECISION SUPPORT

Five interview participants described establishing "command centers" to help guide strategic planning; in these cases, COVID-19 data was used internally in an operational manner to provide decision support for clinical and administrative executives by monitoring positive and negative cases, test volumes, admissions, and age and other demographics through detailed inpatient data. Further, participants noted that many hospitals utilized data to predict volumes and develop plans to convert or add hospital space to accommodate COVID-19 patients if needed.

MODELING

Throughout the evolution of COVID-19-related data requests, what remained constant was a need for modeling to project the future number of cases and impact on the state's healthcare system. However, groups doing modeling reported that the model components and the parameters used to model future outcomes have evolved substantially, as assumptions made have been continually updated as more has been learned about COVID-19. Early models were basic and used case counts, though these quickly pivoted to more complicated models incorporating parameters surrounding transmission and disease progression. While NCDHHS primarily uses time trend modeling for such things as predicting peak surge capacity and informing resource allocation, it is partnering with other experts for more predictive modeling.

Data types and uses of models reported by interviewees include:

- Patient census models to predict caseloads
- Epidemiological models enhanced with Bayesian methods to inform facility bed demand
- The Apache model to predict mortality and guide treatment protocols
- Transmission models to calibrate various scenarios (e.g., lockdown, reopening, etc.) with respect to transmission rates
- Normalized data models that can be applied across different facilities, such as using EHR data across different hospitals to make operational decisions

DISSEMINATION OF COVID-19-RELATED DATA

STATE AND FEDERAL REPORTING

COVID-19 data was requested and required to be reported at the state and federal levels from NCDHHS. One example of COVID-19 data dissemination includes informative sessions for governors and U.S. Senate hearings. One participant reported that they work to integrate COVID-19 data and policy information to provide policymakers with non-technical explanations of how legislation could impact various scenarios. Others noted that assessing current public reaction to COVID-19 informs future planning for other novel pandemics.

INFORMATIVE DASHBOARDS

As previously mentioned, informative dashboards were a common use for COVID-19 data. Interviewees reported using dashboards internally within organizations, externally outside of and across organizations, and to inform the public. Dashboards incorporated data from EHRs, the web, and other public data sources.

COMMUNITY OUTREACH

The importance of transparency and community education was an important theme that arose among interview participants. Webinars and virtual engagements, publications, and PSAs were some of the methods interview participants used to disseminate COVID-19-related information. Local community leaders, county school systems, local governments, journalists, and underserved populations were among groups targeted by interview participants. One interviewee noted that her group is very cognizant of information overload contributing in part to 'COVID fatigue' of the general public. In response, they are very intentional when considering what information to release, and they attempt to tie information to state or local regions so people can better relate.

COVID-19 DATA COLLECTION CHALLENGES

DATA DEFINITIONS AND CONSISTENCY

Interview participants made clear their frustration with a lack of clear and consistent definitions across organizations, even up to the federal level. For example, there are several ways organizations can define "capacity," and there are different methods for calculating positivity rates. There existed significant variation in interpreting data within the state, resulting in skepticism surrounding the data quality. It would be advisable to convene a group to define key concepts that proved useful soon after the current pandemic is under control in order to prepare for future pandemics; in other words, we need to take advantage of the "lessons learned" from this pandemic.

COLLECTION PROCESS

Participants expressed several frustrations with the COVID-19 data collection process. Many issues were related to the amount of time needed to complete data collection, including:

- Access to data and delays caused by waiting for DUAs.
- Continually evolving data requirements. Each new request requires effort to determine what aspects of existing systems need to be changed or updated.
- Lack of clear authority and defined roles for who to contact for approval of data sharing or to have questions resolved in a timely manner. This impacts the ability to meet regulatory requirements.
- Disconnects between HHS requests (some of which originated external to HHS) and the amount of time required to adjust established processes to comply with new or modified data requests.
- Staff fatigue and burnout resulting from the number of data requests, changes in data requests, and the urgent nature of these requests.

All of these issues are especially problematic for those working at smaller labs, hospitals, and facilities operating with limited staff and resources.

MODELING

As previously mentioned, data lags are a common issue in research and have impacted COVID-19 models, which often require more data to be more accurate. The need for data use agreements (DUAs) has led to frustration among modelers, with one group reporting that if data had been more available to them in the first 90 days or less of building the model, it could have been built faster and more precisely. Others claim to have a better understanding of what information can be requested and shared than they did in March, and they now request data that does not require a DUA. One key informant remarked that the type of modeling his group has been doing typically takes years, and doing so amidst a pandemic where information needs are urgent and parameters are constantly changing is a significant added stressor.

CONTACT TRACING

Interview participants mentioned major obstacles in conducting contact tracing. Overall, there has been, over time, an increase in the number of cases considered “lost to follow up” because people are unwilling to cooperate with public health officials. When people are located as part of contact tracing efforts, they can be reluctant to name who they were in contact with during two weeks before symptom onset because those contacts will be required to quarantine, and this results in a decreasing number of named close contacts among traced individuals. Not only are people less willing to cooperate once contacted, but one participant noted that it has become increasingly difficult to reach people at all for contact tracing efforts.

