NEST
Coordinating Center
An initiative of MDIC

Active Surveillance Roadmap

A report produced by the NEST Coordinating Center, an initiative of MDIC

August 12, 2021

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1 EXECUTIVE SUMMARY

The National Evaluation System for health Technology (NESTcc) Active Surveillance Roadmap reports on progress made and seeks public comment on NESTcc's work to establish capabilities and tools to support active surveillance of medical devices. As part of the U.S. Food and Drug Administration's (FDA) 2018 Medical Device Safety Action Plan (Gottlieb & Shuren, 2018), NESTcc was tasked with developing, testing and operationalizing solutions to enable ongoing, prospective, active monitoring of postmarket medical device performance based on real-world data (RWD). Key priorities for the work include: delivering a national system to advance safety signal detection (generation) and refinement; establishing procedures to ensure data quality and reliability; building technical services in secure, scalable and modular environments to accommodate the addition of future innovative methodologies; developing applications beyond postmarket safety; and designing business practices and controls to encourage use by diverse constituents of the device ecosystem.

This Roadmap includes a brief background of NESTcc and an overview of the strategy and tactics used to develop active surveillance services by creating of the NESTcc Cloud, including descriptions of the current architecture, schedules for the phased-build and management based on the U.S. Department of Health and Human Services' (HHS) Enterprise Performance Life Cycle (EPLC) framework. The document also outlines NESTcc analytic methodologies and procedural technical services currently under development to ensure data quality and appropriate signal communication.

2 BACKGROUND

2.1 About NESTcc

The Medical Device Innovation Consortium (MDIC) established NESTcc in 2016 with funds awarded by FDA through a cooperative agreement. NESTcc is tasked with driving quality and efficiency in the use of RWD to inform medical device evaluation, as well as support clinical, patient, regulatory and reimbursement decisions throughout the total product life cycle (TPLC). NESTcc is currently engaged in pioneering work to develop an Active Surveillance capability for medical device safety.

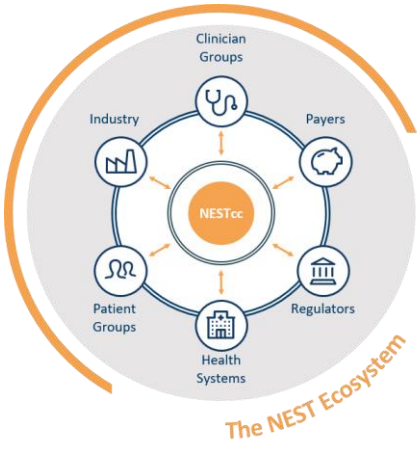

Figure 1. Medical device and health technology ecosystem.

In 2018, FDA tasked NESTcc with establishing postmarket active surveillance capabilities to more quickly detect emerging signals for patient safety by "continuously generating, accessing and evaluating large data sets on device performance and clinical outcomes associated with device use in routine clinical practice" and "improv[ing] the FDA's ability to link adverse events with specific devices." (Gottlieb & Shuren, 2018). This evidence-based approach to proactively monitoring devices in real-world settings is a key area of focus in FDA's 2018 Medical Device Safety Action Plan (U.S. Food and Drug Administration, 2019). To advise the NESTcc staff, the NESTcc Governing Committee approved the creation of the NESTcc Active Surveillance Task Force, a select group of experts with a breadth of experience across the NEST ecosystem (see **Figure 1**).

2.2 Alignment of NESTcc Mission With Active Surveillance

The capability to continuously monitor and evaluate large RWD sets generated during routine clinical device usage has the potential to enable stakeholders to make timelier, evidence-based decisions for patient safety (Gottlieb & Shuren, 2018) and is directly aligned with NESTcc's Mission (**Figure 2**). This capability can transform how medical device benefit-safety profiles are monitored using RWD. While the end goal of the program is to benefit patients through improved device safety, equally important is the focus on patient inclusion in advisory capacities throughout implementation and future operations to build in elements that deliver tangible value to them.

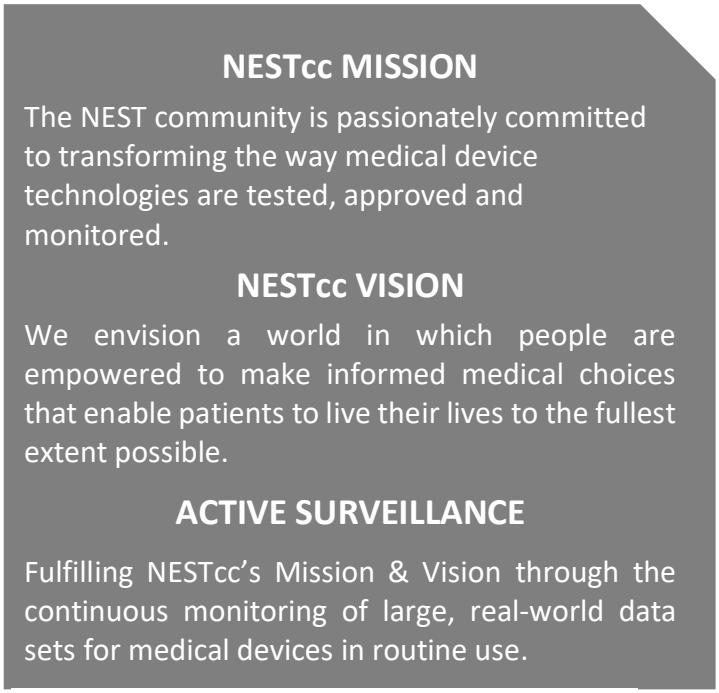

Figure 2. Mission and vision of NESTcc organization.

