Infodemics and Vaccine Confidence: Protocol for Social Listening and Insight Generation to Inform Action

Jessica Kolis, Kathryn Brookmeyer, Yulia Chuvileva, Christopher Voegeli, Sarina Juma, Atsuyoshi Ishizumi, Kaytlin Renfro, Elisabeth Wilhelm, Hannah Tice, Hannah Fogarty, Irma Kocer, Jordan Helms, Anisha Verma

Submitted to: JMIR Public Health and Surveillance
on: August 18, 2023

Table of Contents

Infodemics and Vaccine Confidence: Protocol for Social Listening and Insight Generation to Inform Action

Jessica Kolis[1] MPH; Kathryn Brookmeyer[2] PhD; Yulia Chuvileva[3] PhD, MA, MSc; Christopher Voegeli[4] PhD, MPH; Sarina Juma[5] MPH; Atsuyoshi Ishizumi[1] MSc, MPH; Kaytlin Renfro[6] PhD; Elisabeth Wilhelm[1] MA; Hannah Tice[7] BA; Hannah Fogarty[8] MPH; Irma Kocer[9] MPH; Jordan Helms[9] MPH; Anisha Verma[10] MPH

[1]Global Immunization Division Global Health Center Centers for Disease Control and Prevention Atlanta US
[2]Office of the Director National Center for HIV, Hepatitis, STD, and TB Prevention Centers for Disease Control and Prevention Atlanta US
[3]Division of Population Health National Center for Chronic Disease Prevention and Health Promotion Centers for Disease Control and Prevention Atlanta US
[4]Office of the Director National Center for Immunization and Respiratory Diseases Centers for Disease Control and Prevention Atlanta US
[5]Division of Workforce Development National Center for State, Tribal, Local, and Territorial Public Health Infrastructure and Workforce Centers for Disease Control and Prevention Atlanta US
[6]Division of STD Prevention National Center for HIV, Hepatitis, STD, and TB Prevention Centers for Disease Control and Prevention Atlanta US
[7]Division of State and Local Readiness Office of Readiness and Response Centers for Disease Control and Prevention Atlanta US
[8]Division of Overdose Prevention National Center for Injury Prevention and Control Centers for Disease Control and Prevention Atlanta US
[9]Division of HIV Prevention National Center for HIV, Hepatitis, STD, and TB Prevention Centers for Disease Control and Prevention Atlanta US
[10]Tanaq Support Services Atlanta US

Corresponding Author:
Jessica Kolis MPH
Global Immunization Division
Global Health Center
Centers for Disease Control and Prevention
1600 Clifton Rd
Atlanta
US

Abstract

Background: In the fall of 2020, the COVID-19 infodemic began to affect public confidence in and demand for COVID-19 vaccines in the United States. While national polls and surveys indicated what Americans felt about COVID-19 vaccines, they did not provide an understanding of why they felt that way nor the social and informational influences that factored into vaccine uptake. It was essential for U.S. Centers for Disease Control and Prevention (CDC) to gain a better understanding of the information environment and how it affected Americans' confidence in and demand for COVID-19 vaccines.

Objective: CDC established an Insights Unit within its COVID-19 Response Vaccine Task Force in January 2021 to assist CDC in acting more swiftly to address the questions, concerns, perceptions, information voids, and circulating mis- and disinformation that affected people's health decision-making, particularly around COVID-19 vaccines. The Insights Unit established a novel approach and methodology to rapidly detect and report on the trends in vaccine confidence and demand. The goal was to guide communication efforts and improve programmatic quality in near real-time through a continual and iterative listening and feedback process.

Methods: To inform the first COVID-19 State of Vaccine Confidence (SoVC) Insights Report, we identified and assessed available data sources. CDC's Vaccinate with Confidence framework and the World Health Organization's (WHO) Behavioral and Social Drivers for vaccine decision-making framework (BeSD) were selected as guiding principles for interpreting the impacts of findings on vaccine confidence and uptake. The Insights Unit established qualitative thematic analysis methods and a consensus-building approach to identify prevailing and emerging themes, assess their potential threat to vaccine confidence, and propose actions to increase confidence and demand.

Results: As of August 2022, the Insights Unit produced and distributed 34 reports to over 950 recipients within CDC and

externally. State and local health departments, nonprofit organizations, professional associations, and congressional committees have referenced and used the reports as resources for learning about COVID-19 vaccine confidence and demand, developing communication strategies to address low vaccine acceptance, and demonstrating how CDC monitored and responded to COVID-19 vaccine mis- and disinformation. Our methodology underwent continuous process improvement as workforce capacity and contextual changed to increase the rigor of the research process, the validity of the findings, and the usability of the reports.

Conclusions: While not a solution alone for infodemics, this methodology is a first step in managing an infodemic and can serve as a diagnostic technique for rapidly identifying opportunities for intervention and prevention. The methodology could be adapted and scaled for use in a variety of public health settings. It can be applied beyond acute public health crises to support adherence to guidance and recommendations and could be considered within routine monitoring and surveillance systems.

(JMIR Preprints 18/08/2023:51909)
DOI: https://doi.org/10.2196/preprints.51909

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?

✔ **Please make my preprint PDF available to anyone at any time (recommended).**
Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.
Only make the preprint title and abstract visible.
No, I do not wish to publish my submitted manuscript as a preprint.

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

✔ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**
Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain v
Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in <a href="http

Original Manuscript

Proposed Title

Infodemics and Vaccine Confidence: Protocol for Social Listening and Insight Generation to Inform Action

Authors

Kolis, Jessica MPH; Global Immunization Division, Global Health Center, Centers for Disease Control and Prevention, Atlanta, GA, USA, ORCID: 0000-0001-6946-8849
Brookmeyer, Kathryn, PhD; Office of the Director, National Center for HIV, Hepatitis, STD, and TB Prevention, Centers for Disease Control and Prevention, Atlanta, GA, USA, ORCID: 0000-0003-3149-3509
Chuvileva, Yulia, PhD, MA, MSc; Division of Population Health, National Center for Chronic Disease Prevention and Health Promotion, Centers for Disease Control and Prevention, Atlanta, GA, USA, ORCID: 0000-0001-7335-6110
Voegeli, Christopher, PhD, MPH; Office of the Director, National Center for Immunization and Respiratory Diseases, Centers for Disease Control and Prevention, Atlanta, GA, USA, ORCID: 0000-0002-1217-402X
Juma, Sarina MPH; Division of Workforce Development, National Center for State, Tribal, Local, and Territorial Public Health Infrastructure and Workforce, Centers for Disease Control and Prevention, Atlanta, GA, USA, ORCID: 0000-0001-6335-5229
Ishizumi, Atsuyoshi, MSc, MPH; Global Immunization Division, Global Health Center, Centers for Disease Control and Prevention, Atlanta, GA, USA, ORCID: 0000-0001-9678-5372
Renfro, Kaytlin, PhD; Division of STD Prevention, National Center for HIV, Hepatitis, STD, and TB Prevention, Centers for Disease Control and Prevention, Atlanta, GA, USA, ORCID: 0000-0003-4929-2545
Wilhelm, Elisabeth, MA; Global Immunization Division, Global Health Center, Centers for Disease Control and Prevention, Atlanta, GA, USA, ORCID: 0000-0002-4641-516X
Tice, Hannah, BA; Division of State and Local Readiness, Office of Readiness and Response, Centers for Disease Control and Prevention, Atlanta, GA, USA, ORCID: 0000-0003-3582-6470
Fogarty, Hannah, MPH; Division of Overdose Prevention, National Center for Injury Prevention and Control, Centers for Disease Control and Prevention, Atlanta, GA, USA, ORCID: 0000-0001-5307-4693
Kocer, Irma, MPH; Division of HIV Prevention, National Center for HIV, Hepatitis, STD, and TB Prevention Centers for Disease Control and Prevention, Atlanta, GA, USA, ORCID: 0000-0002-2396-4940
Helms, Jordan, MPH; Division of HIV Prevention, National Center for HIV, Hepatitis, STD, and TB, Centers for Disease Control and Prevention, Atlanta, GA, USA, ORCID: 0000-0002-1668-6575

Verma, Anisha, MPH; Tanaq Support Services ORCID: 0000-0003-4540-7675

JK, KB led the development of the methodology and were primary authors of the manuscript.

YC, CV, AI, SJ, KR, EW, and HT contributed to the development and evolution of the methodology and were responsible for research and writing sections of this manuscript.

HF, IK, JH, and AV contributed to the development of the methodology and evaluation and provided comments on manuscript.

Corresponding author: Jessica Kolis, CDC, ywe5@cdc.gov

Disclaimer

The findings and conclusions in this paper are those of the authors and do not necessarily represent the official position of the U.S. Centers for Disease Control and Prevention.

Abstract

Background: In the fall of 2020, the COVID-19 infodemic began to affect public confidence in and demand for COVID-19 vaccines in the United States. While polls indicated *what* consumers felt about COVID-19 vaccines, they did not provide an understanding of *why* they felt that way nor the social and informational influences that factored into vaccine confidence and uptake. It was essential for us to better understand how information ecosystems were affecting the confidence in and demand for COVID-19 vaccines in the United States.

Objective: The U.S. Centers for Disease Control and Prevention established an Insights Unit within its COVID-19 Response Vaccine Task Force in January 2021 to assist the agency in acting more swiftly to address the questions, concerns, perceptions, and misinformation that appeared to be affecting uptake of COVID-19 vaccines. We established a novel methodology to rapidly detect and report on trends in vaccine confidence and demand to guide communication efforts and improve programmatic quality in near real-time.

Methods: We identified and assessed data sources for inclusion through an informal landscape analysis employing a snowball method. Selected data sources provided an expansive look at the information ecosystem of the United States regarding COVID-19 vaccines. CDC's Vaccinate with Confidence framework and the WHO's Behavioral and Social Drivers for vaccine decision-making framework were selected as guiding principles for interpreting generated insights and their impact. We used qualitative thematic analysis methods and employed a consensus-building approach to identify prevailing and emerging themes, assess their potential threat to vaccine confidence, and propose actions to increase confidence and demand.

Results: As of August 2022, we produced and distributed 34 reports to over 950 recipients within CDC and externally. State and local health departments, nonprofit organizations, professional associations, and congressional committees have referenced and used the reports for learning about COVID-19 vaccine confidence and demand, developing communication strategies, and demonstrating how CDC monitored and responded to misinformation. A survey of the reports' end users found that nearly 80% of respondents found the reports "useful" or "extremely useful" and 61% used the reports to inform communication strategies. Additionally, our methodology underwent continuous process improvement to increase the rigor of the research process, the validity of the findings, and the usability of the reports.

Conclusions: This methodology can serve as a diagnostic technique for rapidly identifying opportunities for public health interventions and prevention. As the methodology itself is adaptable, it could be leveraged and scaled for use in a variety of public health settings. Further, it could be considered beyond acute public health crises to support adherence to guidance and recommendations and could be considered within routine monitoring and surveillance systems.

Keywords: Infodemic, Infodemics, Infodemic Management, Vaccine Confidence, Vaccine Demand, Misinformation, Disinformation, infodemiology, insights, mixed methods, thematic analysis, COVID-19

Introduction

Responding to the unprecedented COVID-19 pandemic has been made even more challenging by an extraordinary infodemic, defined as an overabundance of information, including credible and false or misleading information during a disease outbreak [1]. Information voids (i.e., lack of accurate information on a specific topic from credible sources), misinformation (i.e., inaccurate information), and disinformation (i.e., misinformation designed to achieve an agenda) have been hallmarks of the COVID-19 infodemic [2,3,4]. Further, information voids left by evolving science and rapidly spreading misinformation have served as breeding grounds for public confusion [3], posing a serious risk to compliance with public health prevention and the uptake of mitigation measures, such as mask-wearing and vaccination efforts [4,5].

In the fall of 2020, the influence of the infodemic on public confidence in COVID-19 vaccines was evident before any vaccine was granted an emergency use authorization. National polls and surveys conducted in September and December 2020 indicated wide-ranging levels of public confidence in COVID-19 vaccines and their rollout [6]. In December 2020, before any COVID-19 vaccine was authorized for emergency use, more than a quarter of those surveyed said they would not get vaccinated, citing concerns about the side effects and safety, lack of trust in the government, and worries about the speed at which vaccines were developed [7]. Additionally, in December 2020 and January 2021, two-thirds of U.S. adults (65%) said the federal government was doing a “fair” or “poor” job of distributing vaccines [7,8]. While these large-scale surveys signaled *what* Americans felt about COVID-19 vaccines, they did not provide an understanding of the reasons *why* they felt that way and the social and informational influences that factored into vaccine confidence and uptake.