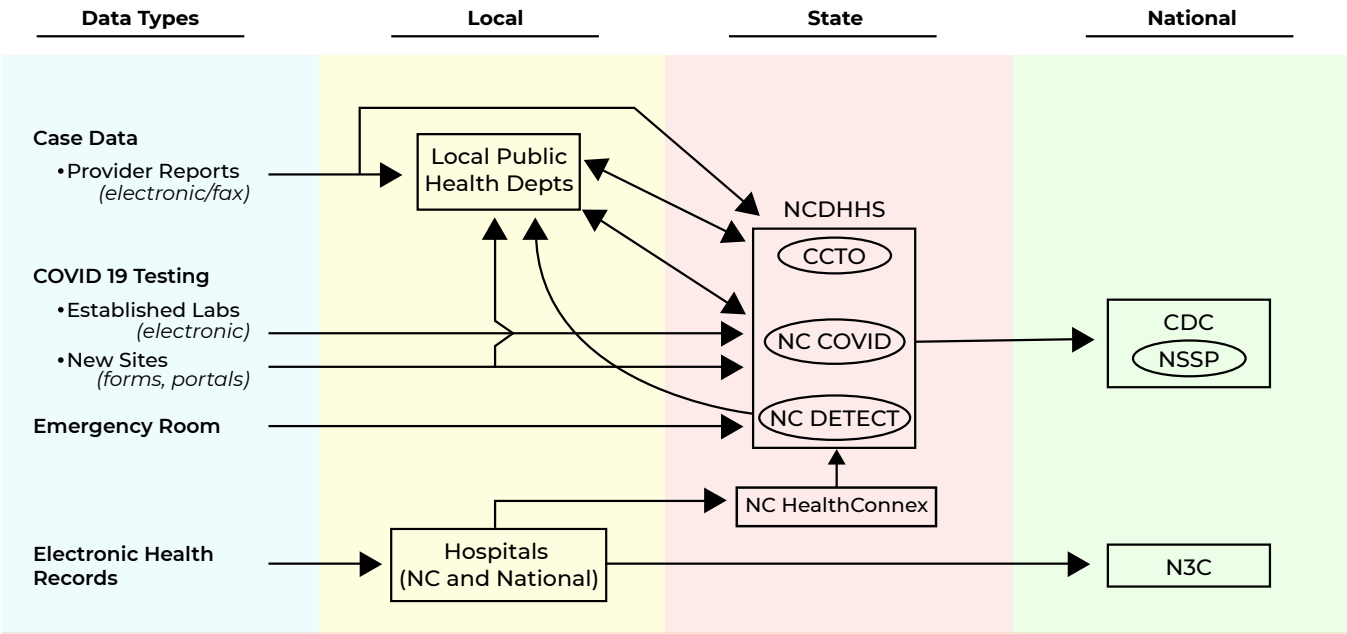
Universities and organizations, mostly health care facilities, are also engaged in contact tracing outside of local health departments. These organizations have comprehensive COVID-19 contact tracing plans with trained staff. Certain organizations carry out contact tracing for employees only. These organizations expressed difficulties in contact tracing outside of their respective organizations, most likely attributable to having no direct metrics to collect COVID-19 regulation compliance among general patient populations.

PATIENT-LEVEL COMORBIDITY & DEMOGRAPHIC DATA

Information regarding comorbidities of patients with COVID-19 typically does not get reported outside of the health systems within which it is collected due to the high degree of personally identifiable information (PII) and protected health information (PHI) attached to the comorbidity data. As this information lives within EHR systems, it is more frequently utilized in separate, internal system analyses for describing patient populations.

COVID-19 DATA FLOW

The project team created the following figure describing the data flow of COVID-19 data across NC based on the interviews.



- **NCDHHS**: North Carolina Department of Health and Human Services
- **CCTO**: COVID-19 Community Team Outreach Tool
- **NC COVID**: North Carolina COVID-19 Surveillance System
- **NC DETECT**: North Carolina Disease Event Tracking and Epidemiologic Collection Tool
- **CDC**: Centers for Disease Control and Prevention
- **NSSP**: National Syndromic Surveillance Program
- **N3C**: National COVID Cohort Collaborative

Figure 1. COVID-19 Data Flow in NC

In NC, the sources of COVID-19 data include case data, COVID-19 testing data, emergency room data, and data from electronic health records (EHRs).

Case data, or data from COVID-19 case investigations, are in the form of medical provider reports, sent both electronically and via fax to local health departments and NCDHHS. The local health departments then have a two-way flow of case data with NCDHHS' **COVID-19 Community Team Outreach (CCTO) Tool** for tracing efforts, and NCDHHS' **COVID-19 Surveillance System (NC COVID)**.

COVID-19 testing data are gathered from established labs electronically and from new testing sites via forms and newly developed portals; the data are sent to local health departments and NC COVID.

Emergency room data are sent directly to the **North Carolina Disease Event Tracking and Epidemiologic Collection Tool, NC DETECT**, NCDHHS' long-established platform for intaking emergency room data. NC DETECT then communicates the emergency room data to local health departments.

COVID-19 data from EHRs reside in hospitals, both locally and nationally, and the NC data are sent from local hospitals to the **NC HealthConnex** platform, NC's Health Information Exchange (HIE) system, and the **National COVID Cohort Collaborative (N3C)**; N3C is a partnership between the National Center for Advancing Translational Sciences (NCATS) and the National Center for Data to Health (CD2H) to create a centralized national data resource, the NCATS N3C Data Enclave, which allows rapid collection and analysis of COVID-19 data (National Center for Advancing Translational Sciences [NCATS], 2021). NC HealthConnex also communicates this information back to NCDHHS.

All of the COVID-19 data received by NCDHHS then gets communicated at the federal level to the **Centers for Disease Control and Prevention (CDC)** and the **National Syndromic Surveillance Program (NSSP)**.

COMPARING NORTH CAROLINA'S REPORTING METRICS

Given the relentless spread and rapid transmission of COVID-19, it is of critical importance that states provide as much up-to-date information to their residents as possible. In order to fulfill this need, many states have created a COVID-19 Dashboard for the general population to reference. Compared to neighboring states (Virginia, Tennessee, Georgia, and South Carolina), North Carolina is reporting most of the critical metrics to assess our ability to 'flatten the curve.' Figure 2 below displays the comparison of the metrics reported on the COVID-19 Dashboards of NC and neighboring states as well as Johns Hopkins and the NY Times.

COVID-19 Dashboards									
Reporting Organization	Number of Daily Cases	Total Number of Completed Tests	Total Number of Cases	Total Number of Deaths	Cases by County	Deaths by County	Current Hospitilzations	PPE Availability	Hospital Beds (Emptied/ Filled)
North Carolina	☑	☑	☑	☐	☑	☑	☑	☑	☑
Virginia	☑	☐	☑	☑	☑	☑	☑	☑	☑
Tennessee	☑	☑	☑	☑	☑	☑	☐	☐	☑
Georgia	☑	☑	☑	☑	☑	☑	☑	☐	☑
South Carolina	☑	☑	☑	☑	☑	☐	☐	☐	☑
Johns Hopkins	☑	☐	☑	☑	☑	☑	☐	☐	☐
NY Times	☑	☐	☑	☑	☑	☑	☐	☐	☐

Figure 2. COVID-19 Dashboards from NC and Surrounding States

On the NC COVID-19 Dashboard homepage (NCDHHS, 2021b), key information is relayed in an easy-to-process layout. This information includes the cumulative and daily number of cases reported, the number of tests conducted each day, the number of people currently hospitalized due to the virus, and the number of positive molecular (PCR) tests per 100,000 residents in each county of the state listed in a table that is easy to navigate. However, in order to see the total number of deaths across the entire state, the viewer must select the "Cases" tab displayed on a side menu. This is one major area of reporting where North Carolina differs from the surrounding states, as all neighboring states have the number of deaths across the state reported on the Dashboard homepage. Also, while North Carolina has a county alert system reporting which counties have a critical level of community spread, the associated graphic seems to only be updated on a monthly basis, not in real-time as the number of new daily cases are reported. In light of new reported cases, real-time updates to this

graphic would be helpful to viewers who may be seeking to understand the current status of their county compared to neighboring counties.