2.3 Scope of This Roadmap

This document outlines the scope, progress to date and roadmap for NESTcc to develop the capability to support active surveillance and evaluation. The document describes a cloud-based, federated data management and analytics system, supported by guidelines and standardized procedures that make the capability scalable, secure and equitable to the multitude of medical device stakeholders committed to device safety.

3 ACTIVE SURVEILLANCE

3.1 Role of Active Surveillance in Device Safety Modernization

Currently, U.S. regulators seek to ensure continued safety and effectiveness of market-authorized medical devices through a number of methods, including medical device reports (MDRs), the Medical

Product Safety Network (MedSun), post-approval studies, postmarket surveillance studies (522), premarket approval application annual reports, scientific literature reviews, manufacturer reports and inspections. For example, regulations in the U.S. require device manufacturers, device user facilities and importers to report adverse events and problems associated with devices, as well as to encourage others (health care providers, patients and caregivers) to voluntarily participate in identifying the occurrence of a problem that may be associated with a device, then submitting Medical Device Reporting (MDR) (U.S. Food and Drug Administration, 2019). Such feedback is critical to current safety monitoring but limited in its ability to quickly identify new risks and changes in frequency for known risks. This vulnerability – likely due to a combination of underreporting, incomplete reporting or lack of denominator/exposure data – can hinder the robust and timely detection of safety signals. The Active Surveillance system that NESTcc is building is intended to complement existing capabilities.

Two of seven Action Items in the FDA's Medical Device Safety Action Plan (U.S. Food and Drug Administration, 2019) are directly related to NESTcc initiatives and expected to shorten the timespan between the initial identification of a medical device concern and epidemiological evaluation. Briefly, these enhancements include:

1. *Real-world evidence (RWE)*: Realize the value of health data produced in the "real-world" from electronic health records (EHR), claims, registries, patient-monitoring devices and outside of typical clinical trials by establishing guidance around data quality, relevance, reliability and applicability to regulatory decision-making.
2. *NEST active surveillance and evaluation system*: Establish a high-quality, end-to-end process to analyze continuously updated and prospectively collected data sets, then apply advanced analytics to facilitate the characterization and detection of safety signals.

Since its inception, NESTcc has been developing the capabilities and leadership needed to generate reliable RWE to support regulatory decision-making (Item #1 above). As part of this work, NESTcc is also currently developing the systems needed to support active surveillance (Item #2 above).

3.2 The Case for Active Surveillance

Postmarket surveillance can support ongoing safety and effectiveness monitoring of devices to better capture how they are used in real-world settings. Six key reasons cited (US Food and Drug Administration, 2015) as the reasons for including postmarket surveillance in a comprehensive monitoring program include:

1. Access to more data: Larger samples will allow for detection of fainter signals.
2. Expansion of diversity of patient population and inclusion of underrepresented populations: Clinical use will reflect device performance in patients known to be underrepresented in biomedical research, as well as in patients traditionally excluded from clinical trials due to comorbidities, disease severity, etc. (US Food & Drug Administration, 2019) (Knepper, 2018), and while not immune to bias, will help expand the base of device exposures.
3. Usage in real-world conditions: Real-world use will expose learning curve considerations, handling and maintenance issues.

4. Quality over time: Consistency in materials/components, manufacture and supply chain is likely to be exposed, as well as long-term device use and maintenance.
5. Low frequency events: Premarket studies do not always reveal adverse events occurring at low frequencies.
6. Unforeseen challenges: Real-world use sometimes exposes unexpected issues that only monitoring will detect. In addition, medical devices have complexities related to the interplay of hardware (plus components and materials), software, user/operator interaction and device identification.

Postmarket safety monitoring may reveal medical device-associated adverse events, such as unforeseen or unimagined adverse events, increases in severity or frequency of known events, new product-to-product interactions, and malfunctions, injuries or reduction in patient benefit related to improper use or device design. Additionally, postmarket monitoring may detect safety signals associated with a certain manufacturer, a class of similar products, or even more broadly, issues with materials or components (US Food and Drug Administration, 2016). For medical devices, passive surveillance data sources include FDA's Manufacturer and User Facility Device Experience (MAUDE) and the European Union's Eudamed database, where submissions are compiled but subject to problems due to incomplete, inaccurate, untimely and unverified data (Ross, 2015). Active surveillance, while initially expensive and narrow in scope, with projects subject to delays (Ross, 2015), has been shown to be capable of operating at scale. The Sentinel Initiative, supported by a broad range of analytic toolkits released over its 12-year lifetime, has 70 million active members in 2021, covering almost 15 billion cumulative pharmaceutical scripts dispensed across its distributed historical database (Sentinel Initiative, 2021). The Sentinel Initiative's technology stack, transparency, distributed approach and operational infrastructure serve as a model from which NESTcc can learn as it tackles the more challenging domain of device safety. The two systems share the objective of continuously monitoring and evaluating large real-world data sets to enable stakeholders to make timelier, evidence-based decisions for patient safety (Gottlieb & Shuren, 2018).

To address the need for active surveillance of medical devices, NESTcc is creating the "NESTcc Cloud." The NESTcc Cloud will serve as the foundational infrastructure supporting FDA's strategy to advance the use of RWE through a multi-stakeholder, national system that employs active surveillance and is available for both public and private studies. A federated network of data partners running in situ analyses is a key architectural aspect of the NESTcc Cloud and is described in more detail below. The environment is being developed with the goal of lowering existing barriers to robust active surveillance of medical devices, including the cost, time and expertise associated with provisioning infrastructure, analytics, workflows to ensure data transparency and traceability, and frameworks to ensure data quality and limit data-sharing risk exposure among data owners.