Creating CDC’s Insights Unit

In late 2020, the leadership of the U.S. Centers for Disease Control and Prevention’s (CDC) COVID-19 Response recognized that the agency needed to act more swiftly to address the questions, concerns, perceptions, information voids, and circulating mis- and disinformation that affected people’s health decision-making, particularly around COVID-19 vaccines. At that time, there was no established, coordinated mechanism for collecting, reviewing, and synthesizing qualitative and quantitative data from multiple CDC-owned and external sources to assess the state of vaccine confidence in the United States. Initially, this unit aimed to help CDC communications and improve programmatic quality in near real-time based on the dynamic needs of Americans through a continual and iterative listening and feedback process.

The Insights Unit was established in January 2021 and became the first infodemic management unit of its kind at the CDC. The Insights Unit was part of the Vaccine Confidence and Demand Team within the Vaccine Task Force of CDC’s COVID-19 Response. The Insights Unit established a novel approach and methodology to rapidly detect trends in vaccine confidence by integrating multiple data sources and using established frameworks to understand the knowledge, attitudes, and behaviors of Americans regarding the COVID-19 vaccines. The aim was to use this methodology to help inform communication and programmatic strategies at a national level [9].

The initial Insights Unit was a three-person, multi-disciplinary team, led by a health communication specialist with experience in immunization campaigns and a behavioral scientist, with additional support from a data analyst with a public health background. The unit remained multi-disciplinary and ranged in capacity between one to two full-time unit leads and one to six full- and part-time data analysts during the first year.

Manuscript Aims and Goals

This manuscript aims to outline the Insights Unit's first year, and the establishment, approach, and methodology for COVID-19 State of Vaccine Confidence (SoVC) Insights Reports so that others can replicate and adapt its processes and methods. We hope to help health authorities and public health professionals explore opportunities for establishing insights units or complex social listening mechanisms for any public health area of concern. They can use these units and their findings to examine the information ecosystems to gain insights into how their community's thoughts and feelings affect critical health decisions and use it to enact policies, design and adapt programs and inform communication campaigns.

Methods

Analysis Plan Development

Planning

Landscape Analysis

Prior to the establishment of the Insights Unit, there were several social listening activities about COVID-19 being conducted by the United States government and external organizations. Many of the activities even within CDC were being conducted and interpreted independently with little to no synthesis or coordination. We conducted an informal landscape analysis in January 2021 to identify and evaluate potential data sources (primary and secondary) for inclusion in our reports. We identified data sources by interviewing colleagues via a snowball sampling strategy within the COVID-19 Response and across CDC with colleagues recommending others for us to interview about potential sources for inclusion. Some interviews were conducted by phone, while others over email. The interviews remained informal with no standard set of questions for each colleague interviewed, however each had the same objective of identifying what social listening data or sources they used, conducted or created, or were aware of internal or external to CDC. These informal interviews and emails helped us identify all currently available social listening data sources within the agency and better understand what types of data sources they used in social listening and monitoring efforts.

The identified data sources were then categorized as mixed methods reports (i.e., reports that included multiple primary data sets, or multiple types of data sets, such as social media and news media monitoring), social listening (i.e., data collected from social media aggregation tools or native platform searches), direct reports (i.e., primary data sets, such as media request line list, CDC-info inquiry line list, etc.), and research (e.g., recently published peer-reviewed and gray literature, polls and surveys from accredited institutions and organizations). We informally assessed 15 identified data sources for accessibility, suitability, and methodological rigor (*see Table A: Evaluation Criteria for Data Sources* which included both primary data sources (i.e., data collected directly by CDC that could be analyzed by our analysts) and secondary data sources (i.e., reports of data already collected and analyzed by analysts outside the Insights Unit). The first criteria focused on accessibility where we considered where or how we could access the data, whether primary or secondary data, how frequently data sets were made available, and if there was a cost or subscription needed to access data (e.g., Meltwater). After accessibility was assessed for each source, we considered the data source's suitability and whether it would be able to help us understand vaccine confidence. Lastly, we considered the methodological rigor of each data source including considering sampling strategies for primary data and analytic approaches for secondary data sources. The first report, produced less than a month from the creation of our unit, included data from 11 sources from the

initial landscape analysis. Our application of evaluation criteria became more rigorous as the unit expanded and we better understood our utilization and quality of data sources.

Table A. Evaluation Criteria for Data Sources

Caption: Criteria categories and their definitions that were utilized to evaluate each potential data source prior to inclusion in social listening and analysis.

Criteria	Definition
Accessibility	●How easily the unit could access the data or the findings. ●How frequently the data or results were made available. ●The cost, if applicable, associated with accessing the data.
Suitability	●Comparing the available data with the intended purpose of reporting on the state of vaccine confidence.
Methodological rigor	●Data collection methods used in pulling the data. ●Analytic approaches used when reports presented findings.

After initial inclusion criteria were evaluated, we categorized each source into one of the following categories: mixed methods reports, social media listening, direct reports, and research. We then considered the type and amount of data within each source to determine appropriate techniques for data collection and analysis and clarified what the data would be able to tell us in relation to vaccine confidence. (*see Table B: Data Source Categories and Intended Utilization*).

Table B. Data Source Categories and Intended Utilization

Caption: Data sources used were categorized into one of the four categories (mixed method reports, social media listening, direct reports, and research) and their potential utilization approaches were identified based on those categories. These approaches outlined what type analyses might be used and what issues they might be able to elucidate.

Data Source Type	Definition	Utilization
Mixed Methods Reports	●Reviewed multiple data primary data sets ●Included multiple types of data sets (i.e., social media listening, news media monitoring, etc.)	●Trending and emerging topics/keywords ●Changes in information seeking patterns ●Understanding sentiment ●Information gaps/voids ●Socio-behavioral indicators ●Vaccination barriers ●Misinformation narratives
Social Media Listening	●Data collected from social media aggregation tools or native platform searches	●Conversation levels by topics (i.e., # of posts, level of engagement with posts, etc.) ●Share of voice by topic or sub-topic, percentage of total conversation occupied by a single topic or sub-topic ●Trending and emerging topics/keywords ●Changes in information seeking patterns ●Understanding sentiment ●Information gaps/voids ●Socio-behavioral indicators ●Vaccination barriers ●Misinformation narratives
Direct Reports	●Primary data sets (i.e., media request line list, CDC-info inquiry line list, etc.)	●Conversation levels by topics (i.e., # of posts, level of engagement with posts, etc.) ●Share of voice by topic, sub-topic ●Trending and emerging topics/keywords ●Changes in information seeking patterns ●Information gaps/voids ●Vaccination barriers

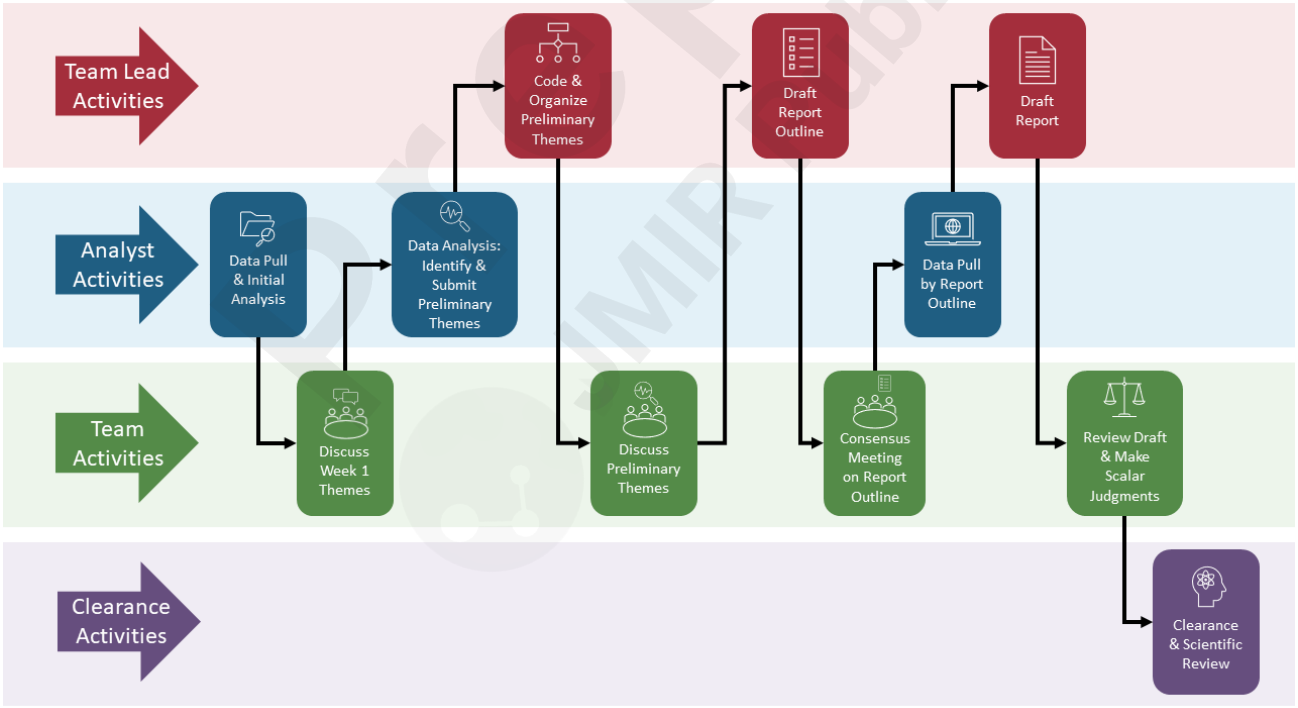
		●Misinformation narratives
Research	●Recently published peer-reviewed and gray literature ●Internal and External polls from accredited institutions and organizations	●Identify socio-behavioral indicators ●Determine vaccination intent ●Identify vaccination barriers

Analysis and Reporting Cadence

During the Insights Unit's first year (January 2021–January 2022), the team created and disseminated SoVC reports biweekly. Data collection and analysis occurred concurrently with report production and clearance, which ensured that the reports provided insights on emerging themes quickly and without gaps in data collection. Analysts performed data collection and analysis for their assigned data segments for two weeks and submitted preliminary findings to unit lead(s). Unit lead(s) then led multiple, consensus-building meetings to develop the report outline, draft report narrative, and complete the scientific review (See *Figure 1: Visualization of Report Production*).

Figure 1. Visualization of report production

Caption: Reports went through a multistep process with phases that occurred at separate levels (e.g., team lead, individual analyst, team, and clearance). Analysts interpedently collected and analyzed their assign data source and using a lone-wolf coder approach team lead triangulated individual analyst findings. Team collectively identified preliminary themes and conducted further data collection and analysis. This data collection and analysis then informed the creation of reports and scalar judgements. A consensus building process was utilized throughout with team lead and analysts working collectively to reach agreement on findings.



Social Listening and Integrated Analysis

Analytical Frameworks

We considered two key frameworks in the development of our approach for analyzing and interpreting our collected data. The first was CDC's Vaccinate with Confidence strategy, which was

created to bolster confidence in COVID-19 vaccines (see *Figure 3: CDC Vaccinate with Confidence Strategy for COVID-19 Vaccines*) [12]. The second was the World Health Organization's (WHO) Behavioral and Social Drivers for vaccine decision-making framework (BeSD), which includes four domains that influence vaccine demand and uptake (see *Figure 4: BeSD Framework*) [13]. Based upon these frameworks and informed by UNICEF's Vaccine Misinformation Management Field guide [14], as well as the UK Government Communication Service's RESIST Counter-Disinformation Toolkit [15], a threat matrix was developed to provide further context to our findings and determine possible actions to address dominant and emerging vaccine confidence issues (see *Figure 5: COVID-19 State of Vaccine Confidence Insights Report Threat Matrix*).

Figure 3. CDC Vaccinate with Confidence Strategy for COVID-19 Vaccines

Caption: Vaccinate with Confidence is the strategic framework of the Centers for Disease Control and Prevention (CDC) to strengthen confidence in COVID-19 vaccines through three strategies: building trust, empowering healthcare personnel, and engaging communities and individuals. Strong confidence in COVID-19 vaccines within communities leads to more adults, adolescents, and children getting vaccinated — which leads to fewer COVID-19 illnesses, hospitalizations, and deaths.