There are some additional discrepancies on the NCDHHS dashboard that may hinder quick access to important information when compared to neighboring states. For example, the state of Georgia has a more interactive and easy-to-understand county-by-county map that utilizes a broader color spectrum to indicate the number of cases reported per 100,000 people in the previous 14 days (Georgia Department of Public Health [DPH], 2021). Georgia also has a graph labeled "COVID-19 Over Time" that the viewer can zoom in and out of to see the specific number of cases and deaths reported dating back to March 2020. This is more user friendly than the graph utilized by NCDHHS, as the viewer cannot zoom into specific dates, making it difficult to see the exact number of cases and deaths reported on a single day.

Virginia is unique among the neighboring states in that it notes the number of multisystem inflammatory syndrome in children (MIS-C) cases reported across the state on their dashboard (Virginia Department of Health [VDH], 2021). South Carolina is unique in reporting the estimated percentage of people who have recovered from COVID-19 based on reporting from each hospital (South Carolina Department of Health and Environmental Control [DHEC], 2021). Tennessee is unique in that it has a quick link on the dashboard homepage to a weekly situation summary (Tennessee Department of Health [TDH], 2021a); this summary is a bulleted list of trends, the state's response, and health reminders to state residents. In addition, Tennessee also has a quick link to the percentage of positive tests among long-term care facilities county-by-county across the state (TDH, 2021b). These are all data sources and information displays that could be applied to the NCDHHS Dashboard website and would aid in disseminating information to NC residents in a more complete and interactive way.

One area of reporting where North Carolina clearly surpasses neighboring states' dashboards is noting the available supply of Personal Protective Equipment (PPE) for hospitals. The average daily requests for PPE such as face shields, gloves, gowns, N95 respirators, and regular medical masks, along with the "estimated days of supplies on hand," the number of orders placed for the previously stated PPE, and how many of the ordered items have been received are all displayed on the dashboard (NCDHHS, 2021d). No other neighboring state, with the exception of Virginia's reporting on general PPE shortages, currently reports the type of PPE available and the PPE orders among hospitals on their respective dashboards.

As mentioned, this is a constantly evolving situation, and new data and data types are regularly incorporated into dashboards. In the coming months, new data types will become available on NC's dashboard, including home testing data and vaccination data. In terms of vaccinations, NC's dashboard is ahead of many states in that it displays race and ethnicity breakdowns of vaccine recipients by county (NCDHHS, 2021e), and this information is critical for making operational decisions to ensure equitable distribution of vaccines among different demographics (Fadulu & Keating, 2021).

NORTH CAROLINA REPORTING COMPARED NATIONALLY

Figures 3 and 4 below illustrate a comparison between national reporting of North Carolina's case and death counts with the numbers that NCDHHS reports. To conduct this analysis, two reputable national reporting organizations were selected, The New York Times and Johns Hopkins University, to compare with the reporting conducted by NCDHHS. Each organization has an active COVID-19 dashboard for every state. The 7-day average was taken on the 2nd full week of each month. The fluctuation in the reported numbers may be due in part to the reporting times and possible delays in gathering this information from NC counties and other reporting and testing agencies.

Figure 3. New COVID-19 Cases Reported by Organization

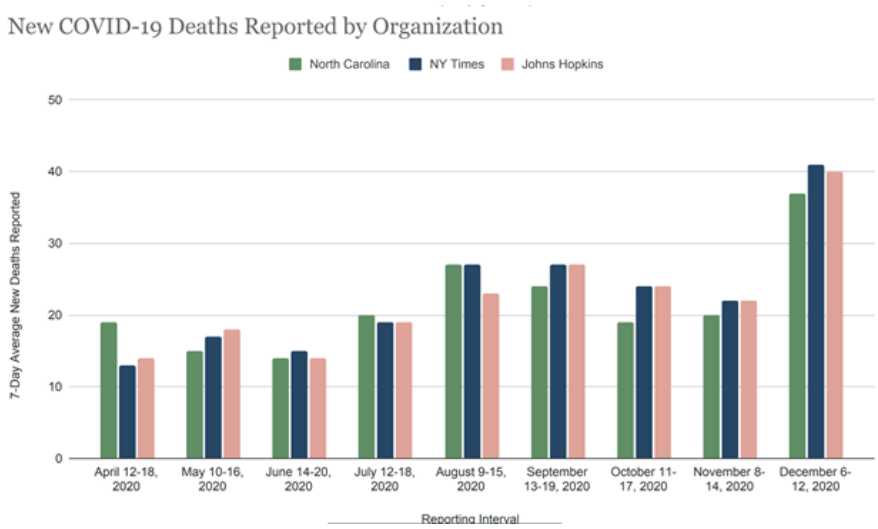


Figure 4. New COVID-19 Deaths Reported by Organization

NCDHHS collects their reporting data from various sources, including the North Carolina COVID-19 Surveillance System (NC COVID), which is a part of the CDC effort to link "web-based health surveillance and reporting systems" (NCDHHS, 2021a). The County Emergency Operation Centers (EOCs) is also a crucial provider of reporting data, along with the U.S. Outpatient Influenza-like Illness Surveillance Network, which is composed of state health departments and healthcare providers across the U.S. Appriss, a data and analytics company, is utilized to note the current ICU capacity of hospitals across the state.