3.3 Scope of NESTcc Active Surveillance

Surveillance is a core practice within public health, encompassing an end-to-end process characterized by the "systematic and continuous collection, analysis and interpretation of data, closely integrated with the timely and coherent dissemination of the results and assessment to those who have the right to know so that action can be taken" (International Epidemiological Association, 2008). NESTcc has been tasked with developing *active* surveillance capabilities where a sample of data is expressly and proactively collected and then analyzed. Again, active surveillance is a tool that is complementary to the

passive surveillance and spontaneous reporting efforts currently used to detect safety signals for medical devices.

Medical device postmarket safety surveillance consists of three stages, as promoted by the Sentinel Initiative: signal detection, signal refinement and signal evaluation (Robb, et al., 2012) (McClure, et al., 2014). These stages are described in **Table 1** and applications of surveillance activities are presented.

NESTcc's active surveillance system will primarily support signal detection and signal refinement. Signal refinement studies are expected to be the most straightforward due to their consideration of known or suspected adverse events with one or more medical devices. This type of hypothesis-driven work will be the initial focus of the active surveillance program. Over time, system development will increase automation throughout the data curation pipeline and expand analytical capabilities. As the system matures, NESTcc expects to support signal detection studies, which are anticipated to be more challenging due to their less-defined nature. The active surveillance system is not expected to perform signal evaluation work; however, NESTcc may execute such studies as follow-up outside active surveillance.

Table 1. Definitions of the three stages of surveillance with application examples.

Stage	Description and Applications
Signal detection *(no known signal)*	• Situation: No signal for a pre-specified device exposure – adverse event signal pair exists, but there is a hypothesis or reason to believe it is important to monitor a device for a safety signal • Retrospective and/or prospective data collection may be used • Application: Apply statistical methods to determine whether associations between devices and reports of adverse events are possibly safety signals • Future application: Utilize an exploratory analysis leveraging machine learning analytics to identify associations between experiences with a medical device and the presence of a signal (adverse event or benefit), across a range of potential outcomes • Future application: Configure recurring, automated monitoring and signal analysis of devices, or classes of devices, for key safety outcomes to automatically find signals warranting refinement
Signal refinement *(known signal)*	• Situation: A potential signal has been identified between a device and one or more adverse events, generating a hypothesis that warrants further investigation and data gathering to confirm the presence of a signal (if any) • Retrospective and/or prospective data collection may be used • Collect and statistically analyze data regarding patient exposure to a device, or class of devices, and outcome(s) of interest that are thought to have a relationship for the purpose of validating the hypothesis • Application: Refinement of "emerging signals" under review, postmarket surveillance in support of an FDA 522 Study, low frequency signals associated with rare events or small populations using device, etc. • Future application: Universal monitoring for known, high-risk devices
Signal evaluation	• Situation: A device exposure – adverse event signal pair exists but no causal relationship has been established • Application: A full epidemiological analysis to determine the presence of and characterization of a causal relationship between a device exposure and a particular adverse outcome

4 STRATEGY AND ROADMAP

4.1 Overview

NESTcc is building an active surveillance capability for the medical device ecosystem that reduces the technical, administrative and financial burden of device surveillance, while providing methodologies,

processes, compliance and security to produce trustworthy RWE for the evaluation of potential safety signals. NESTcc will accomplish this goal by building the NESTcc Cloud environment within a framework of quality-managed procedures across six key areas: architecture, interoperability, analytics, data management, research network, stakeholder collaboration and compliance (**Table 2**).

Table 2. Overview of goals for NESTcc's active surveillance service.

Architecture	Interoperability
• **Build and manage a cloud-based platform orchestrating data curation and analytics services across a federation of distributed data partners, with centralized results aggregation and signal monitoring** ○ A principal cloud orchestrating studies across a federation of data partners ○ Containerized query, analytic and study orchestrator tools for deployment into collaborator environments ○ Analysis and retention of patient-level data in collaborators' local data warehouse, sharing only aggregated results outside of firewall to principal cloud ○ Centralized study communications and orchestration of participation; collaborators opt into or out of studies	• **Establish streamlined and standardized local data management** ○ Develop a hybrid Common Data Model (CDM) by adopting an existing CDM and extending it to accommodate medical device data elements, patient-reported information, capacity for wearables/digital/software as a device ○ Provide quality control scripts to validate conformance of data to CDM ○ Support Open APIs and FHIR standards
Analytics	**Data Management**
• **Distribute packaged analytics to data partner environments to ensure consistency** ○ Eight tools from Lahey's Data Extraction and Longitudinal Analysis (DELTA) analysis engine modified to work in a federated environment ○ Addition of data science and statistical workbench application to enable ideation, iteration, and development of statistical or AI-based algorithms ○ Simultaneous accommodation of multiple parallel surveillance projects without service degradation	• **Offer standardized data curation capabilities (transform, query, log, validate)** ○ Support for raw, staging and processed data flow with metadata and logging ○ User tools to facilitate transformation to common data models, terminology harmonization ○ Reusable data quality checks (e.g., model compliance, valid values, missing data, logical and integrity checks) ○ Feasibility of study (data availability) queries ○ Automated PII detection/rejection ○ Support for patient-level traceback, when necessary ○ End-to-end data modification audit trail
Research (and Data) Network	**Security**
• **Expand current network of collaborators** ○ Initial focus is U.S-based collaborators with EHR data and expertise for specific device validation use cases ○ Target is 30 collaborators for active surveillance ○ Create modular contractual agreements to support research collaboration, participation in federated data network, and publications/data use ○ Support collaborators who wish to establish their own data/analytic environment ○ Add non-US-based collaborators	○ **Offer project sponsors and collaborators confidence that private and proprietary information is secure** ○ Aligned with ISO 27001/ISO 27701/ISO 13485/HIPAA/GDPR ○ System architecture and implementation confirmed via independent security audit

After Phase I implementation (estimated Q3 2021), a centralized operational model will exist, with de-identified, study-specific extracts of patient-level data pooled in the single Primary Cloud, as depicted on the left side of **Figure 3**. Phase I is not the target state for operations; it is an intermediate milestone. The right side of the figure illustrates the targeted federated operating model, which Phase II development will deliver. In Phase II, all patient-level data is retained within the collaborator-managed environments, behind their existing firewalls, and only aggregated results data are passed to the Primary Cloud. A data retention policy, with guidance pertaining to data reuse, reproducibility of analyses and version control/data provenance, will be developed.