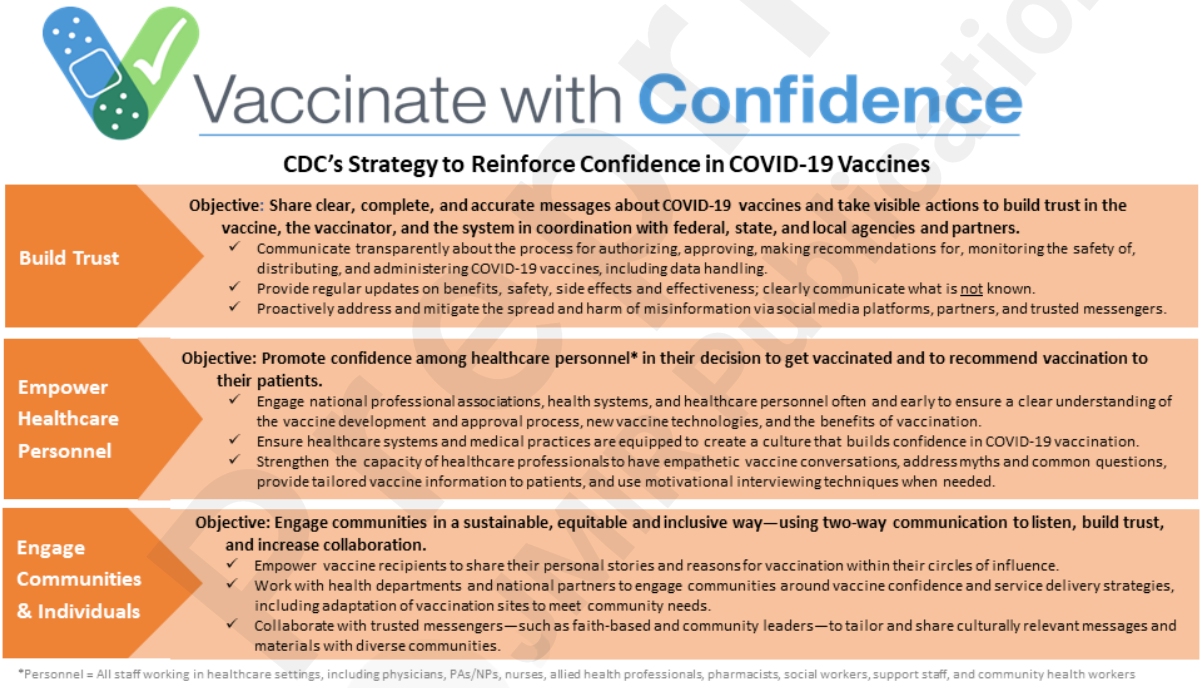


Figure 4. The Behavioral and Social Drivers (BeSD) Framework.

Caption: The Behavioral and Social Drivers (BeSD) framework, based on Brewer et. al's Increasing Vaccination Model, is a tool used to understand how the four domains (thinking and feeling, social processes, motivation, and practical issues) impact intent to vaccinate and low vaccine uptake.

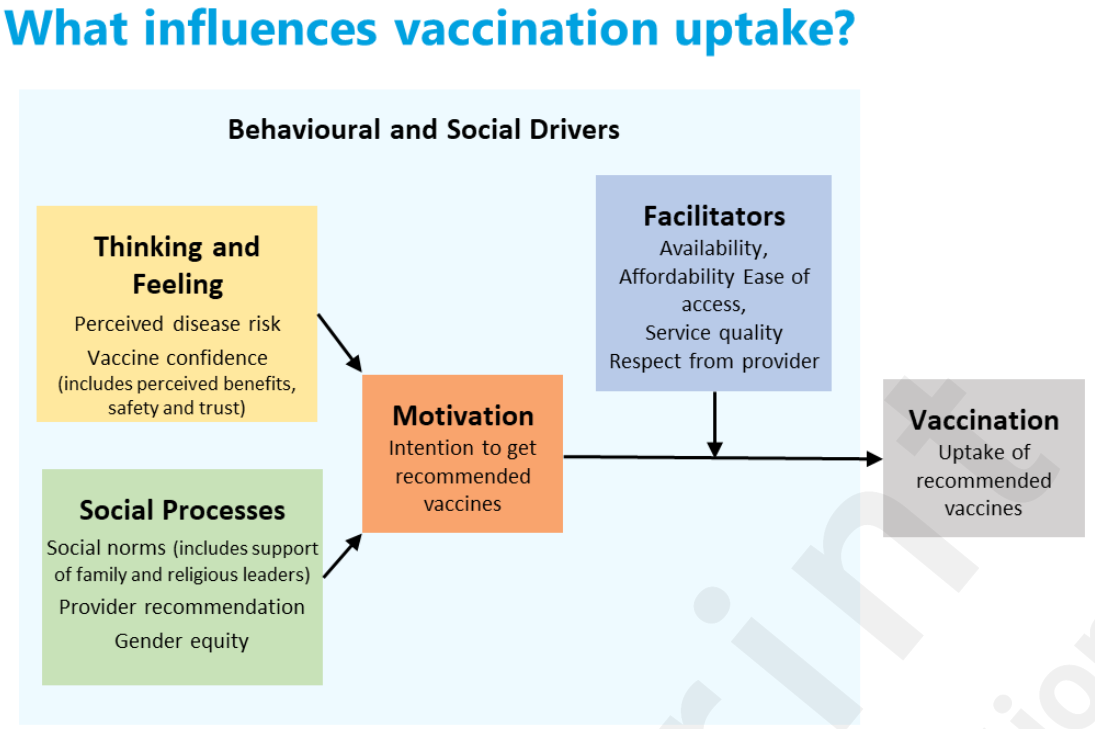
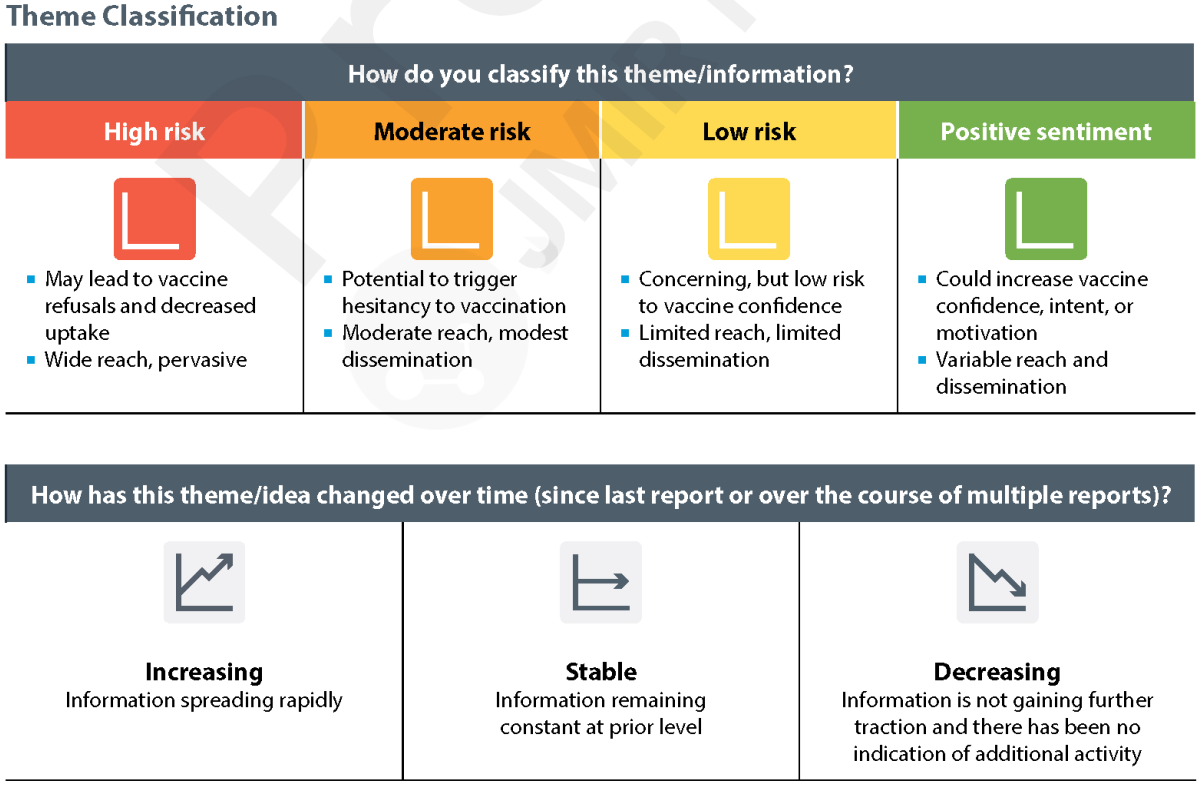


Figure 5. COVID-19 State of Vaccine Confidence Insights Report Threat Matrix.

Caption: CDC's COVID-19 Vaccine Confidence Insight Unit developed a threat matrix to help denote the potential threat that an identified theme would have on one's intent to vaccinate. Themes were classified by level of threat to vaccination (e.g., high risk to impact vaccination) and directionality which indicated the relative volume and prevalence of a theme across information systems.



Data Segment Analysis

Analysts were assigned to at least one designated data segment. Data segments were grouped as categories of data and included external social listening (i.e., social media broadly), internal social listening (e.g., engagement on CDC-owned social media platforms, calls or email inquiries to CDC-INFO), media monitoring (i.e., news media monitoring, media request line list), research (i.e., literature, surveys, etc.), and internet trends (e.g., Google trends, website traffic, etc.). Each analyst was responsible for data collection and preliminary analysis of their assigned data sources. Analysis methods varied slightly depending on the data source category and how the data source could be utilized based on the prior evaluation (see *Table C: Data Segment Collection and Analysis*). All analyses used established qualitative theme identification techniques [10]. For data segment analysis, the techniques of repetition, such as looking for high volume of key terms of phrases in inquiries to CDC-INFO or mentions in news media, and indigenous categories, such as looking for unique terms or phrases like "viral-shedding," were predominantly used to identify preliminary themes (see *Table C: Applied Qualitative Data Analysis Methods)* [11]. Additionally, analysts also used a mixed deductive and inductive approach to identify preliminary themes. Deductively, analysts organized data into major categories represented by the aforementioned analytical frameworks: vaccines, vaccinators, vaccine system, and outside the vaccine system (see Table E. COVID-19 State of Vaccine Confidence Code Book). Inductively, they allowed themes to emerge organically not allowing the assembled code book to limit their findings and potentially miss new emerging themes, topics, or terms. Following initial identification, analysts further examined their identified themes to determine if an individual theme was a single development with a loud "echo" (e.g., multiple media sources around the country reported on the same new workplace vaccination policy rolled out by a company, and repeated posts, inquires, and searches about that particular policy) or a repeated development with multiple iterations of the theme across the data segment (e.g., different media sources around the country reported on multiple companies rolling out a similar workplace vaccination policies and conversations online amplified each of the individual policies separately) and documented which type it was.

Table C. Data Segment Collection and Analysis

Caption: All data sources were categorized into segments with similar data. The table shows what each data segment was comprised of and lists the exact source, how data was collected and analyzed and if there were additional inclusion/exclusion criteria.

Data Segment	Source	Collection Method	Additional Criteria	Analysis Method
Primary Data				
News Media[1]	Meltwater[2]	Boolean search string (See *Table E: COVID-19 State of Vaccine Confidence Insights Boolean Strings*)	English only, United States Only	- Identify top and unique keywords in headlines and copy through established Boolean search strings by number of mentions - Identify themes by looking for repletion following sorting mentions by reach as calculated by Meltwater until saturation (extremely high volumes meant review of all lines of data impossible given time frame) - Focused searches with top and unique keywords via snowball methods for context determination

[1] Due to the high volume of data, it was not possible to review every data point. Analysts reviewed data points until new themes could no longer be identified having reached a point of data saturation.
[2] Social Listening Aggregator Tool

	CDC Media Office	Weekly line list	Inquires routed to COVID-19 Vaccine Task Force	- Identify themes based on top and unique keywords - Determine directionality of emerging themes based on share of voice week to week
External Social Listening[3]	Meltwater[2] CrowdTangle[4]	Boolean search string	English only	- Identify top and unique keywords in headlines and copy through established Boolean search strings - Identify themes by looking for repetition following sorting mentions by reach as calculated by Meltwater until saturation (extremely high volumes meant review of all lines of data impossible given time frame) - Focused searches with top and unique keywords via snowball methods for context determination
Internal Social Listening[5,6]	Facebook Twitter Instagram	Native platform searches	N/A	- Identify top key words and unique key words within comments or post interactions on CDC-owned social media channels
	CDC-Info[7]	Weekly line list	- Contains COVID-19 and "vaccine" or "vaccinate" - Duplicates removed	- Random sample of 10% of the questions to identify themes in the questions - Identify top and unique keywords within samples
Research & Academics	Peer Reviewed and Preprint Literature	PubMed LitCovid Google Scholar	- Published or made available during the reporting period - Focus on the US population at national, state, or other jurisdictional levels - Opinion pieces excluded	- Articles reviewed to determine if findings related to variables within BeSD framework and CDC's Vaccinate with Confidence Strategy for COVID-19 vaccines.
	Polls and Surveys	National Immunization Survey Kaiser Permanente Foundation Google	- Data collected align with indicators in BeSD framework and CDC's Vaccinate with Confidence Strategy for COVID-19 vaccines - Polls collected or published during the reporting period	- Longitudinal polls reviewed to see changes in vaccine confidence over time and to achieve point-in-time estimates for reporting period - Polls evaluated for focus on specific populations or geographic areas
Web Metrics[8]	Internal web traffic	Adobe Analytics	- N/A	- Identify most frequently viewed CDC webpages and frequently asked questions - Identify changes in weekly top