The New York Times notes that their "data is the product of dozens of journalists working across several time zones to monitor news conferences, analyze data releases and seek clarification from public officials on how they categorize cases" (NYT, 2021). Johns Hopkins has created their own data repository that collects information not only from state health departments, but from individual counties that have the infrastructure required to report their daily number of cases and deaths (JHU, 2021). In light of these multiple sources of data collection and reporting, the largest discrepancy seemed to occur during the summer surge in July through September, where reporting from NCDHHS, the New York Times, and Johns Hopkins varied by as many as 77 new reported cases. However, the variations in case counts are not significant enough to warrant concern in reporting discrepancies.

Johns Hopkins, in addition to reporting the 7-day average of cases and deaths, reports granular county-level information on the total number of ICU beds, the number of people age 65 and older, the number of people living in poverty, and even the types of insurance the population of the county is estimated to be enrolled in (private, medicaid, direct purchase, etc.) (JHU, 2021). This information is gleaned from the American Community Survey (Esri Demographics, 2021) and is by far the most detailed county-by-county information available by any publicly reporting agency.

It is encouraging to note the overall consistency of the number of new cases and deaths reported; however, the lag in reporting data and the different times of the day when each agency chooses to post their information seem to contribute to minor but regular discrepancies in numbers.

DISCUSSION

Based on the above results, we have compiled the below list of recommendations to improve COVID-19 data-related processes and intercommunication in NC.

1. **Future pandemic response should be centralized through the North Carolina Department of Health and Human Services.**

   Standardized and coordinated information sharing is the foundation of effective pandemic response. Key informants expressed their appreciation for the leadership exemplified by NCDHHS following the COVID-19 outbreak and a desire for an even more streamlined process when preparing for and responding to future pandemics. Key informants expressed frustration over requirements imposed by the federal government that were made without appropriate guidance and with very short timelines for compliance. Even in such cases, the key informants emphatically asserted that the leadership and coordination provided by NCDHHS helped alleviate these difficult circumstances.

2. **The state should work to support and sustain the cross-sector collaborative networks established during the COVID-19 outbreak.**

   A consistent theme mentioned by key informants is that cross-sector collaboration was initially a major facilitator in the collection and use of COVID-19-related data. NC was particularly well-positioned to collect data regarding the pandemic because of prior collaborations, for example, NC DETECT. Thus, many of the new collaborations developed from prior and existing relationships and a desire to maximize the combined impact of the work being done by colleagues working at different institutions. NC is fortunate to have a number of strong research institutes and will benefit from formalizing many of the collaborative networks that have organically developed since March and encourage even more data synergy and consistency in data collection processes moving forward.

3. **Pandemic-related data should be publicly accessible and available in a format that is easy to download and utilize.**

   Proactive data collection and timely analysis facilitates identification of patterns and the dissemination of timely information. To increase access to such data, NC should publicly release data in an easy to download format to not only inform the public but also to facilitate analysis by the broad community of data scientists. Further, data should be translated in a manner useful to both data scientists and the greater public.

4. **Contact tracing and surveillance methods need to be more robust and systematic.**

   Early detection of cases remains the most important factor in effective management of an infectious outbreak. This effort becomes more challenging for communities that are more remote and/or economically disadvantaged. Development of a standardized protocol that supports extensive testing and reporting will empower communities by providing them with guidance on best practices and appropriate response in the event of an infectious outbreak. Starting a process now to develop and strengthen community-level or city/town-level surveillance methods should function to limit the impact of future infectious outbreaks by increasing the chances for early detection. This, in turn, should minimize economic disruptions at the source of the problem. Robust and practical guidelines for clinical and administrative workforces will need to be created and standardized at the state level, and then effectively communicated and funded.

In addition, there are several more general observations and lessons learned that arose from the analysis:

CROSS-SECTOR COLLABORATION

A positive byproduct of the COVID-19 pandemic has been the ability and demand for cross-sector collaboration. Cross-sector collaboration was identified as something that North Carolina did very well. Collaborative efforts were mentioned by every interview participant. Some of the groups involved in these collaborations included school systems, government organizations, health systems, pharmaceutical and medical supply companies, think tanks and consulting firms, non-profit institutions, researchers and educators, health professionals, and foundations. The collaborations were effective in proactively establishing mechanisms to receive state and federal data, facilitating data centralization, and synergizing modeling efforts. On the other hand, the fast-paced and always evolving environment created by COVID-19 was at times difficult to navigate among collaborators. In addition, some interviewees reported there were lost opportunities for collaboration, such as when a lack of awareness of work being done by others resulted in duplicative efforts.

REDUCTION IN REGULATORY BARRIERS

Due to the risk involved with in-person visits, regulatory agencies made temporary exceptions for telehealth services. This allowed individuals to receive necessary health services and avoid the risk of exposure to COVID-19. Insurance payers should evaluate the services provided over the year and consider whether it is prudent to continue coverage of telehealth services in the future.

TECHNOLOGY INTEGRATION

Technology plays a critical role in effective data collection and reporting. Several organizations noted successes in terms of software or system integrations between the state health department and electronic labs reporting interfaces.

IT systems and services were forced to improve and/or stabilize their products as a byproduct of their data collection and reporting efforts. Further, NCDHHS responded quickly to develop and deploy electronic methods for providers and laboratories to upload data. System integrations may play a pivotal role in future reporting successes, and the public health infrastructure would benefit from additional funding for data-related health IT projects at the state and federal levels.

STRENGTHS & WEAKNESSES ASSESSMENT

At the start of any major initiative, a needs assessment or gap analysis can inform the development of the project's activities. Throughout this study, interviewees identified several strengths but also gaps that, if filled, could serve to advance progress in data collection. For example, early on, communication between two state-level organizations was identified as a gap. Once identified, both entities sought to intentionally improve communication practices to promote progress and to come together for a common agenda--reporting of lab data. One interviewee identified a strength in terms of partner organizations' ability to secure and gather data and then disseminate or report on discrete data elements. Another expressed that supply chain and manufacturing information might best be managed at the federal level, noting NCDHHS and Centers for Medicare and Medicaid Services (CMS) may be better able to manage these data effectively, as opposed to ad hoc university health systems or researchers.

Several notable weaknesses were also mentioned. First, testing responsibility and contact tracing was sometimes too burdensome to be managed by local health departments alone. For testing, manufacturing responsibility could be scaled and shared with the private sector. Additionally, while cross-sector collaboration is noted above as a strength, a gap remains in terms of disaster preparedness. Partners were able to align efforts quickly, but many challenges may have been avoided if there was a plan in place prior.