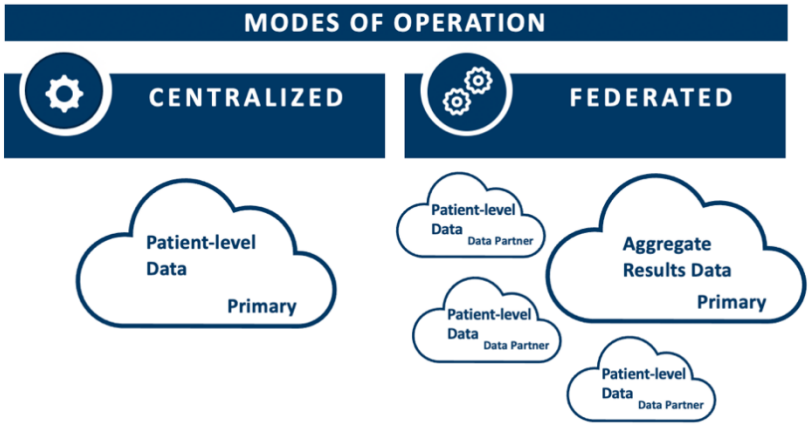

Figure 3. Centralized vs. federated modes of operation and patient-level data.

4.2 General Approach

4.2.1 ADVISORY WORKING GROUPS

The NESTcc Governing Committee established the NESTcc Active Surveillance Task Force to provide guidance on the development of an active surveillance capability. The Task Force, in turn, approved the establishment of two specialized Working Groups (WG): the IT Cloud WG and the Active Surveillance Methodology WG. A third working group, the Data Curation WG, is planned.

Each of the three WGs has a specific charge, set of advisory responsibilities and regular cadence of meetings. The IT Cloud WG has been responsible for advising the technical implementation of Phase I and will be refreshed to perform a similar activity for Phase II. Because the Phase I IT Cloud WG also included the test and validation study principle investigators, this WG has also been involved in Phase I validation and testing. The Active Surveillance Methodology WG has been tasked to identify a list of analytic methodologies and/or tools that would be deployed into the Phase II NESTcc Cloud environment. The soon-to-be-formed Data Curation WG will focus on aspects of data curation including, but not limited to, data standardization, organization and integration.

4.2.2 PROJECT MANAGEMENT: TECHNICAL IMPLEMENTATION

The build of the NESTcc Cloud will follow the Agile version of the Enterprise Performance Life Cycle (EPLC) process by which HHS plans, manages and oversees IT projects. The project team is also using associated FDA versions of document templates to ensure completeness, especially pertaining to privacy and security assessments and related compliance. **Figure 4** shows a diagram representing the general process, blending a waterfall approach with Agile development cycles. Panel members for the various Stage Gate Reviews are drawn from members of the Task Force and/or WGs with the request that those who volunteer for each specific Review believe they are knowledgeable and experienced to bring a critical eye to the materials being reviewed.