[3] Public posts on social media not interacting directly with CDC-owned social media channels on Twitter, Facebook, Instagram, Pinterest, Reddit, and YouTube.
[4] A public insights tool from Meta.
[5] Public posts on social media interacting directly with CDC-owned social media or direct inquires to the agency via CDC-INFO.
[6] Due to the high volume of data, it was not possible to review every data point. Analysts reviewed data points until new themes could no longer be identified having reached a point of data saturation.
[7] CDC's national contact center, CDC-INFO. CDC-INFO offers live agents by phone and email to help people find the latest, reliable, and science-based health information on more than 750 health topics.
[8] Analyses of web search data looked to understand trends in online search activity, which can be indicators of emerging concerns, questions, or information voids.

				search terms
	External web searches	Google Trends	- Data from Google Trends, using an automated web scraping package available in R software looking at rising search queries and topics related to a pre-specified set of keywords across Google and YouTube.	- PowerBI[9] was used to create tree map charts (visual representation of search interest data, using chart size dimensions proportional to percent increase) for quickly identifying keywords or topics experiencing the largest increase in search frequency on a given day.
Confirmatory Data Segments				
Web Metrics		- SEMrush	N/A	- SEMrush was used for targeted analysis of specific topics once themes were identified through primary sources
Third Party Reports		- COVID-19 Joint Information Center Communication Surveillance Report - Federal Emergency Management Agency Social Listening Report - Virality Project - Project Vector - NCIRD social listening report - First Draft News Vaccine Misinformation Insights Report - CrowdTangle content insights report	Reports published during report period or considered data from data within the report period.	- Identify themes based on top and unique keywords and themes - Used as means to validate findings after triangulation to avoid double counting

*Table D. Applied Qualitative Data Analysis Methods**

Caption: Analysts employed 4 key types of qualitative analysis techniques to identify preliminary theme. Below are the different techniques utilized and how they were defined for our use.

Technique	Definition
Repetition	Themes that come up repeatedly in a single input or across our data inputs.
Similarities and Differences	Data points that differ or are like each other (which may make up sub-themes to a larger overarching theme); themes that are like others that were presented in prior reports (continuing themes); or themes that are similar but with unique elements from others that were presented in prior reports (evolving themes).
Indigenous Categories	Technical or slang-sounding terms that are used in new or unique ways by the community under study, such as "viral shedding" or "medical segregation."
Missing Data	What should or could be talked about but is not, such as when media reports on a development but social media is quiet on it or a discussion about a development is missing critical information that exists (helping us identify information gaps).

[9] Data visualization software

Table E. COVID-19 State of Vaccine Confidence Code Book[10]

Caption: A qualitative code book was established because of our process evaluation that broke up our themes into various codes and developments. The description column describes inclusion criteria for each development.

Code	Development	Description
Data	Data (general)	● Any data related to vaccination or vaccination research
	Data Issues	● Lack of data ● Data inequities ● Unsettled science or contradictory data ● Lack of data disaggregation (e.g., by ethnic group) ● Lack of specific data (e.g., specific vaccination outcomes.) ● Misinformation about: ○ VAERS data ○ Debunked data or studies
	Supply and demand data	● Allocation/availability of vaccines by state/county ● Demand outstripping supply ● Supply outstripping demand ● Excludes vaccine shortages or wastage
	Coverage/uptake data	● Vaccine uptake/vaccine coverage data ● Vaccine series completion data
	Confidence data	● Polls and other data on elements of vaccine confidence (e.g., intention to vaccinate or how events affected confidence)
	Non-vaccine data decreasing urgency to vaccinate	● Non-vaccination research/data that may decrease urgency to vaccinate (e.g., declining COVID-19 cases and deaths or studies showing lower severity illness on certain groups)
	Non-vax data increasing urgency to vaccinate	● Data unrelated to vaccine safety and effectiveness (e.g., COVID-19 risk data and negative psychological and socioeconomic effects of the pandemic)
Vaccines	Vaccines (general)	● Verified developments relating to the vaccines themselves (e.g., vaccine effectiveness and safety) ● Any developments that damage the safety/effectiveness reputation of vaccines have an outsized impact on intentions to vaccinate.
	Effectiveness	● Vaccine effectiveness data in general, for specific populations (e.g., immunocompromised individuals), or against new variants (e.g., Alpha, Delta, etc.) ● Discussions that affect the perception of effectiveness (e.g., breakthrough infections, need for booster doses, etc.) ● Discussions of administration affecting effectiveness (e.g., mixing of vaccine brands, longer than recommended lag between doses, etc.) ● Discussions of vaccine effectiveness compared to "natural immunity" from a previous COVID-19 infection

[10] Inclusion criteria for all codes required development or finding to have occurred during the reporting period specific for a single report.

		• Misinformation about effectiveness of vaccines.
	Safety	• Positive and negative developments/data about safety of COVID-19 vaccines • Perceptions of safety, including during and post-vaccination discomfort, side effects, and adverse events • Misinformation about adverse events after vaccination; safety concerns for specific groups of people or for vaccine ingredients
Vaccinators	Vaccinators (general)	• Institutions and people involved in administering vaccines (e.g., doctors, nurses, etc.) • Vaccination sites (e.g., clinics, mass vaccination sites, etc.) • Organizations and entities (e.g., pharmacy retail programs, hospitals, etc.) focused on vaccination administration • Non-traditional vax sites (e.g., schools, religious spaces, etc.)
	Trusted vaccinators	• Recommendations to get vaccinated from trusted healthcare providers (e.g., doctors, nurses, and healthcare workers) • Stories of trusted vaccinators recommending and administering vaccines • Expressed hesitancy among trusted vaccinators
	Practical issues	• Issues accessing vaccination (e.g., challenges in booking appointments or physically getting to appointments)
	Inequity in vaccine access	• Stories specifically about differences in vaccine access for specific groups
	Administrative errors	• Concerns about stories of errors made by the vaccinators (e.g., administering incorrect doses, incorrect duration between doses, improperly storing or refrigerating vaccines, etc.) • Does not include intentional errors (e.g., June 2021 story of a pharmacist tampering with 500 vaccines)
	Unfilled appointments	• Issues of underutilization of available vaccines
Vaccine System	Vaccine system (general)	• Institutions, policies, and processes that bring vaccines into being and distribute them to the population (excluding vaccines and vaccinators)
	Research and development	• Creation of vaccines (e.g., vaccine development, clinical trials, booster dose development, etc.) • Does not include stories about the need for additional testing for different populations that were inadequately represented in clinical trials
	Licensing and authorization	• Developments around the licensing, authorization, and continuous safety tracking/monitoring of the vaccines (e.g., emergency use authorization, advisory panels for CDC and FDA, etc.) • Does not include misinformation about VAERS
	Supply	• Manufacturing to distribution • Storage and handling • Vaccine shortages

		• Vaccine wastage and disposal
	Demand generation and barriers	• Developments that aim to increase demand for vaccines (e.g., incentives) • Does not include expansion of vaccination efforts
	Official guidance	• Discussion or reports of official guidance from trusted public health agencies (e.g., CDC, WHO, etc.) • Guidance for vaccination and people who are fully vaccinated
	Malpractices	• Stories of malpractices within the vaccine system can damage its reputation and reduce the trust that the public places in its institutions to do the best by public health and not by some other interest like profit and to act from a place of care rather than from prejudice and bias. Includes both new and past mistakes and bad practices.
	Conspiracies	• Claims about government conspiracies, “Big Pharma conspiracies”, and previously debunked malpractices
Outside Vaccine System	Outside vaccine system (general)	• Actions taken by economic entities, such as businesses or employers, influencers, and the general public • Developments within digital media where powerful rhetoric can become a development all on its own
	Requirements	• Policies about vaccination requirements (e.g., employee or school requirements) • Proof of vaccination systems (i.e., "vaccine passports") • Travel requirements
	Economic closing or re-opening	• Restrictions to slow the spread of COVID-19 • Lifting of COVID-19 restrictions and “return to normal”
	Public behaviors	• Actions that support or undermine vaccinations in others (e.g., fake vaccine scams, vaccine-card fraud, etc.) • Discussions of vaccine (brand) "shopping", "vaccine hunting", and "vaccine tourism” • Violence over masking or vaccination requirements • Social norms among the vaccinated (e.g., continuing to wear masks and social distancing post-vaccination) or among unvaccinated
	Misinformation policing	• Stories and discussions of digital company policies to remove or flag problematic content or users
	Influencer narratives	• Discussions by prominent public figures that may support or undermine vaccine confidence (e.g., expert critiques, the lack of intention to get vaccinated by prominent politicians or celebrities, etc.) • Stories or discussions of getting vaccinated
	Public narratives	• Discussions by members of the public that either support or undermine vaccine confidence (e.g., vaccine selfies, personal accounts of magnetism or adverse effects, public vaccine refusals, etc.) • Persistent anti-vaccination tropes (e.g., medical segregation, experimentation, etc.)
	Cultural clashes	• Discussions of mandates impinging on individual freedoms

		and lack of vaccination as a sign of resistance to being controlled ● May include broader tropes brought in to justify vaccinations or decisions not to vaccinate as being rooted in cultural ideas about the role of government, business, or science in society
	COVID-19 denialism	● Tropes that underplay the seriousness of COVID-19 (e.g., denying that it exists, arguing that only certain people are vulnerable, spread is exaggerated, etc.) ● Discrediting preventive measures to control COVID-19 (e.g., masking, social distancing not effective) ● Does not include discussions about vaccines providing protection
	Unverified COVID treatments	● Information claiming that COVID-19 can be prevented or treated by unverified treatments or that because of such treatments, vaccines are not necessary ● Includes unverified or debunked claims that certain herbs, essential oils, or chemicals can treat COVID or reduce its symptoms or their severity.

Techniques to Identify Themes. Field Methods. 2003;15(1):85–109.

Table F: COVID-19 State of Vaccine Confidence Insights Unit Boolean Strings

Caption: Below is an example of the Boolean string used in Meltwater to collect news and social media mentions. All strings used the “Base Boolean String” and “sub-strings” were added using the “AND” operator to help dive deeper into the subtopics of interest.

Base Boolean String[11]	
((((("vaccine" OR "vaccination" OR "vaccines" OR “vaccinate” OR "immunization" OR "immunizations" OR "immunize" OR "vax" OR "vaxx" OR "shot" OR "shots" OR "va((ine" OR "va((!ne" OR "#vaccine" OR "vaccinated" OR "vacine" OR "vaccin" OR "vacines") AND ("COVID-19" OR "COVID19" OR "COVID" OR "Coronavirus" OR "Corona" OR "China virus" OR “Chinavirus” OR “Chinese virus” OR “Wuhan virus”)) OR ("covid19vaccine" OR "covidvaccine" OR "covidvax" OR "coronavaccine" OR "covid19vaccines"))	
Selected sub-strings[12,13]	
Effectiveness	("effective" OR "effectiveness" OR "effectivness" OR "efficacy" OR "transmission" OR “transmit” OR "spread" OR "prevent" OR "prevents"))
Side Effects	("side effect" OR "side effects" OR "pain" OR "fever" OR "headache" OR "redness" OR "red" OR "sore" OR "fatigue" OR "exhausted" OR "sleepy" OR "muscle" OR "aches" OR "chills" OR "tiredness" OR “symptom” OR “tired” OR “headache” OR “nausea” OR “nauseous”))
Adverse Events	("damaged" OR "injured" OR "maimed" OR "bells palsy" OR "bell's palsy" OR "bells' palsy" OR “bellspalsy” OR "injury" OR "blood disorder" OR "covid arm" OR "covidarm" OR "paralyzed" OR "paralyze" OR "adverse" OR "ruin" OR "seizure" OR "seizures" OR “faint” OR “hospital” OR “allergic reaction” OR “paralysis” OR “paralyzed” OR “unresponsive” OR “severe reaction” OR “VAERS”))
Fertility/Pregnancy	("pregnancy" OR "pregnant" OR "miscarriage" OR "fertile" OR "fertility" OR "fetus" OR

[11] Used for all search strings.
[12] Added to base string via AND operator.
[13] Note: Search strings evolved over time.