STUDY LIMITATIONS

We note several limitations in this study. The main limitation is that qualitative research does not provide generalizability nor statistical representation of larger populations. While we have obtained and summarized common themes among interview participants, these themes cannot be generalized to the larger population of North Carolina. The information in this report is descriptive and meant to provide insight to the experiences and opinions of stakeholders represented by the sample population. The research team made several efforts to obtain a geographically diverse sample of interviews, but the representation of rural health efforts is limited in this study. Another limitation is the variability between interviews. All interviews were scheduled for 30 minutes, but it proved difficult for the interviewers to get through all the questions in this timeframe depending on the specific circumstances of each interview. Some interviews did run over time, but in multiple cases, the interviewer was not able to cover all the questions, especially in early interviews.

Relatedly, the evolution of the interview guide during the study was another limitation; as noted, the interview guide proved too lengthy in the initial interviews, so it was condensed into the most essential questions to be asked of each key informant. Thus, different questions were asked of interviewees scheduled before and after this change. This was taken into consideration during the analysis, and we attempted to standardize between interviews to limit the effects of this change, but it is a significant limitation, nonetheless.

Further, due to the ever-changing nature of the pandemic, there was a sense of urgency in getting this information out as quickly as possible to inform data collection efforts. Therefore, the team experienced time constraints in the development of the study. Additionally, interviewees experienced significant challenges with competing work priorities, and several were not available to be interviewed at the time of the study. The response rate for interview requests was 59.7%.

Though this study is not without its limitations, it has undeniably produced and compiled a unique, highly valuable collection of rich information about COVID-19 data from some of the utmost experts and stakeholders in NC. As we have laid out, a great deal has been learned from the analysis of these key informant interviews, and we hope that this work can be utilized to make better informed decisions around COVID-19 data collection, reporting, dissemination, and intercommunication among stakeholders. Most importantly, we hope our work will directly benefit the many lives impacted by the virus in our state and beyond.

REFERENCES

Centers for Disease Control and Prevention (CDC). (2021). *CDC COVID Data Tracker*. https://covid.cdc.gov/covid-data-tracker/#cases_casesper100klast7days.

Centers for Disease Control and Prevention (CDC). (2004). *Public health guidance for community level preparedness and response to severe acute respiratory syndrome (SARS)*. Washington: Department of Health and Human Services (US).

Consolidated Appropriations Act, 2021, U.S.C. (2021).

Coronavirus Aid, Relief, and Economic Security Act (CARES Act), U.S.C. (2020).

Coronavirus Preparedness and Response Supplemental Appropriations Act, 2020, U.S.C. (2020).

Esri Demographics. (2021). American Community Survey (ACS). Arcgis.com. https://doc.arcgis.com/en/esri-demographics/data/acs.htm.

Fadulu, L. and Keating, D. (2021). Few states are accurately tracking coronavirus vaccinations by race. Some aren't at all. https://www.washingtonpost.com/local/social-issues/race-data-vaccine-states/2021/01/27/fc9df6de-6001-11eb-afbe-9a11a127d146_story.html

Families First Coronavirus Response Act, U.S.C. (2020).

Georgia Department of Public Health (DPH). (2021). *Georgia Department of Public Health Daily Status Report*. https://dph.georgia.gov/covid-19-daily-status-report.

Johns Hopkins University (JHU). (2021). *COVID-19 Dashboard by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University (JHU)*. Arcgis.com. https://www.arcgis.com/apps/opsdashboard/index.html#/bda7594740fd40299423467b48e9ecf6.

Katz, J., Sanger-Katz, M., & Quealy, K. (2020, July 29). *A Detailed Map of Who Is Wearing Masks in the U.S.* The New York Times. https://www.nytimes.com/interactive/2020/07/17/upshot/coronavirus-face-mask-map.html

Majumder, M.S., & Rose, S. (2020, October 6). *Health Care Claims Data May Be Useful For COVID-19 Research Despite Significant Limitations*. Healthaffairs.Org. https://www.healthaffairs.org/do/10.1377/hblog20201001.977332/full/.

National Center for Advancing Translational Sciences (NCATS). (2021). *About the National COVID Cohort Collaborative*. https://ncats.nih.gov/n3c/about.

NC Exec. Order No. 116. (2020).

NC Exec. Order No. 117. (2020).

NC Exec. Order No. 118. (2020).

NC Exec. Order No. 121. (2020).

NC Exec. Order No. 138. (2020).

NC Exec. Order No. 181. (2020).

NC Exec. Order No. 188. (2021).

The News & Observer (N&O). (2021). *Coronavirus*. https://www.newsobserver.com/news/coronavirus.

The New York Times (NYT). (2021). *Coronavirus (Covid-19) Data in the United States.* Retrieved from https://github.com/nytimes/covid-19-data.

North Carolina Department of Health and Human Services (NCDHHS). (2021a). *About the Data*. https://covid19.ncdhhs.gov/dashboard/about-data.

North Carolina Department of Health and Human Services (NCDHHS). (2021b). *COVID-19 North Carolina Dashboard*. https://covid19.ncdhhs.gov/dashboard.

North Carolina Department of Health and Human Services (NCDHHS). (2021c). *North Carolina Electronic Disease Surveillance System (EDSS)*. Ncedss.Com. https://ncedss.ncpublichealth.com/login.do.

North Carolina Department of Health and Human Services (NCDHHS). (2021d). *PPE*. https://covid19.ncdhhs.gov/dashboard/ppe.

North Carolina Department of Health and Human Services (NCDHHS). (2021e). Vaccinations. https://covid19.ncdhhs.gov/dashboard/vaccinations.

North Carolina Department of Health and Human Services (NCDHHS) & UNC School of Medicine. (2021). *NC DETECT*. Ncdetect.Org. https://ncdetect.org/.

North Carolina General Assembly. *COVID-19 Information and Resources*. https://sites.ncleg.gov/library/north-carolina-covid-19-information-and-resources/.

PolicyMap. (2021). *PolicyMap*. Policymap.com. https://www.policymap.com/.

SafeGraph. (2021). *SafeGraph*. Safegraph.Com. https://www.safegraph.com/

South Carolina Department of Health and Environmental Control (DHEC). (2021). *South Carolina County-Level Data for COVID-19*. https://scdhec.gov/covid19/south-carolina-county-level-data-covid-19.