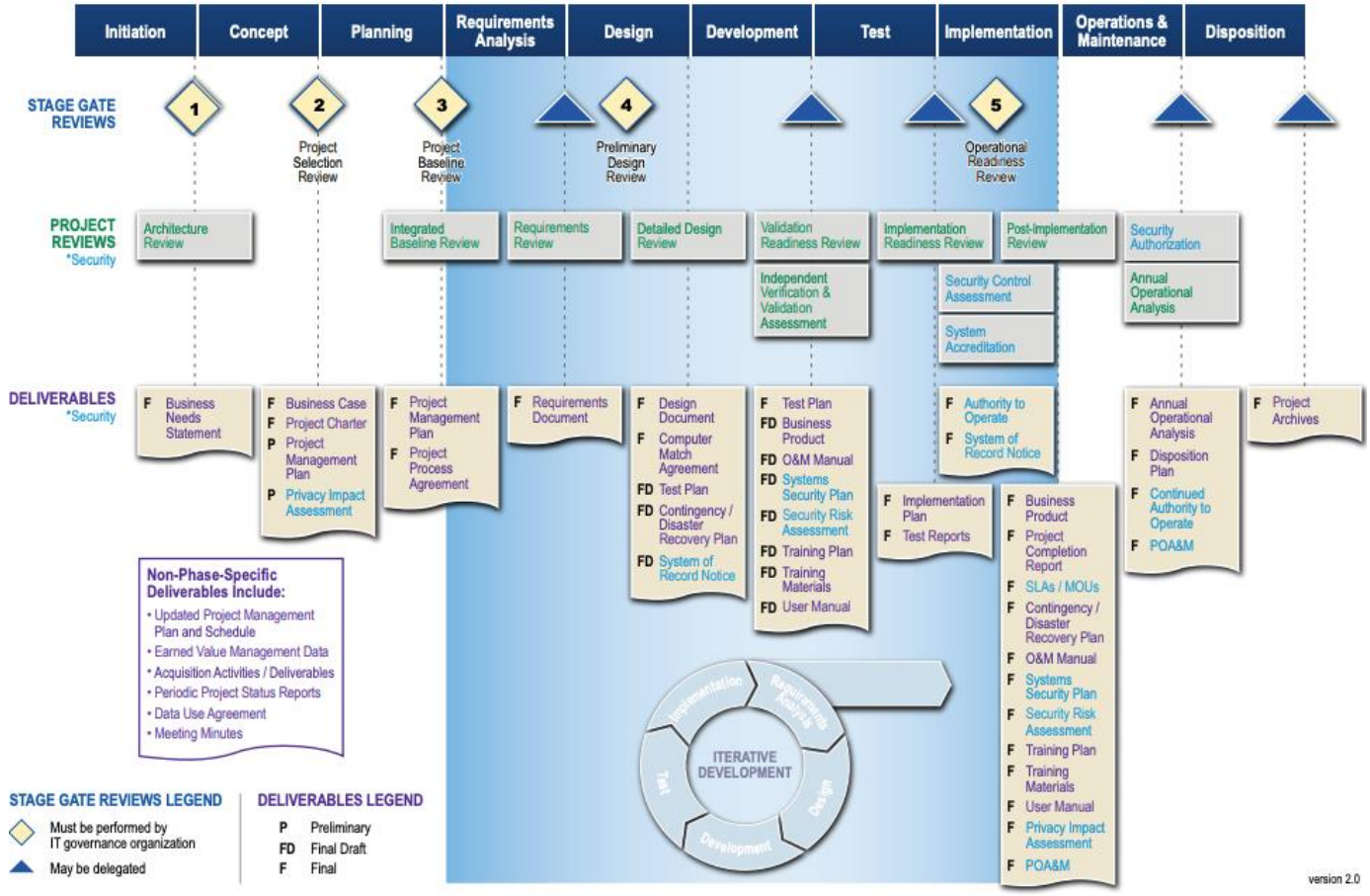

Figure 4. EPLC process flow.

4.2.3 PROJECT MANAGEMENT: SURVEILLANCE STUDIES

Building on the research project management experience NESTcc has developed since its inception, surveillance studies using the NESTcc Cloud environment will follow essentially the same operational flow. NESTcc's intent is to enable the complex analysis of big data in a secure and audited environment with methodological and ethical guidelines, multi-stakeholder collaboration and project management processes governed by an ISO 13485-based quality management system. Active surveillance project operations are planned to be assessed against additional standards, including HIPAA, GDPR, and ISO 27001/27701. **Figure 5** illustrates the NESTcc process flow by which managed studies follow.

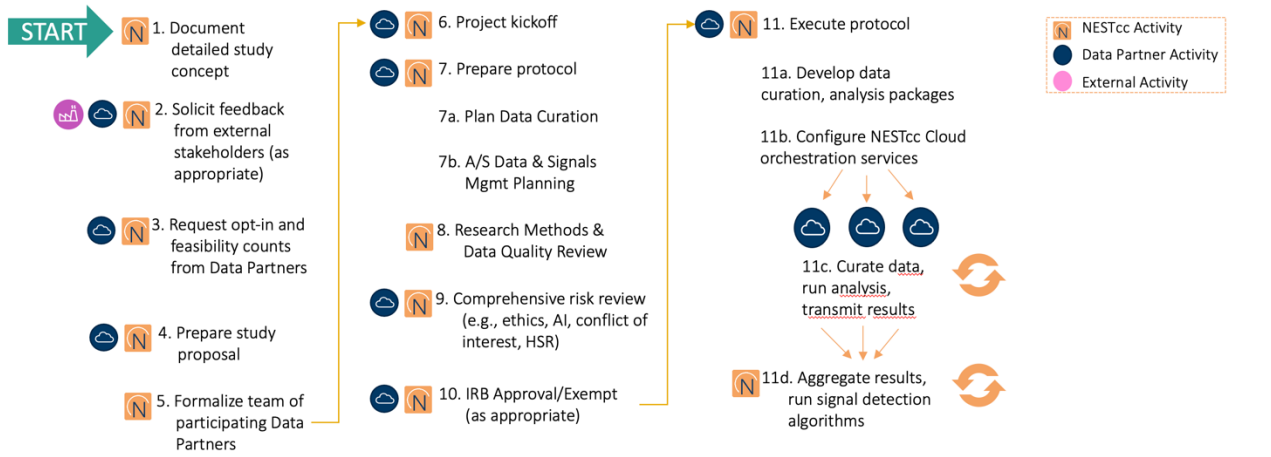

Figure 5. NESTcc active surveillance project process flow diagram.

4.3 Building the Active Surveillance Data Network

NESTcc has assembled a Research Network of 16 internationally recognized Network Collaborator institutions with access to RWD sourced from EHRs, registries, pharmaceutical and medical claims, patient-generated data, billing, supply chain and genomic data. As of June 2021, this group represents more than 161 million patients, about half as many patients represented in the Sentinel Distributed Database (Sentinel Initiative, 2021). Developing and expanding a network of Data Partners specifically for active surveillance work is an essential component of building the capability of active surveillance. It is NESTcc's vision that a diverse group of Data Partners be engaged to contribute data and research expertise to active surveillance, and current Network Collaborators may choose to participate as well. Broad representation of community and research institutions with RWD generated in urban and rural settings across the U.S., and eventually internationally, will contribute to the program's ability to monitor large volumes of data representative of real-world device usage. A robust federation of collaborators will enable NESTcc to efficiently execute studies, while providing flexibility to collaborators to choose whether to participate in specific studies.

NESTcc is in the process of expanding its Research Network to include active surveillance Data Partners. NESTcc plans to issue a request for information (RFI) to both present a value-based case for joining NESTcc's Research Network for active surveillance and to collect information on the capabilities the collaborator would bring to the program. Plans are to issue the RFI during Q3 2021.

4.4 Activity to Date and Upcoming Activities

Details of the NESTcc Cloud build and its phases are discussed in the Technology section of this document.