	"fetal" OR "birth" OR "defect" OR “unborn” OR “womb”))
Access	("appointment" OR "appointments" OR "appt" OR "appts" OR "schedule" OR "sign up" OR "register" OR "slot" OR "time" OR "scheduled" OR "scheduling" OR "registering" OR “cancel” or “access”))
Natural Immunity/Previous Infection	(“previously infected” OR “natural immunity” OR “recovered”))
Safety	(“safe” OR “safety” OR “protect” OR “protected” OR “prevent”))
Variants	(“variant” or “strain” or “variants” or “varients” or “ varient” or “B.1.617” or “P.1” or “B.1.351” or “B.1.1.7” or “UK variant” or “india variant” or “Indian variant” or “brazil variant” or “south Africa variant” or “SA variant”))

Integrated Thematic Analysis

For each reporting period, analysts submitted preliminary findings from their assigned data segment(s) into a spreadsheet before the first consensus-building meeting. Each analyst could identify and submit an unrestricted number of preliminary themes. Each line of the spreadsheet contained one preliminary theme with a high-level summary of its content, links to illustrative examples, and identified the data source(s) originating the theme.

Following the submission of all preliminary themes for each data segment, unit lead(s) reviewed all entries in the spreadsheet and applied codes to categorize the entries by themes that spanned across multiple data segments. The unit lead(s) then sorted the codes to guide discussion during the first of three consensus-building meetings that focused on theme generation. The categories with the most entries in the spreadsheet received the greatest dedicated time for discussion, with each analyst sharing their analysis as the theme appeared (or did not) in their assigned data segment(s). The unit then spent the rest of the time discussing unique or smaller emerging themes.

Based upon this initial consensus-building meeting, unit lead(s) assembled an outline before the second consensus-building meeting that focused on collaboratively confirming and sorting themes into three categories: main, emerging, or continuing and evolving. Main themes were defined as the most pervasive themes identified during that period, appearing to affect vaccine confidence broadly or among a specific group of people. Emerging themes were identified as new, lower-volume themes that were gaining traction or higher-volume themes that did not appear in previous reports. Continuing and evolving themes included topics that had been covered in previous reports but had shifted in nuance or emerged in different data segments from when it was originally identified. Lead(s) explained how and why they categorized themes based on the findings and discussion from the theme generation meeting. They also explained why the themes were categorized as main, emerging, or continuing and evolving categories. Analysts then voiced agreement or disagreement with the categorization by providing additional context and feedback. The lead(s) used the discussion to re-organize and re-frame the themes to more accurately represent how they emerged in the data—in some cases, dividing or combining preliminary themes. The team finished these meetings by coming to a consensus about the final outline for the report. Following the meeting, analysts pulled relevant data for each theme that they had already identified from their data segment(s) and conducted a secondary deep-dive analysis on each of the outlined themes. The analysts collected raw data from all segments into a single document, including illustrative visual and text/quote examples of the themes in social media posts, poll results, and news headlines. Unit lead(s) reviewed all relevant data provided by the analysts on each designated theme for the report and performed additional analysis to uncover nuances within each identified theme.

Figure 2. Example from SoVC Report #11: Process from Data Segment Analysis to Final Outline

Caption: This figure highlights the steps conducted for integrated analysis for each report. This figure contains actual data from the 11th report. Step 1 represents the preliminary themes identified by data segment. Step 2 represents the themes submitted in step 1 sorted by the code book to look for congruence. Step 2's results help guide step 3's consensus-building meeting helps which is used to draft a report's outline in step 4. During step 5, the insights unit reviews the outline developed in step 4 and once confirmed, return to their data sources to pull relevant samples of data based on the themes in the outline. In step 7, reports are drafted using the data collected in step 6 and report is review with team to ensure accuracy and assess threat level. In step 8, report outline is refined, and report finalized and submitted to clearance.

Step 1: Analysts report preliminary themes by data segment.						
Data Segment	External Social Media (e.g., Twitter, Reddit, etc.)	Internal Social Media (e.g., CDC-owned social media channels)	Third Party Reports (e.g. Communication Surveillance Report)	News Media	Polls & Literature	Web Metrics (e.g., Google Trends)
# Themes	6	6	9	39	15	8

Step 2: Preliminary themes coded by team leads.								
Top Codes (≥ 5 prelim. themes)	Vaccines/ Safety	Outside vaccine system/ Requirements	Data/ Data Issues	Vaccine System/ Guidance for fully vaccinated	Data/ Vaccine Uptake	Vaccines/ Effectiveness	Outside vaccine system/ COVID-19 restrictions	Data/ Confidence Data
# Themes	13	12	10	8	7	7	5	5

Step 3: Consensus-building meeting to develop outline.

Step 4: Proposed report outline developed by team leads.

Category	Theme
Major	Spread of Delta variant in US causing concern about vaccine effectiveness • Coverage of/discussions about delta variant • Concerns and questions about breakthrough cases, booster doses, asymptomatic spread and transmission among fully vaccinated people, and effectiveness of specific vaccines • Mis/disinformation about vaccines causing the virus variants
	Concerns vaccine mandates infringe on individual liberties and medical freedom. • Coverage of/discussions about vaccine requirement policies (e.g., employers, travel, etc.) and plans, vaccine-related legislation, mask requirements, and vaccine status/privacy • Consumer concerns and questions about legality of vaccine requirements, requirements of vaccines under emergency use authorization, exemptions and "right to opt out," and tracking vaccination status and breakthrough cases • Polls & Research about opinions on mandates/requirements, societal impacts of requirements, and impact on vaccination motivation/intent • Mis/disinformation about vaccination requirements, emergency use authorizations, and licensing of vaccines
Emerging	COVID-19 vaccines resulting in improvement in other illnesses/co-morbidities
Continuing & Evolving	VAERS misrepresentation (e.g., claims of increased vaccination-related death)
	Concerns about length of protection, booster doses, and guidance for immunocompromised people
	Vaccination Administration issues (e.g., errors in handling and administering vaccines, guidance)
	Impacts of misinformation - Information seeking behaviors, social media behaviors, impact on one's plan to get vaccinated.
	COVID-19 virus origins – claims COVID-19 was in US prior to reports but covered up.

Step 5: Consensus-building meeting to confirm proposed outline.

Step 6: Analysts pull data and perform secondary search based on proposed outline.

Step 7: Team leads draft and review report with team and determine threat classification.

Step 8: Consensus-building meeting to confirm final report outline.

Category	Theme	Classification
Major	Consumers are confused and concerned about the effectiveness of currently available vaccines against the Delta variant.	High risk, increasing
	Some consumers feel that vaccine mandates and proof-of-vaccination systems infringe on individual liberties.	Moderate risk, stable
Continuing & Evolving	Adverse events after vaccination.	
	Concerns from people with compromised immune systems.	
	Origin of virus that causes COVID-19.	

Report Production

Drafting Theme Narratives

Unit lead(s) drafted a narrative for each theme that described consumer questions, concerns, frustrations, information voids, message penetration issues, and misunderstanding of science or guidance based on their analysis of the pulled data. Narratives also highlighted circulating mis- and disinformation and clarified which demographics appeared to be more affected by that theme if it could be discerned. Each theme's narrative included links to illustrative examples, the latest literature, and recent polls to provide real-world context to each theme and was crafted through the lens of our analytical frameworks to connect the theme to known behavioral impacts on vaccine confidence, demand, and uptake. Themes determined to be "main themes" had the most substantial narratives, whereas continuing and evolving narratives were limited to a few sentences, focusing on the changes from when that theme was previously featured.

Each theme's narrative concluded with "ways to take action," providing concrete recommendations for how communicators, community leaders, healthcare providers, and health authorities could improve COVID-19 vaccine confidence, demand, and uptake based on evidence-informed practices in behavioral science and communication theory. The section typically included communication, programmatic, and engagement opportunities. The communication actions provided suggestions for filling information voids, correcting message penetration issues through reframing, and engaging specific trusted messengers to disseminate and amplify messages. The programmatic suggestions identified how new or existing programs, processes, or systems could change or be engaged. Finally, research suggestions clarified what research gaps and unanswered questions could be further explored.

When the Insights Unit had two leads, each lead was responsible for interpreting and drafting distinct theme narratives. After they completed the process described above for their assigned themes, they met to discuss and review their findings and proposed actions. When the Insights Unit had one lead, the unit lead drafted all the theme narratives and met with an analyst with the most experience to participate in a similar feedback loop before meeting with the larger unit. After a report was fully drafted, a third consensus-building meeting focused on reviewing the narratives with the entire team to ensure content integrity.

Applying the Threat Matrix

Following confirmation of each theme's narrative with the unit, the lead(s) facilitated a discussion to determine the threat level and directionality of each main and emerging theme. Using the theme matrix, the team discussed the level of risk to vaccine confidence (i.e., whether the theme would directly or indirectly lead to vaccine refusals and decreased uptake, which groups of people would have their confidence and uptake most affected by the theme, and the volume of spread the theme appeared to have across data segments and within different groups of people). The team came to a consensus on the level of potential risk to vaccine confidence and uptake for each theme represented: high, moderate, low, or positive. After the risk level was confirmed, we discussed the directionality of the main themes to determine if the themes were *increasing*, *stable*, or *decreasing* in volume.

All emerging themes were classified as increasing due to the nature of the category. The analysts identified the directionality of a theme based on the data from their individual data segments when compared to previous reporting periods. If there was no consensus between the data segments on the directionality of a theme, then we determined the directionality for each data segment and the overall trend across all data segments. For example, if a theme decreased slightly in frequency in the news but increased significantly on social media and CDC-INFO, then we labeled the theme as increasing. Themes that remained at the same volume and spread but had been previously overshadowed by

high-risk or volume themes and developments across multiple reporting periods relative to each data segment's baseline were classified as stable in directionality.

Assembling the Report

Once all themes were classified through consensus, the order of the report outline was confirmed, with the main theme posing the highest threat to vaccine confidence, demand, and uptake listed first. Subsequently, a written executive summary provided an overview of the main themes identified in the report and the primary ways to act.

Ethical Considerations

The methodology was determined to be nonhuman subjects research through CDC's ethics and project approval process. Additional ethical approval was not required as the data being considered was deidentified during the analysis step. Primary data sets, such as social listening data collected via Meltwater, were comprised of publicly available data falling under the public domain. Secondary data sets, such as survey data, did not contain any personal identifiable information and data collection and analysis for those data sets underwent their own independent ethical approval processes.

Quality Improvement

Process Review

The urgency of COVID-19 pandemic and accompanying vaccine confidence issues necessitated creating reports as the methodology was being built. Our methodology underwent continuous process improvement as workforce capacity and contextual changes ebbed and flowed while analysts and unit leads innovated with every report edition. The report also underwent several distinct evaluation efforts to increase the rigor of the research process, the validity of the findings, and the usability of the reports.

For example, in June 2021, to identify and reduce threats to the validity of the report's findings, the Insights Unit team undertook a formal process evaluation, outlining the methods used for identifying and classifying themes, and assessed the rigor of data collection and analysis. The team evaluated how data was collected, themes were categorized, the prevalence of data was determined, and how they determined the threat to vaccine confidence. An evaluation expert interviewed past and current team members, conducted concept mapping exercises with unit leads, observed a full cycle of a report, and reviewed standard operating procedures documents and literature. Through this evaluation, the team was able to identify and articulate the process they had developed to date, using established team-based applied qualitative data collection and analysis techniques and best practices. The evaluation also resulted in several changes to the data collection, analysis, and interpretation of future reports, including:

1. Developing a codebook (See *Table E: COVID-19 State of Vaccine Confidence Insights Report Code Book*) and onboarding training to ensure team members shared a common framework for understanding the data they encountered.
2. Expansion of preliminary theme submissions for each data source to include relevant quantitative metrics related to prevalence of theme within that data source.
3. Reducing reliance on third-party reports by utilizing these reports to triangulate findings from raw data inputs rather than analyzing third-party reports as raw data input.
4. Requiring team members to update their data segment's standard operating procedure at the end of deployment to ensure their process innovations were passed on to their replacements.

Some evolutions of our methodology were not a result of this evaluation, but rather contextual shifts. Data sources included and analyzed for each report varied over time largely because of the increased capacity of the Insights Unit due to more analysts joining the unit over time. This allowed for more person-time to be dedicated to the analysis of primary data sources, which in turn led the team to identify additional primary sources to utilize, such as Meltwater and SEMRush. Additionally, some of the secondary data sources, mostly insights reports produced by external agencies and organizations, halted their reports over time. This development and the Unit's ongoing use of continuous-quality improvement principles decreased the necessity of relying on third-party reports. The process also led to the creation of a more substantial on-boarding process to ensure consistent application of our qualitative and quantitative analysis methods and deductive and inductive approach as outlined in relevant standard operating procedures.