Tennessee Department of Health (TDH). (2021a). *Epidemiology and Surveillance Data*. https://www.tn.gov/health/cedep/ncov/data.html.

Tennessee Department of Health (TDH). (2021b). *Long Term Care Facility Data*. https://www.tn.gov/health/cedep/ncov/data/clusters-in-long-term-care-facilities.html. U.S.

Government Publishing Office. *Coronavirus (COVID-19)*. https://www.govinfo.gov/features/coronavirus.

Virginia Department of Health (VDH). (2021). *COVID-19 in Virginia*. https://www.vdh.virginia.gov/coronavirus/covid-19-in-virginia/.

WHO Media Centre. (2003). *WHO issues a global alert about cases of atypical pneumonia: Cases Of Severe Respiratory Illness May Spread To Hospital Staff* [News releases]. https://www.who.int/mediacentre/news/releases/2003/pr22/en/.

World Health Organization (WHO). (2020a). *WHO Director-General's statement on IHR Emergency Committee on Novel Coronavirus (2019-nCoV)*. https://www.who.int/dg/speeches/detail/who-director-general-s-statement-on-ihr-emergency-committee-on-novel-coronavirus-(2019-ncov).

World Health Organization (WHO). (2020b). *WHO Director-General's opening remarks at the media briefing on COVID-19 - 11 March 2020*. https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020.

WRAL. (2021). *NC Coronavirus Updates and Latest News*. WRAL.Com. https://www.wral.com/coronavirus/.

APPENDICES

APPENDIX A. Timeline of State & Federal COVID-19 Legislation

Timeline of State and Federal COVID-19 Legislation		
Date	**Federal**	**North Carolina**
March 4, 2020	Coronavirus Preparedness and Response Supplemental Appropriations Act, 2020	
March 10, 2020		Executive Order No. 116
March 14, 2020		Executive Order No. 117
March 17, 2020		Executive Order No. 118
March 18, 2020	Families First Coronavirus Response Act	
March 20, 2020		Executive Order No. 119
March 23, 2020		Executive Order No. 120
March 27, 2020	CARES Act (Coronavirus Aid, Relief, and Economic Security Act)	Executive Order No. 121
March 30, 2020		Executive Order No. 122
March 31, 2020		Executive Order No. 124
April 1, 2020		Executive Order No. 125
April 7, 2020		Executive Order No. 129
April 8, 2020		Executive Order No. 130
April 9, 2020		Executive Order No. 131
April 17, 2020		Executive Order No. 133
April 20, 2020		Executive Order No. 134

April 23, 2020		Executive Order No. 135
April 24, 2020	Paycheck Protection Program and Health Care Enhancement Act	
April 28, 2020	Student Veteran Coronavirus Response Act of 2020	
May 5, 2020		Executive Order No. 138
May 12, 2020		Executive Order No. 139
May 20, 2020		Executive Order No. 141
May 30, 2020		Executive Order No. 142
June 4, 2020		Executive Order No. 143
June 5, 2020	Paycheck Protection Program Flexibility Act of 2020	Executive Order No. 144
June 24, 2020		Executive Order No. 147
July 4, 2020	An act to extend the authority for commitments for the paycheck protection program and separate amounts authorized for other loans under section 7(a) of the Small Business Act, and for other purposes	
July 13, 2020	Emergency Aid for Returning Americans Affected by Coronavirus Act	
July 14, 2020		Executive Order No. 150
July 16, 2020		Executive Order No. 151
July 24, 2020		Executive Order No. 152
July 28, 2020		Executive Order No. 153
August 3, 2020	Protecting Nonprofits from Catastrophic Cash Flow Strain Act of 2020	

August 5, 2020		Executive Order No. 155
August 14, 2020	Safeguarding America's First Responders Act of 2020	
August 31, 2020		Executive Order No. 162
September 1, 2020		Executive Order No. 163
September 21, 2020		Executive Order No. 165
September 30, 2020		Executive Order No. 169
October 21, 2020		Executive Order No. 170
October 28, 2020		Executive Order No. 171
November 10, 2020		Executive Order No. 176
November 23, 2020		Executive Order No. 180
December 8, 2020		Executive Order No. 181
December 21, 2020	Consolidated Appropriations Act, 2021	Executive Order No. 183
December 30, 2020		Executive Order No. 184
January 6, 2021		Executive Order No. 188
January 27, 2021		Executive Order 189; Executive Order 190

References:

- U.S. Government Publishing Office. Coronavirus (COVID-19). https://www.govinfo.gov/features/coronavirus.
- North Carolina General Assembly. COVID-19 Information and Resources. https://sites.ncleg.gov/library/north-carolina-covid-19-information-and-resources/.

APPENDIX B. Interview Request

Template Interview Request: Email Body

Dear [insert name],

I am writing to request a Zoom interview with you at your earliest convenience. I am part of the "COVID-19 Data Travels" project, led by the Renaissance Computing Institute (RENCI) at UNC-Chapel Hill and funded by the NC Policy Collaboratory. My research team aims to understand how COVID-19 data are being collected and reported across the state; thus, we are seeking the help of leading public health and medical professionals at universities and health organizations

across the state to help with this endeavor. Given your expertise/position [pick one] in [insert area of expertise/position here], we identified you as an invaluable source to provide insight into this complex and constantly evolving situation.

Please let me know some dates and times you are available for a 30-minute interview to help RENCI collect information on COVID-19 data. We have included the list of questions we generally ask below for your review. Additionally, we have attached a more in-depth overview of the project and the interview.

Thank you in advance for considering our request. Please let us know if you think of other people we should reach out to, as well.

Most importantly, we hope that you, your family, and your colleagues are staying safe and healthy in this challenging time.