To date, the following Phase I foundational activities have been completed or are due to be completed in the near term:

- Establishment of NESTcc Active Surveillance Task Force (Dec 2018)

- Competitive selection of a vendor to build the NESTcc Cloud platform (Apr 2020)
- Designation of two Working Groups (WG) – IT Cloud WG and Methodology WG – to provide multi-stakeholder input into NESTcc's active surveillance work (Jul 2020)
- Task Force completion of first EPLC Stage Gate Review – Project Selection Review (Sep 2020)
- Kickoff of NESTcc Cloud Phase I project with contractor (Nov 2020)
- Validation projects using synthetic data with health system collaborators Mercy Health and Lahey Hospital and Medical Center (May 2021)
- Demonstration projects using de-identified EHR data from two collaborators (Replication of bare metal vs. drug-eluting cardiovascular stent research, extension of stent research to compare those using zotarolimus and everolimus) (June/July 2021)
- Delivery of NESTcc Cloud Phase I (July 2021)
    - Implication: NESTcc Cloud (Phase I environment) is available for research projects that can use DELTA's analysis tools. At this phase, collaborators must be willing to pool relevant, study-specific, de-identified, patient-level data in the "Primary" cloud given the federated architecture will not be available until Phase II delivery

The NESTcc Cloud Phase I project schedule is outlined below in ***Figure 6***.

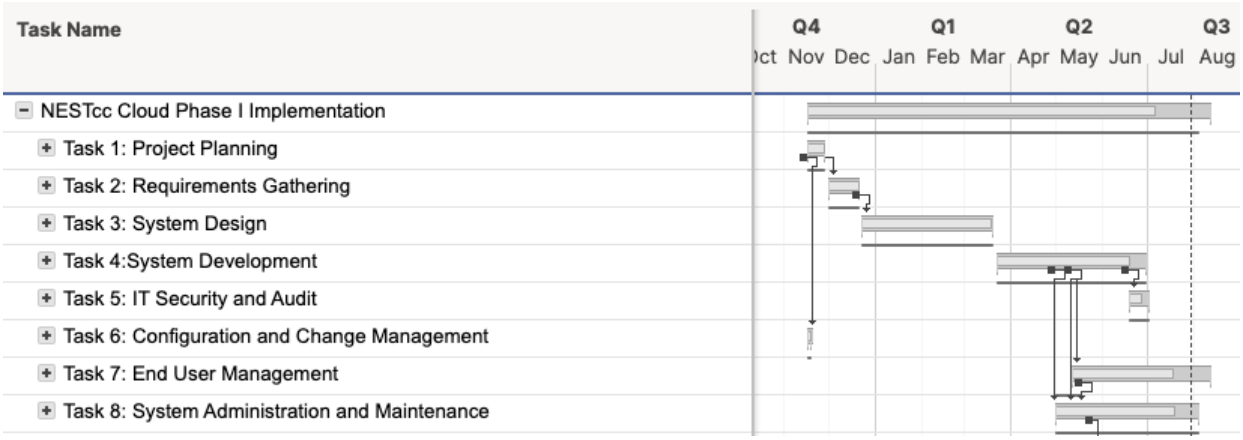

Figure 6. Active surveillance project schedule (as of July 2021).

The following key milestones are upcoming:

- Establishment of the Data Curation WG (estimated September 2021)
- Kick-off of NESTcc Cloud Phase II project with contractor (estimated Q4 2021)
- Delivery of NESTcc Cloud Phase II (estimated Q3 2022)

For NESTcc to achieve the Phase II kickoff milestone noted above, the NESTcc active surveillance team and WGs are working in parallel to shape, define and decide upon multiple details with long-term strategic implications. ***Table 3*** shows workstreams of strategic importance where work is currently in progress or will commence in the future.

Table 3. NESTcc active surveillance strategic workstreams.

Category	Workstream
Research	• Adopting and extending a Common Data Model (CDM) • Phase II validation use case selection and study definition • Inclusion of the patient perspective • Selection/acquisition of additional analytic capabilities to include in Phase II • Data curation (query, CDM, data QA, pipeline processing, metadata) • Expansion of active surveillance collaborators and issuance of a request for information (RFI) to solicit interest • Governance
Operations	• NESTcc team growth and development • Phase II technical requirements and draft statement of work • Legal and contractual strategy and templates for new collaborators • Support for regulatory studies, inquiries and submissions • Technical operations planning

5 TECHNOLOGY

5.1 Overview

The NESTcc Cloud solution is designed to be the resource of choice for regulatory scientists, as well as investigators from industry, patient groups, providers and health systems conducting active surveillance (or general) research on medical devices. The NESTcc Cloud orchestrates studies across the federated network of collaborators. Patient-level records will remain under the control of the collaborator, with NESTcc providing resources to standardize distributed data curation and analytics, as well as results-level data sharing. Each collaborator-managed data warehouse is expected to contain up-to-date information from the collaborator's source data systems (EHR, claims system, registry or other) with an enclave for processed study data. This technical architecture meets several objectives. First, collaborators retain full ownership and control of patient-level records. Second, NESTcc can centrally issue requests to collaborators to opt into or out of study participation. Third, the architecture facilitates data harmonization through the consistent application of a CDM and quality checks and allows for the replication of the same query and analysis settings/parameters across participating collaborators to minimize variation. Finally, it lowers technical barriers to participation in surveillance research.

5.2 Two-phase Build

Based on input from the Task Force, NESTcc has taken a phased approach to the technology build to expedite the instantiation of a fully operational system for active surveillance while continuing to thoughtfully enhance and expand system capabilities. Key distinctions between Phase I and II capabilities are outlined in **Figure 7**.