Expert Review

Additionally, in July 2021, an external subject matter expert with a background in infodemics, fact-checking, and digital media, was contracted to conduct a strengths, weaknesses, opportunities, and threats (SWOT) analysis of the Insights Units and its reports. The expert conducted in-depth interviews with leads and analysts and observed the process of producing a report from inception to publication. The analysis found that the methodology used and subsequent reports were unique in terms of diverse subject matter expertise of the unit membership, speed and regularity, constant reflection, and scale (i.e., considering the whole information ecosystem, including listening for unanswered questions and topics leading to confusion, as well as mis- and disinformation, the use both of many different data sources and data sources unique to CDC such as CDC-INFO, etc.). However, the analysis also found that the reports lacked a clear objective, as they attempted to serve as an early warning system for misinformation, a rapid response system to help manage official communications, and/or a means to identify content gaps. The reports also had a wide target audience, and mechanisms to ensure intercoder reliability, as it relied on a "lone wolf" (i.e., single coder) coding strategy through unit lead(s). The review found additional weaknesses to address, such as the need to clarify definitions and the codebook so that it was uniformly applied, and the challenges of comparing data from disparate sources with widely varying denominators. Steps were taken in the second year of the Insights Unit to ensure the code book was uniformly applied through enhanced onboarding training for Insights Unit deployers and staff while expanding the "lone wolf" coding strategy by increasing the number of staff trained.

Results

Production and Distribution

As of August 2022, the Insights Unit has produced 27 full reports, five rapid reports, and two special reports. Rapid and special reports employed the same methodology applied to full reports except with a shorter data collection period. Rapid reports pertain to a specific and time-sensitive issue, such as the Emergency Use Authorization of COVID-19 vaccines in adolescents.

SoVC reports were distributed by email to a listserv of 956 readers, including CDC staff and external partners (e.g., federal agencies, professional associations, community-based organizations, and non-profits, in addition to the health authorities and government representatives from around the world). Reports were also published online and have been downloaded 22,361 times. Findings and action steps were also shared through presentations to local, national, and international stakeholders from March 2021 to December 2021.

Utilization

Within the agency, the reports were used by VTF and the emergency operation center to directly inform strategies for communication and community engagement. VTF's communication team used the reports to enhance CDC's COVID-19 vaccine social media content, update digital content on myths and frequently asked questions, and improve search engine optimization. As the reports highlighted challenges faced by specific demographics, VTF staff used the findings to support partner organizations and states and jurisdiction efforts for on-the-ground vaccine demand generation. The reports also served as a unique mechanism for providing feedback to VTF's vaccine distribution teams, by highlighting underlying practical issues impacting vaccine uptake, such as challenges navigating digital systems to schedule vaccine appointments.

Between March 2021 and August 2022, SoVC reports were referenced by groups outside of CDC 22 times online. Organizations and agencies who referenced the reports included state and local health departments (n= 6), non-profit organizations (n= 5), professional associations (n= 3), and congressional committees (n= 2). The reports were most referenced as a resource for learning about COVID-19 vaccine confidence and demand, with nearly half of the identified references (n= 11) linking directly to reports. Many of these references focused on increasing uptake of the COVID-19 vaccine, sharing the latest information on COVID-19 vaccinations, and improving public health communications. SoVC reports were also cited by research studies (n= 4) to justify study goals and findings. For example, one journal article used a report as context to describe how perceptions of vaccine effectiveness can impact vaccine confidence [16-19].

In addition to serving as an informative resource, the reports were also used to demonstrate how CDC monitored and responded to COVID-19 vaccine mis- and disinformation. Then CDC director, Dr. Rochelle Walensky, referenced the reports and the Insights Unit at two congressional hearings, describing how CDC is working to understand the public's perceptions of COVID-19 vaccines and address the public's information needs to increase vaccine uptake [20]. Furthermore, we identified two sources that applied findings and themes from the reports, one of which was the African Infodemic Response Alliance, which issued a report in June 2021 on COVID-19 infodemic trends, containing recommendations adapted from an SoVC report.

Evaluation

To better understand the consumers of the reports, the value they derive from them, and how to further improve their utility, we conducted an outcome evaluation in October 2021 via a survey of the audience. We received 58 survey responses (7% response rate from the email distribution list) and solicited feedback from six stakeholders and one focus group session with four Insights Unit staff. We found that nearly 80% of respondents said they found the reports "useful" or "extremely useful," and 61% of respondents used the reports to inform communication strategies. The respondents also made suggestions for improvements (e.g., creating bulleted summaries of findings, providing resources for readers to operationalize reports' recommendations, etc.), which led to several changes to make the report more readable and useful. Subsequent reports, beginning in early 2022, incorporated bulleted summaries of themes rather than descriptive narratives and resources. These were provided within the reports themselves and links to additional resources for communicating about the themes were also provided.

Discussion

Our approach and methodology intentionally responded to urgent, interconnected needs. The primary priority was to assist the agency and our partners in acting more swiftly to address the questions,

concerns, perceptions, information voids, and circulating mis- and disinformation around COVID-19 vaccines and inform communication and programmatic action. The second priority was to establish a coordinated mechanism for collecting, reviewing, and synthesizing qualitative and quantitative data from multiple internal and external sources. Our unit and reports were intended to bridge the gap between CDC's current social listening efforts, reports produced by external organizations, and current surveys and polls conducted by CDC and third parties. Being one of the first integrated infodemic management products of such scale, scope, and regularity to be produced at the CDC, these reports interpreted findings in a novel way through the lens of behavioral theory and public health best practices. Similar published approaches and methodologies often focus on singular data streams and typically do not utilize a deductive approach to provide context, understand, and use the implications of findings.

While our methodology was developed specifically in the context of understanding vaccine confidence and demand, and the information ecosystem around COVID-19 vaccines in the United States, it can be adapted for any public health issue. Further, the varied use of our reports suggests that they met the information needs of a diverse audience, served an informative purpose, and conveyed an urgency to address information challenges, implying that similar reports on other public health topics would support both routine and emergency functions. However, this methodology should not be directly reproduced without contextualization. It is important to conduct needs assessments and landscape analyses to understand the current state of the topic at hand and the population that is going to be examined, in addition to identifying data sources for consideration. Irrespective of these contextualization needs, our integrated analysis and the consensus-building processes are replicable, and others could create similar reports. Additionally, the reports were produced by a rotating staff with different backgrounds and changing capacities, which implies that the analysis process and techniques used can be learned and scaled up or down depending on the skills and bandwidth of the staff.

Limitations

There were several limitations of this methodology both for generating insights themselves, but also in its utility.

Data and Granularity

A limitation for SOVC reports themselves was the data sources utilized and their inherent limitations. As widely documented, data gleaned from social media may present a skewed picture of the population of interest due to a lack of representativeness. For instance, people of certain age groups, languages, or geographic locations may be over- or under-represented in social media data depending on how digitally connected they are [9,21,22]. Demographic biases in the data are also likely further pronounced by the limited availability of data from closed social media (e.g., private groups in WhatsApp, Telegram, or Facebook) or dominance of more easily accessible data (e.g., Twitter conversations) in our analyses [3,23]. Further, demographic data is not always linked to user-generated content on social media [21]. Sophisticated artificial intelligence and algorithms would be needed to deduce user demographics based on profile content generation and digital footprints, which could present ethical challenges. Additionally, users may present themselves differently in online spaces than in face-to-face interactions; therefore, their content may not always be accurate proxies for their true intentions, perceptions, or behaviors [9]. This makes it difficult to detect local contexts or nuances in large data sets compiled for national-level analysis [3]. For example, data aggregator tools such as Google Trends exclude smaller trends (i.e., searches made by very few people) [24]. While we have tried to minimize the impact of some of these limitations by integrating multiple data sources beyond aggregated social media or web analytics data, they remain important

to bear in mind when interpreting findings. One potential way to mitigate this issue is to conduct a comprehensive analysis of all data sources prior to inclusion to better understand the demographics being sampled and to ensure that the data sources are accurately reflecting the demographics of the population of focus. Another potential tactic would be to establish thresholds for specific key terms for each data source to help identify when mentions are above baseline for each topic. Research should also be done to explore what type of infodemic management is being done at subnational and local levels to identify emerging best practices to test and validate.

Potential Bias and Subjectivity

Social listening and analysis tools, especially focused on the digital space, have primarily been built for the purposes of marketing and external communication campaign tracking. Adapting such tools for use in the public health space can be challenging and, in the case of our methodology, we had to adapt tools to meet our analytical needs. Additionally, there are no systematic, near real-time approaches to consider the risk of a qualitatively identified theme in the context of a behavioral health model. This was further complicated by the lack of indicators to directly link exposure to information ecosystems to vaccine refusal. We attempted to mitigate the impacts of subjectivity through consistent training which involved overlap between deployers to ensure adequate knowledge transfer and making decisions based on consensus. While consensus-based decision making is not perfect, it did reduce the potential impacts of individual bias on assessing the threat of themes to vaccine confidence. Those considering adapting these methods for their own use should consider how they could further reduce potential bias and ensure that theoretical frameworks are deployed. Further, artificial intelligence (AI) and machine learning could potentially be leveraged to compliment the human component of our process to reduce bias and subjectivity of analysis by performing the initial analysis of raw data sources.

Adaptability

The methodology itself may also be limited in its ability to be operationalized in different public health settings. First, the scale of the unit and process described in this manuscript required several dedicated staff and significant personnel time. This may not be accessible or replicable outside of an emergency response. Research needs to be conducted to better determine ways to adapt this process for routine use. Further, the process may not be achievable exactly as described for most health authorities until tools are developed and deployed to reduce the amount of personnel time required to generate insights. Different iterations of this methodology should be designed, tested, and evaluated in different public health settings to determine minimal and optimal staffing, funding, and mechanism. Researchers should consider how to leverage artificial intelligence and related technology in the collection and analysis of these types of activities and reports.

Impact and Outcomes

Additional evaluation is also needed to better understand the utility of rapid or deep dive reports compared to standard reports and how the different types of reports could be utilized more impactfully. Additionally, because of how the unit was established, we did not have decision-making authority in acting based on the reports' findings and recommendations. Units such as this should be more intrinsically linked to program and mechanisms should be put into place to maximize the utilization of the reports.

The impact of these reports on program and campaign activities leading to improvement of vaccine coverage also needs to be assessed. Limited indicators exist to directly link exposure to information ecosystems or communication campaigns to vaccine uptake and coverage. The guiding analytical frameworks we used do not currently directly include information ecosystems and campaigns in their

models. Further, most socio-behavioral measures related to vaccination typically measure intent to vaccinate but cannot directly link to whether behavior was completed. Identifying indicators that more directly link action with exposure to infodemics, communication campaigns, and interventions is necessary to better understand their impact behaviors as well as establish best practices. Indicators outside of public health should be considered for adaptation, tested, and validated in the field.

Conclusion

The Insights Unit's methodology is not a solution for infodemics, but rather a diagnostic technique that can rapidly identify opportunities for intervention. In addition, identifying mechanisms to act on the findings need to be created at the onset to ensure success. This methodology is also an adaptable process that can be scaled and used in a variety of public health settings for all public health issues. It can be applied beyond acute public health crises, such as within routine monitoring and surveillance. The methodology as outlined provides a roadmap for organizations focused on improving public health to conduct an integrated analysis of disparate data sources to generate insights and leverage data insights during outbreaks and routine programming to inform population-level health. The insights generated can inform the deployment of effective interventions to fill information voids and counter misinformation, including partnering with trusted messengers.

The results of integrated analyses can be considered within already established, evidence-based conceptual frameworks and leveraged to deliver timely and responsive communication, programming, and policy. The continual listening and feedback process of Insights Units such as the one detailed here allows for health authorities to clarify guidance and improve programmatic quality in near real-time, based on changing public needs, especially within the context of a rapidly evolving public health emergency. Institutionalizing such units could enable public health messages, guidance, programs, and policies to better resonate and meet the information needs of the public. More responsive public health action could build trust among public health authorities and the public, which could lead to higher sustained uptake of health promotion strategies.

Public health authorities are especially well-positioned to utilize their staff, including behavioral and social scientists, communication scientists, informatics specialists, ethicists, and epidemiologists, to institute Insights Units for population health promotion. These units can be on the frontlines of not only identifying and diagnosing infodemics, but also of building and promoting the science to test, evaluate, and scale promising interventions to rapidly respond to community needs, counter misinformation, bolster resiliency, and reinforce trust [25]. These units could benefit from being closely connected to traditional public health functions, both within an institution and with external partners, including those focused on public health surveillance, field epidemiology, and health communications, to further contextualize and validate data trends [26,27]. Such collaborations allow health authorities to work more proactively.

Future Directions

The methodology and work of the Insight Units and similar units are critical to the burgeoning field of infodemic management. This methodology has centered on data within the United States, a high-income and predominantly English-speaking country; these methods should be extended to better incorporate non-English language digital listening in the future because information environments are not limited by language or geographic borders. The methodology presented here provided a national view of vaccine confidence and demand issues; localized sub-national applications will

require improved digital listening tools and investigation to determine the best online and offline sources from which to collect data. Additionally, the development of new technology and platforms at low cost is essential to moving this work beyond high-income countries that can adapt current tools to perform these social listening and analysis functions. Artificial intelligence and machine learning could help reduce the burden of data collection and preliminary analysis with promising approaches already being tested, such as using machine learning to help identify bots and measure infodemic risk [28]. Finally, more research needs to be conducted to better understand the impact that infodemics and infodemic management have on health behaviors. Indicators should be explored and examined to allow for quantifiable measurement of the burden of infodemics which is turn will allow for improved evaluation of interventions and infodemic management itself.