Best regards,

[insert name], on behalf of project team

Interview Questions:

1. When did you begin collecting COVID-related data?
2. What were your objectives when you started collecting data?
3. Has the objective evolved? In what ways?
4. What guidance, if any, have you received from other organizations?
5. What were the biggest barriers in your work?
6. What type of patient-level/individual data is your organization collecting?
7. What challenges have you experienced in collecting individual-level data?
8. How does your organization collect data on patient contact/contact tracing?
9. How are hospital capacities being reported?
10. How are hospital utilizations being reported?
11. How is comorbidity being addressed?
12. How are the results of data collection being reported up to the NCDHHS (Department of Health and Human Services)?
13. How are COVID-19 diagnoses and outcomes being centralized?
14. What is the purpose of data models you use?
15. Is there data that you need, but don't have, for your models to be more accurate?
16. How are decisions made by your organization regarding data accessibility and dissemination?
17. What are some ways in which data dissemination has informed on or positively impacted

THE UNIVERSITY of NORTH CAROLINA at CHAPEL HILL

A Case Study of Data Travels

The Renaissance Computing Institute (RENCI) at UNC-Chapel Hill received funding from the NC Policy Collaboratory (a center at UNC that focuses on policy issues for the state of NC) through the award, "A Case Study of Data Travels." As we face the COVID-19 pandemic caused by the novel coronavirus, medical professionals, technologists, community leaders, and policy makers need to leverage data to grapple with the current crisis while also using data to build predictive models for the coming months and years. We are all collecting data, at various levels, from the individual to communities and organizations across our state. As we collect and analyze data from various sources in the fight against this pandemic, the average North Carolinian is left wondering how to make sense of the role of data in dealing with the COVID-19 pandemic. Thus, we are seeking the help of leading public health and medical professionals at universities and health organizations across the state of NC to develop a report which will answer many of the outstanding questions regarding the collection and use of COVID-19 data.

We plan to record our interviews and transcribe the audio recording; however, the content of the interview will be de-identified. Any information used in the report will not cite an interviewee by name unless you indicate that you would prefer that your name is used. Please note that the recording will be deleted after transcription. If you have any concerns, please do not hesitate to let us know.

We have included the list of interview questions we prepared below.

Interview Questions:

1. When did you begin collecting COVID-related data?
2. What were your objectives when you started collecting data?
3. Has the objective evolved? In what ways?
4. What guidance, if any, have you received from other organizations?
5. What were the biggest barriers in your work?
6. What type of patient-level/individual data is your organization collecting?
7. What challenges have you experienced in collecting individual-level data?
8. How does your organization collect data on patient contact/contact tracing?
9. How are hospital capacities being reported?
10. How are hospital utilizations being reported?
11. How is comorbidity being addressed?
12. How are the results of data collection being reported up to the NC DHHS (Department of Health and Human Services)?
13. How are COVID-19 diagnoses and outcomes being centralized?
14. What is the purpose of data models you use?
15. Is there data that you need, but don't have, for your models to be more accurate?
16. How are decisions made by your organization regarding data accessibility and dissemination?
17. What are some ways in which data dissemination has informed on or positively impacted the state of the pandemic?

APPENDIX C. Interview Guide

Final COVID-19 Data Interview Guide

#	Question
1	When did you begin collecting COVID-related data?
2	What were your objectives when you started collecting data?
3	Has the objective evolved? In what ways?
4	What guidance, if any, have you received from other organizations?
5	What were the biggest barriers in your work?
6	What type of patient-level/individual data is your organization collecting?
7	What challenges have you experienced in collecting individual-level data?
8	How does your organization collect data on patient contact/contact tracing?
9	How are hospital capacities being reported?
10	How are hospital utilizations being reported?
11	How is comorbidity being addressed?
12	How are the results of data collection being reported up to the NCDHHS (Department of Health and Human Services)?
13	How are COVID-19 diagnoses and outcomes being centralized?
14	What is the purpose of data models you use?
15	Is there data that you need, but don't have, for your models to be more accurate?
16	How are decisions made by your organization regarding data accessibility and dissemination?
17	What are some ways in which data dissemination has informed on or positively impacted the state of the pandemic?

Original COVID-19 Interview Guide

Intro. General Questions:

1. When did you begin collecting COVID-related data?
2. What were your objectives when you started collecting data?
3. Has the objective evolved? In what ways?
4. What is your protocol for collecting data?
5. What guidance, if any, have you received from other organizations?
6. When you meet with your colleagues to discuss data collection, what kind of things do you talk about?
7. How, if at all, has this process for collecting the data changed since last March (or when you began your collection)?
8. What were the biggest barriers in your work?
9. What were the greatest facilitators in your work?

Group 1. Patient Data Sources:

1. What type of patient-level/individual data is your organization collecting?
2. How are the data being collected by your organization?
3. How do you track demographics for patients?
4. How is comorbidity data being addressed?
5. What challenges have you experienced in collecting individual-level data?

Group 2. Other Data Types:

1. How did you define, identify and track COVID testing sites?Can you explain what COVID testing sites are?
2. Where are COVID testing sites?
3. What are the materials important for COVID testing?
4. Can you tell me about COVID-19-related treatment materials and resources?
5. How does your organization collect data on patient contact/contact tracing?

Group 3. Data Collection & Reporting:

1. What are the different sources of data that you collect?
2. How are decisions made regarding data collection?
3. What challenges have you had regarding data collection?
4. What, if anything, do you wish you had done differently regarding data collection?
5. How are hospital capacities being reported?
6. How are hospital utilizations being reported?
7. How is comorbidity being addressed?
8. How are the results of data collection being reported up to the NCDHHS (Department of Health and Human Services)?
9. How are COVID-19 diagnoses and outcomes being centralized?

Group 4. Data Modeling:

1. What is the purpose of data models you use?
2. What are the larger plans for data use?
3. How are the data being aggregated?
4. What models are you using for the data you collect?
5. What challenges are you experiencing with using the models?

Group 5. Data Connections

1. How are decisions made by your organization regarding data accessibility and dissemination?
2. How are data transferred from one source to another?
3. How are data transferred beyond local limits? To state organizations?
4. What are some ways in which data dissemination has informed on or positively impacted the state of the pandemic?
5. What advice, if any, would you give others in your role regarding collecting COVID data?

APPENDIX D. Glossary

A Priori Approach**:** Deductive; relating to or derived by reasoning from self-evident propositions

Apache Model**:** Apache is a mortality prediction model used in the ICU, and it gives you a score (Apache Score) of a patient on admission that gives you likelihood of mortality.

Appriss**:** Based in Louisville, KY, Appriss Health provides trusted technology solutions to federal and state governments, payers, health systems, clinicians, pharmacies, and health information exchanges working to improve public health.