PHASE I	PHASE II
TEST PLANS	**TEST PLANS**
• Synthetic dataset: Limited covariate complexity (n~20), hidden associations in data elements, missing data, adverse events for signal detection, ~10K records to simulate a "real" population from one health system • Cardiovascular: De-identified EHR data from two sources with a) known results and b) new data comparing two drug-eluting stents	• Synthetic dataset: EHR data with advanced covariate complexity, hidden associations in data elements, missing data, adverse events for signal detection, learning curve effects, data entry/mapping errors, ~100K records to simulate real population • "Real" scenario(s) of interest to FDA: ventilators, robotic surgery, breast implants, fully disposable duodenoscopes
INGEST	**INGEST**
• Standardized flat file containing data from two collaborators • De-identified, patient-level data aggregated in principal cloud	• Collaborator receives study request, query, data curation details, data quality validation logic, study instructions, analysis package • NESTcc Cloud receives opt-in/opt-out, analysis results
PROCESSING	**PROCESSING**
• Automated check and rejection of files containing PII • Data validation pipeline on formatting, data types, valid values	• Collaborator's pipeline processing converts raw data into NESTcc CDM; cleans and standardizes; imputes, calculates, and derives according to instructions, writes metadata • Quality and harmonization checks run
ANALYSIS	**ANALYSIS**
• Run through containerized DELTA 3.7 analytic engine [all analyses (e.g., descriptive statistics, propensity matching)] • Format of results output in standard format (e.g., CSV)	• Collaborator receives analysis package from NESTcc Cloud, analysis executed locally • Collaborator pushes results and statistics to NESTcc Cloud for aggregation and assessment
REPORTING	**REPORTING**
• DELTA and NESTcc Cloud Dashboards	• Additional standard dashboards and reports in added to NESTcc Cloud

Figure 7. Phase I and II test plans and system capabilities.

Overall, Phase I will deliver basic cloud infrastructure capabilities of data ingestion, processing, analysis, and reporting with limited features and options for a single cloud instance. The Phase I validation use cases assume pre-defined standardization of de-identified patient-level data, limited validation and processing of the raw data before injection into the open-source DELTA statistical analysis tool containerized and deployed within the NESTcc Cloud, with results displayed in DELTA's dashboarding tool. Phase II will add the communication and orchestration functions for participation needed to work in a federated analytics environment, enhance the data curation and data validation processes, install additional analytic capabilities and enhance dashboards.

5.3 System Architecture

Conceptual diagrams of the Phases I and II system architecture are illustrated in **Figure 8**.

At the top of the diagram, the Phase I architecture is depicted with the two Data Partners participating in the validation and test projects, sharing de-identified patient-level data to the NESTcc Cloud. Data validation, processing and metadata curation occur in the NESTcc Cloud database (DB) storage. Analysis is run using DELTA 3.7 deployed in a container. The bottom half of **Figure 8** illustrates the proposed

Phase II NESTcc Cloud conceptual architecture interfacing with one (of many proposed) Data Partner environments. The Phase II NESTcc Cloud is shown with orchestration services added for managing the coordination and communications with the newly added remote Data Partner environments. Planned in Phase II is the addition of analytic capabilities beyond DELTA, ranging from traditional statistical applications like SAS to support for open-source data science packages based on Python and R to offer capability and flexibility. The Phase II diagram also conceptually illustrates the '*n*' independent, firewalled Data Partner environments, with orchestration commands flowing from the NESTcc Cloud and responses and aggregate results and statistics flowing back into the NESTcc Cloud. While not depicted, study investigators and sponsors may have access to the aggregate results and dashboards in the NESTcc Cloud, while patient-level data remains behind the firewall of the source environment. At the bottom of the Phase II figure, an exception to the rule is noted, when FDA requests a specific study-related extract of patient-level data in order to conduct its public health activities. Please see Section 7.2 for additional discussion.