Acknowledgments

We thank all members of the Insights Unit for their contributions to this manuscript and our methodology. We also thank the leadership of the COVID-19 Vaccine Confidence and Demand team and COVID-19 Vaccine Task Force for their support and championing our work, especially Neetu Abad, James Lee, David Fitter, Amanda Cohn, Demetre Daskalakis, and Dimitri Prybylski. We want to express our gratitude to our colleagues who were responsible for coordinating, reviewing, formatting, and disseminating our reports for their patience and support as we navigated a novel, near-real-time process. We are grateful to our CDC colleagues in the Global Health Center, National Center for Immunization and Respiratory Diseases, Office of Readiness and Response, and Office of Communication for partnering with us to identify and utilize the diverse sources our methodology relied on and turning our insights into action, especially Keri Lubell, Carol Crawford, Rosemary Bretthauer-Mueller, Kristen Nordlund, Kathleen Laporte Conley, James Singleton, and Richard Quartarone. Lastly, we thank our stakeholders who used our reports and worked tirelessly to improve confidence in and demand for COVID-19 vaccines and combating health misinformation broadly, especially the Office of the Surgeon General, #ThisIsOurShot, and all the states and jurisdictions we supported and worked with.

Conflicts of Interest

JK, KB, VC, CV, SJ, KR, HT, HF, IK, and JH are staff of the U.S. Centers for Disease Control and Prevention. AI is an employee of Global Health Task Force and supports the Demand for Immunization team through a contract. EW is a visiting fellow at Brown University in the Information Futures Lab. AV is an employee of Tanaq Support Services and supports the Insights Unit through a contract. These authors alone are responsible for the views expressed in this paper, and they do not represent the views of their organizations.

Abbreviations

BeSD: Behavioral and Social Drivers for the Uptake of Vaccines

CDC: Centers for Disease Control and Prevention

GHC: Global Health Center

ORR: Office of Readiness and Response

COVID-19: Coronavirus disease

NCIRD: National Center for Immunization and Respiratory Diseases

OC: Office of Communication

SoVC: State of Vaccine Confidence

WHO: World Health Organization

References

1. World Health Organization. Infodemic. Accessed January 3, 2023. https://www.who.int/health-topics/infodemic
2. Donovan J. Concrete Recommendations for Cutting Through Misinformation During the COVID-19 Pandemic. *American Journal of Public Health*. 2020;110(S3):S286-S287. doi:10.2105/ajph.2020.305922
3. Purnat TD, Vacca P, Czerniak C, et al. Infodemic Signal Detection During the COVID-19 Pandemic: Development of a Methodology for Identifying Potential Information Voids in Online Conversations. *JMIR Infodemiology*. 2021;1(1):e30971. doi:10.2196/30971
4. Briand SC, Cinelli M, Nguyen T, et al. Infodemics: A new challenge for public health. *Cell*. 2021;184(25):6010-6014. doi:10.1016/j.cell.2021.10.031
5. Islam MS, Sarkar T, Khan SH, et al. COVID-19–Related Infodemic and Its Impact on Public Health: A Global Social Media Analysis. *The American Journal of Tropical Medicine and Hygiene*. 2020;103(4):1621-1629. doi:10.4269/ajtmh.20-0812
6. Nguyen KH, Srivastav A, Razzaghi H, et al. COVID-19 vaccination intent, perceptions, and reasons for not vaccinating among groups prioritized for early vaccination — United States, September and December 2020. *American Journal of Transplantation*. 2021;21(4):1650-1656. doi:10.1111/ajt.16560
7. Hamel LK, Ashley , Muñana C, Brodie M. KFF COVID-19 Vaccine Monitor: December 2020. Kaiser Family Foundation. Updated December 15. https://www.kff.org/coronavirus-covid-19/report/kff-covid-19-vaccine-monitor-december-2020/
8. Hamel L, Kirzinger A, Lopes L, Kearney A, Sparks G, Brodie M. KFF COVID-19 Vaccine Monitor: January 2021. Kaiser Family Foundation. https://www.kff.org/coronavirus-covid-19/report/kff-covid-19-vaccine-monitor-january-2021/
9. Hou Z, Tong Y, Du F, et al. Assessing COVID-19 Vaccine Hesitancy, Confidence, and Public Engagement: A Global Social Listening Study. *Journal of Medical Internet Research*. 2021;23(6):e27632. doi:10.2196/27632
10. Braun V, & Clarke, V. Thematic Analysis. In: H. Cooper PMC, D. L. Long, A. T. Panter, D. Rindskopf, & K. J. Sher, ed. *APA Handbook of Research Methods in Psychology, Vol 2: Research Designs: Quantitative, Qualitative, Neuropsychological, and Biological*. American Psychological Association; 2012:57-71.
11. Guest G MK, Namey EE. *Applied Thematic Analysis* SAGE Publications, Inc.; 2012.
12. Prevention CfDCa. Vaccinate with confidence: strategy to reinforce confidence in COVID-19 vaccines. US Department of Health and Human Services, CDC. 2021.
13. BeSD Working Group. Based on , Brewer NT, Chapman GB, Rothman AJ, Leask J, Kempe A. Increasing Vaccination: Putting Psychological Science Into Action. *Psychological Science in the Public Interest*. 2017;18(3):149-207. doi:10.1177/1529100618760521
14. United Nations Children's Fund Vaccine Misinformation Management Field Guide. New York, NY2020.
15. United Kingdom Government Communication Service. *RESIST Counter Disinformation Toolkit*. 2019. https://gcs.civilservice.gov.uk/publications/resist-counter-disinformation-

toolkit/
16. Kittitas County Public Health Department. 2021 Kittitas County COVID-19 Vaccine Community Survey Results Report. 2022.
17. Rosen AD, Beltran J, Thomas E, et al. COVID-19 Vaccine Acceptability and Financial Incentives among Unhoused People in Los Angeles County: a Three-Stage Field Survey. *Journal of Urban Health*. 2022;99(3):594-602. doi:10.1007/s11524-022-00659-x
18. Janzen S, Orr C, Terp S-J. Cognitive Security and Resilience: A Social Ecological Model of Disinformation and other Harms with Applications to COVID-19 Vaccine Information Behaviors. 2nd Workshop Reducing Online Misinformation through Credible Information Retrieval, ROMCIR 2022; 2022.
19. Federal Emergency Management Agency. Vaccine Hesitancy & Approach to Action An Anthropological Study in Southern Colorado. 2021.
20. *An Update on the Ongoing Federal Response to COVID-19: Current Status and Future Planning* US Senate Committee on Health, Education, Labor & Pensions, Testimony of Rochelle Walensky sess (2022). https://www.help.senate.gov/imo/media/doc/Walensky8.pdf
21. Scott I. (Mis)communication? Social listening and the exclusion of marginalised voices. *The Humanitarian Leader*. 2022;doi:10.21153/thl2022art1558
22. Jang H, Rempel E, Roe I, Adu P, Carenini G, Janjua NZ. Tracking Public Attitudes Toward COVID-19 Vaccination on Tweets in Canada: Using Aspect-Based Sentiment Analysis. *Journal of Medical Internet Research*. 2022;24(3):e35016. doi:10.2196/35016
23. Spitale G, Biller-Andorno N, Germani F. Concerns Around Opposition to the Green Pass in Italy: Social Listening Analysis by Using a Mixed Methods Approach. *Journal of Medical Internet Research*. 2022;24(2):e34385. doi:10.2196/34385
24. Google News Initiative. Google News Initiative Training Center. Accessed September 19, 2022. https://newsinitiative.withgoogle.com/training/lesson/4876819719258112?image=trends&tool=Google%20Trends
25. Chou WS, Gaysynsky A, Vanderpool RC. The COVID-19 Misinfodemic: Moving Beyond Fact-Checking. *Health Educ Behav*. Feb 2021;48(1):9-13. doi:10.1177/1090198120980675
26. Badell-Grau RA, Cuff JP, Kelly BP, Waller-Evans H, Lloyd-Evans E. Investigating the Prevalence of Reactive Online Searching in the COVID-19 Pandemic: Infoveillance Study. *Journal of Medical Internet Research*. 2020;22(10):e19791. doi:10.2196/19791
27. Tangcharoensathien V, Calleja N, Nguyen T, et al. Framework for Managing the COVID-19 Infodemic: Methods and Results of an Online, Crowdsourced WHO Technical Consultation. *Journal of Medical Internet Research*. 2020;22(6):e19659. doi:10.2196/19659
28. C. N. Hang, P. -D. Yu, S. Chen, C. W. Tan and G. Chen, "MEGA: Machine Learning-Enhanced Graph Analytics for Infodemic Risk Management," in IEEE Journal of Biomedical and Health Informatics, vol. 27, no. 12, pp. 6100-6111, Dec. 2023, doi: 10.1109/JBHI.2023.3314632.

Supplementary Files

Untitled.
URL: http://asset.jmir.pub/assets/47e7d517574f558f65c4f5701d8bcdd4.docx

Untitled.

COVID-19 State of Vaccine Confidence Insights Report

Report 11 | July 7, 2021 | Date Range: June 8 – 21, 2021

Summary

Findings. Consumers expressed concern about the Delta variant of the virus that causes COVID-19 potentially causing more severe illness. While this was a motivating factor for some consumers to get vaccinated, others felt the variants were a sign of the virus becoming endemic, decreasing the perceived importance of vaccination. As the United States continues to return to prepandemic-style life, some consumers expressed concern that employer- or education-based vaccination requirements and proof-of-vaccination systems to attend large events or travel infringed on their individual liberties. Consumers were divided, with some claiming that that those who are vaccinated do not need to worry about the vaccination status of others. Finally, concerns about myocarditis and pericarditis following mRNA COVID-19 vaccination were widespread. More information about consumers' perceptions and opinions regarding myocarditis and pericarditis following vaccination can be reviewed in the Rapid COVID-19 State of Vaccine Confidence Insights Report, released on June 29, 2021.

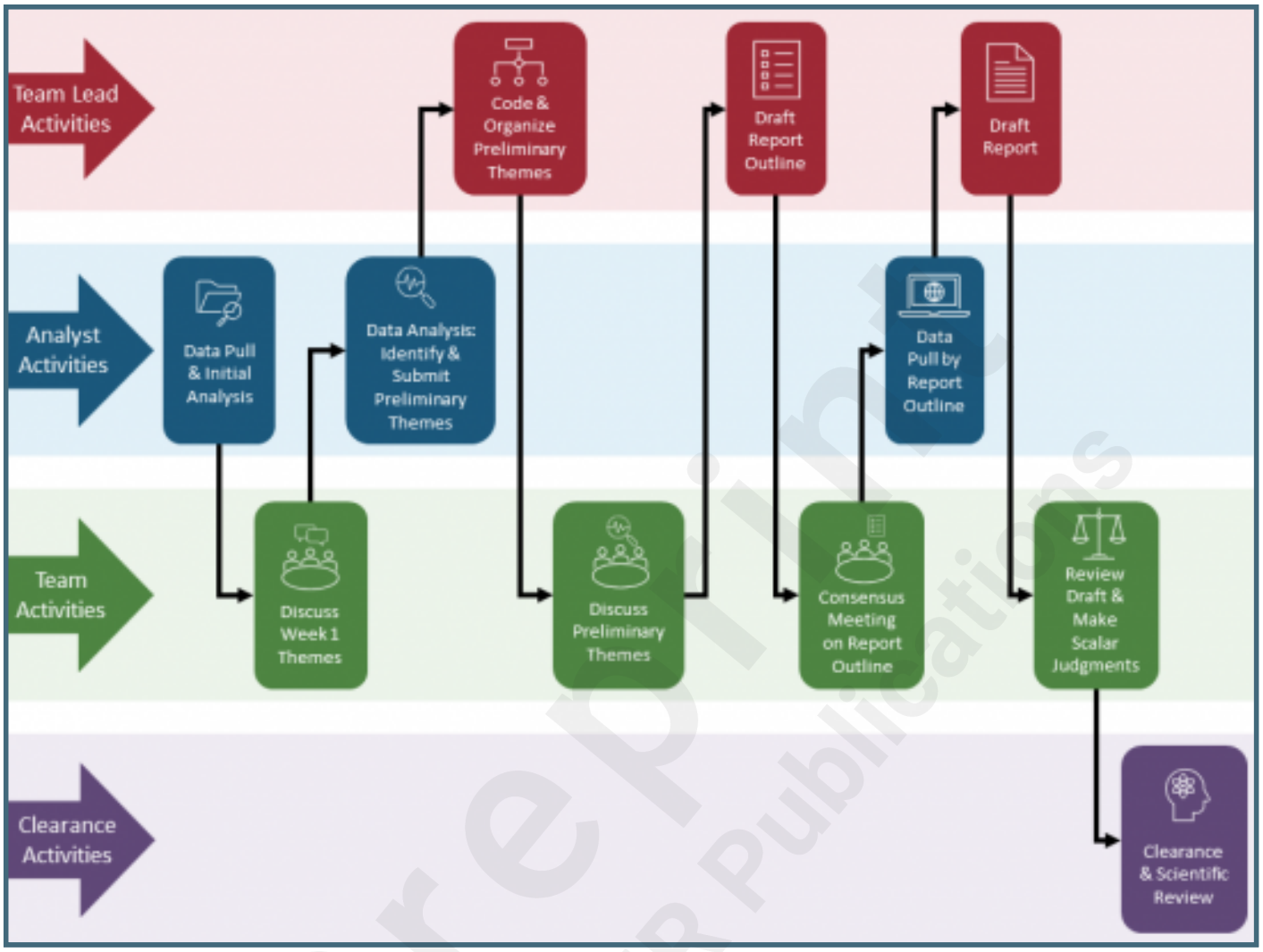
Ways to take action. Federal, state, and local partners should continue to work together to increase transparency around rationale for updated guidance, respond to gaps in information, and confront misinformation with evidence-based messaging. The goal of these efforts is to increase confidence in COVID-19 vaccines and expand vaccine uptake more broadly. Messages should be developed and disseminated to provide known information about the effectiveness of available COVID-19 vaccines against emerging variants, as well as the role vaccination plays in reducing the spread and emergence of new variant viruses. Policies regarding travel and large gatherings should be evaluated to identify opportunities to reduce requirements for fully vaccinated people that could motivate consumers within the moveable middle to get vaccinated.