Blue Cross Blue Shield**:** The Blue Cross Blue Shield Association is a national association of 36 independent, community-based and locally operated Blue Cross Blue Shield insurance companies.

CARES Act**:** The Coronavirus Aid, Relief, and Economic Security (CARES) Act was passed by Congress and signed into law by President Trump on March 27th, 2020. The CARES Act suspended student loan payments and sent a $1,200 stimulus check to all qualifying citizens, along with other mitigation strategies.

CCTO Platform**:** COVID-19 Community Team Outreach Tool. Developed by NCDHHS to assist with monitoring and follow-up of COVID-19 cases and known close contacts. The tool allows Health Departments to automate workflows, streamline case and known contact communication, and align and integrate with statewide COVID-19 response systems.

CDC**:** Centers for Disease Control and Prevention. Serves as the national focus for developing and applying disease prevention and control, environmental health, and health promotion and health education activities designed to improve the health of the people of the United States.

CMS**:** The Centers for Medicare & Medicaid Services, CMS, is part of the Department of Health and Human Services (HHS).

COVID-19 Surveillance System**:** NCDHHS' surveillance tool known as "NC-COVID" for short.

CTSA**:** The Clinical and Translational Science Awards (CTSA) Program supports a national network of medical research institutions — called hubs — that work together to improve the translational research process to get more treatments to more patients more quickly.

Data use agreement (DUA)**:** A data use agreement (DUA) is an agreement that is required under the Privacy Rule and must be entered into before there is any use or disclosure of a limited data set to an outside institution or party.

EHR**:** Electronic Health Record; a digital version of a patient's paper chart. EHRs are real-time, patient-centered records that make information available instantly and securely to authorized users.

EPIC**:** Epic Systems is one of the largest providers of health information technology, used primarily by large U.S. hospitals and health systems to access, organize, store and share electronic medical records.

FDA**:** Food and Drug Administration (FDA); an agency within the U.S. Department of Health and Human Services. It consists of the Office of the Commissioner and four directorates overseeing the core functions of the agency: Medical Products and Tobacco, Foods and Veterinary Medicine, Global Regulatory Operations and Policy, and Operations.

HHS**:** The mission of the U.S. Department of Health and Human Services (HHS) is to enhance the health and well-being of all Americans, by providing for effective health and human services and by fostering sound, sustained advances in the sciences underlying medicine, public health, and social services.

HIE**:** Health Information Exchange; allows health care professionals and patients to appropriately access and securely share a patient's medical information electronically.

Inductive Approach**:** Method of qualitative data analysis which allows for themes to arise directly from the data.

Key Informant Interviews**:** A qualitative method utilized to discover a range of attitudes, opinions, or experiences of a particular phenomenon or event

Medicaid**:** Medicaid provides health coverage to low-income people and is one of the largest payers for health care in the United States.

N3C**:** National Covid Cohort Collaborative; a partnership among the NCATS-supported Clinical and Translational Science Awards (CTSA) Program hubs and the National Center for Data to Health (CD2H) (link is external), with overall stewardship by NCATS. Collaborators will contribute and use COVID-19 clinical data to answer critical research questions to address the pandemic.

N95 Respirator**:** An N95 respirator is a respiratory protective device designed to achieve a very close facial fit and very efficient filtration of airborne particles. Note that the edges of the respirator are designed to form a seal around the nose and mouth. Surgical N95 Respirators are commonly used in healthcare settings and are a subset of N95 Filtering Facepiece Respirators (FFRs), often referred to as N95s.

NC COVID**:** North Carolina COVID-19 Surveillance System; NC COVID creates a central repository of person-based public health data.

NC DETECT**:** The North Carolina Disease Event Tracking and Epidemiologic Collection Tool (NC DETECT) is North Carolina's statewide syndromic surveillance system.

NCDHHS**:** The North Carolina Department of Health and Human Services (DHHS) is a large state government agency in the U.S. state of North Carolina, somewhat analogous to the United States Department of Health and Human Services.

Negative pressure rooms**:** Negative pressure rooms, also called isolation rooms, are a type of hospital room that keeps patients with infectious illnesses, or patients who are susceptible to infections from others, away from other patients, visitors, and healthcare staff.

NSSP**:** the National Syndromic Surveillance Program; a collaboration among CDC, federal partners, local and state health departments, and academic and private sector partners who have formed a Community of Practice. They collect, analyze, and share electronic patient encounter data received from emergency departments, urgent and ambulatory care centers, inpatient healthcare settings, and laboratories.

PHI**:** Protected health information (PHI); The HIPAA Privacy Rule provides federal protections for personal health information held by covered entities and gives patients an array of rights with respect to that information. At the same time, the Privacy Rule is balanced so that it permits the disclosure of personal health information needed for patient care and other important purposes.

PII**:** personally identifiable information; any data that could potentially identify a specific individual.

PolicyMap**:** PolicyMap is a data and mapping tool for accessing data about communities across the U.S. to make better-informed decisions without any prior experience.

PPE**:** Personal protective equipment (PPE) refers to protective clothing, helmets, gloves, face shields, goggles, facemasks and/or respirators or other equipment designed to protect the wearer from injury or the spread of infection or illness.

PSA**:** Public service announcement; an announcement made for the good of the public

PUI**:** Patient Under Investigation; a patient presenting COVID-19 symptoms or a patient who has been in close contact with someone who tested positive for COVID-19.

SafeGraph**:** Provides a Points of Interest (POI) database and neighborhood patterns related to COVID-19.

Surge capacity**:** A measurable representation of ability to manage a sudden influx of patients. It is dependent on a well-functioning incident management system and the variables of space, supplies, staff and any special considerations (contaminated or contagious patients, for example).

Telehealth**:** The delivery and facilitation of health and health-related services including medical care, provider and patient education, health information services, and self-care via telecommunications and digital communication technologies.

Transmission rate**:** Also known as "attack rate." The rate of transmission is determined, primarily, by the hospitalization rate and total number of coronavirus diagnoses. A complex computer model is used to calculate the current rate. A transmission rate of 1.0 means that each person who has the virus passes it to one other person. A projection above a rate of 1.0 could result in a "spread". A rate below 1.0 means that the virus is spreading to fewer and fewer people.

WRAL**:** An NBC-affiliated television station licensed to Raleigh, North Carolina and serving the Triangle region.