Contents

Centers for Disease Control & Prevention, COVID-19 Response, Vaccine Task Force
Vaccine Confidence & Demand Team, Insights Unit

Figures

Visualization of report production.



Example from SoVC Report #11: Process from Data Segment Analysis to Final Outline.

Step 1: Analysts report preliminary themes by data segment.

Data Segment	External Social Media (e.g., Twitter, Reddit, etc.)	Internal Social Media (e.g., CDC-owned social media channels)	Third Party Reports (e.g. Communication Surveillance Report)	News Media	Polls & Literature	Web Metrics (e.g., Google Trends)
# Themes	6	6	9	39	15	8

Step 2: Preliminary themes coded by team leads.

Top Codes (≥ 5 prelim. themes)	Vaccines/ Safety	Outside vaccine system/ Requirements	Data/ Data Issues	Vaccine System/ Guidance for fully vaccinated	Data/ Vaccine Uptake	Vaccines/ Effectiveness	Outside vaccine system/ COVID-19 restrictions	Data/ Confidence Data
# Themes	13	12	10	8	7	7	5	5

Step 3: Consensus-building meeting to develop outline.

Step 4: Proposed report outline developed by team leads.

Category	Theme
Major	Spread of Delta variant in US causing concern about vaccine effectiveness • Coverage of/discussions about delta variant • Concerns and questions about breakthrough cases, booster doses, asymptomatic spread and transmission among fully vaccinated people, and effectiveness of specific vaccines • Mis/disinformation about vaccines causing the virus variants
	Concerns vaccine mandates infringe on individual liberties and medical freedom. • Coverage of/discussions about vaccine requirement policies (e.g., employers, travel, etc.) and plans, vaccine-related legislation, mask requirements, and vaccine status/privacy • Consumer concerns and questions about legality of vaccine requirements, requirements of vaccines under emergency use authorization, exemptions and "right to opt out," and tracking vaccination status and breakthrough cases • Polls & Research about opinions on mandates/requirements, societal impacts of requirements, and impact on vaccination motivation/intent • Mis/disinformation about vaccination requirements, emergency use authorizations, and licensing of vaccines
Emerging	COVID-19 vaccines resulting in improvement in other illnesses/co-morbidities
Continuing & Evolving	VAERS misrepresentation (e.g., claims of increased vaccination-related death)
	Concerns about length of protection, booster doses, and guidance for immunocompromised people
	Vaccination Administration issues (e.g., errors in handling and administering vaccines, guidance)
	Impacts of misinformation - information seeking behaviors, social media behaviors, impact on one's plan to get vaccinated.
	COVID-19 virus origins – claims COVID-19 was in US prior to reports but covered up.

Step 5: Consensus-building meeting to confirm proposed outline.

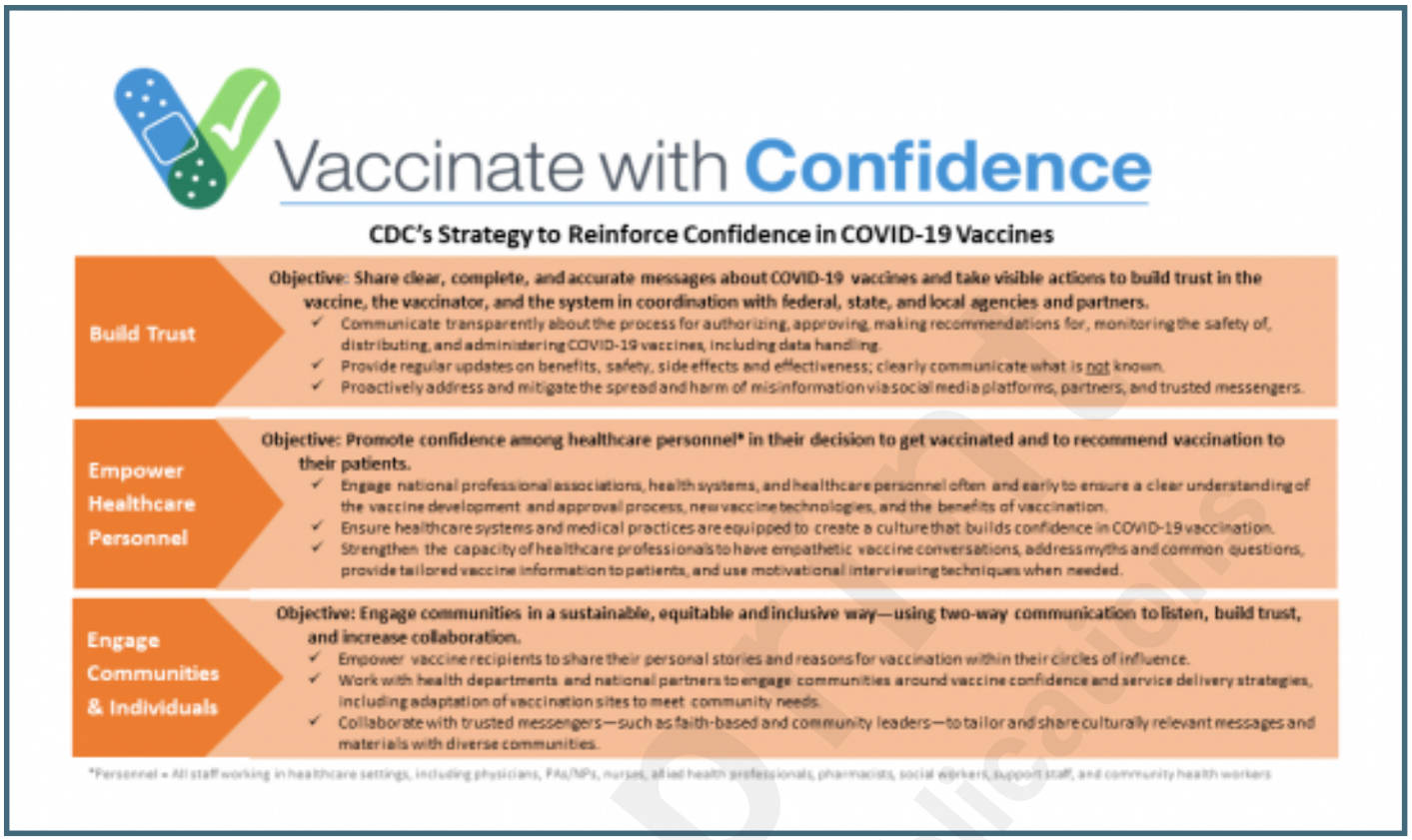
Step 6: Analysts pull data and perform secondary search based on proposed outline.

Step 7: Team leads draft and review report with team and determine threat classification.

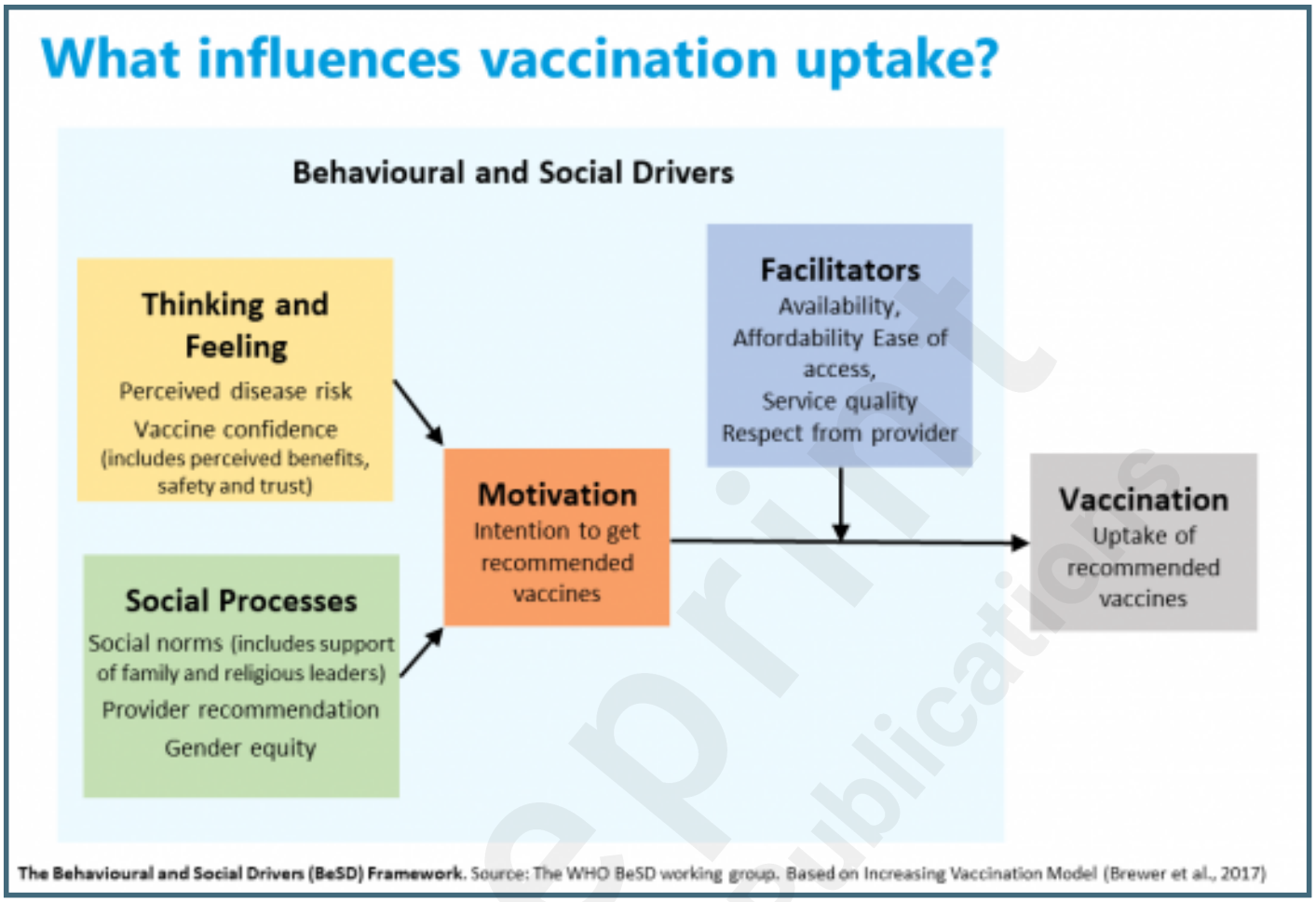
Step 8: Consensus-building meeting to confirm final report outline.

Category	Theme	Classification
Major	Consumers are confused and concerned about the effectiveness of currently available vaccines against the Delta variant.	High risk, increasing
	Some consumers feel that vaccine mandates and proof-of-vaccination systems infringe on individual liberties.	Moderate risk, stable
Continuing & Evolving	Adverse events after vaccination.	
	Concerns from people with compromised immune systems.	
	Origin of virus that causes COVID-19.	

CDC Vaccinate with Confidence Strategy for COVID-19 vaccines.



The Behavioral and Social Drivers (BeSD) Framework.



COVID-19 State of Vaccine Confidence Insights Report Threat Matrix.