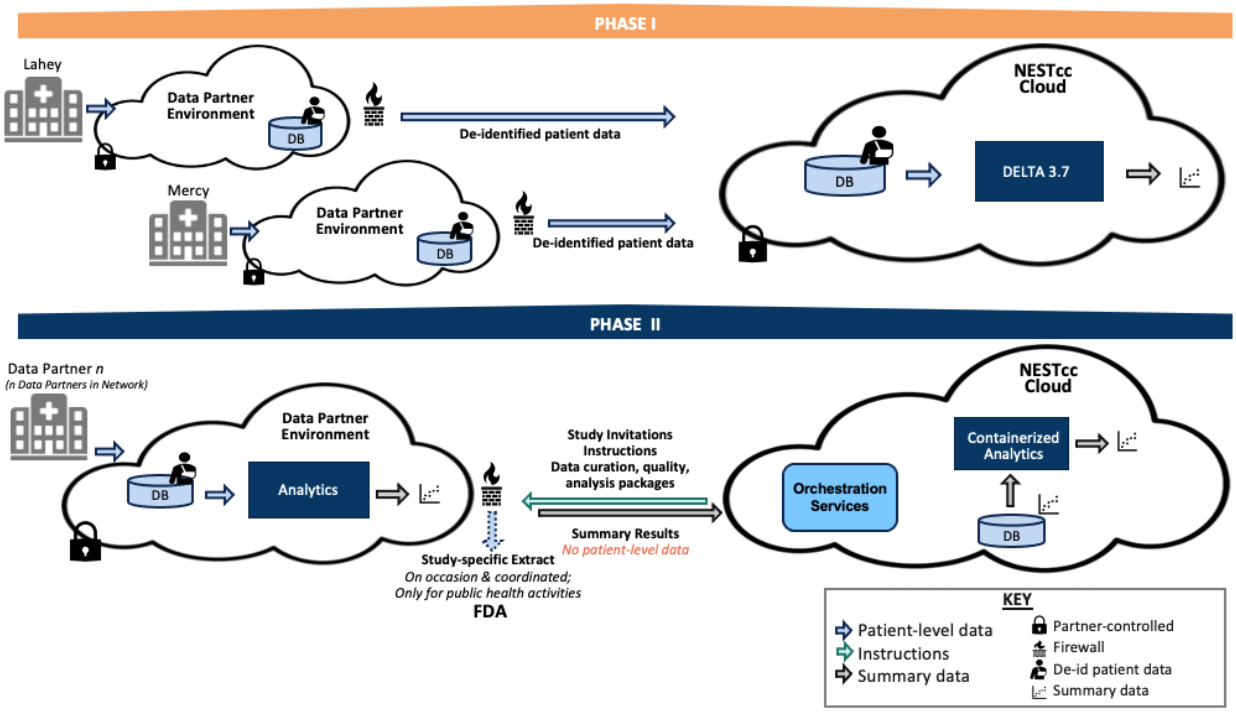

Figure 8. Conceptual architecture of NESTcc Cloud Phases I and II (as of July 2021).

A more detailed diagram of the Phase II system architecture, showing various zones and services to match the conceptual diagram above, is provided in **Figure 9**.

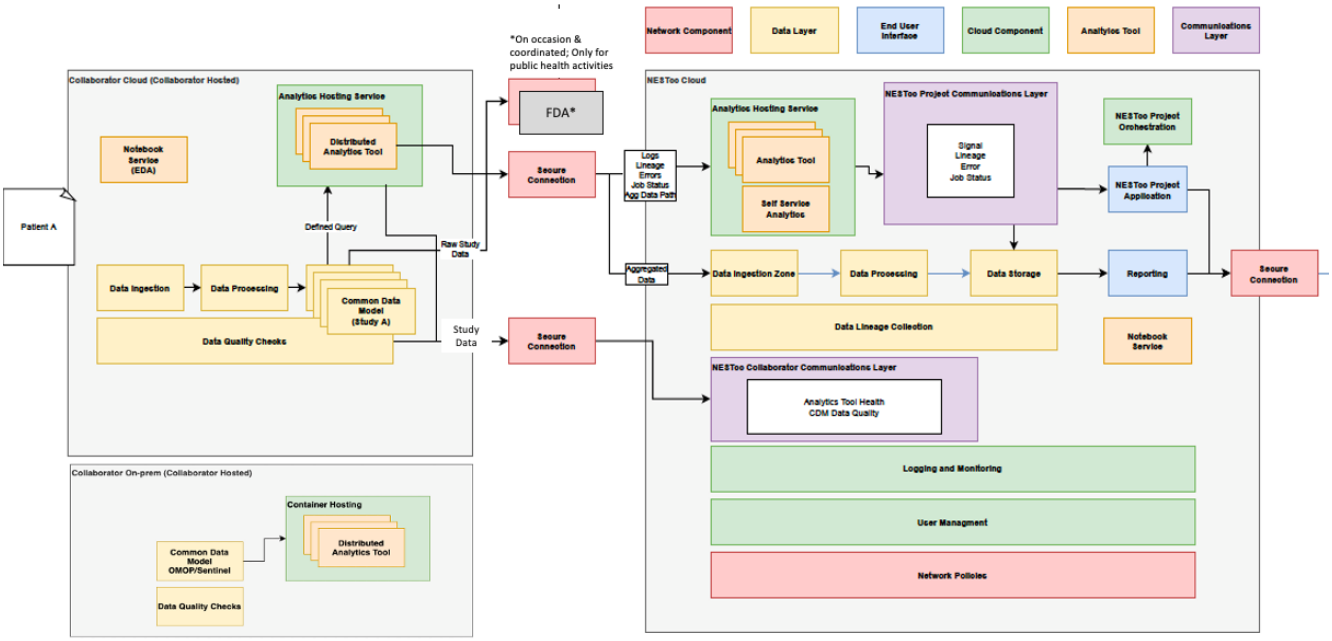

Figure 9. Phase II System architecture schematic diagram.

5.4 Data Characterization

5.4.1 COMMON DATA MODEL

For efficient and consistent execution of federated data analyses, NESTcc will use a hybrid Common Data Model (CDM), adopting a to-be-determined existing data model and extending it to include tables and data elements for under-represented data, such as characteristics of a medical device. In addition, NESTcc intends to have Data Partners start with patient-level data in their rawest form, with NESTcc providing options to reduce the data engineering/curation effort by accommodating different formats' starting points and leveraging transformation tools. The standardization of data element meaning and, to a lesser extent, structure, is critical for using distributed datasets.

The core of NESTcc's CDM will leverage key features selected from among four existing CDMs: Observational Medical Outcomes Partnership (OMOP), Sentinel, Informatics for Integrating Biology & the Bedside (I2B2), and the National Patient-Centered Clinical Research Network (PCORnet). The criteria to evaluate candidate data models are under development and are expected to include: the data type and purpose for which the CDM is commonly used, the data tables and elements, the feasibility of extending the CDM to accommodate expected covariates, the interest of the CDM managing organization in working with NESTcc, and the availability of open-source helper tools and code to facilitate data transformation, processing, queries and analytics. NESTcc is seeking input from the public, including collaborators and ecosystem stakeholders.

The extensions needed to accomplish NESTcc's goals will vary based on tables and data elements included in the selected core CDM. It is expected that any core CDM will require enhancements to accommodate device-related fields. Other extensions that have been identified for consideration include: support for both key outcomes (mortality, need for a repeat-procedure, re-admission) and

other device-specific outcomes, patient reported outcomes and measures, metadata, operator data, cost data and others. NESTcc's approach will be to consider the use cases at hand and identify a parsimonious set of data requirements to meet the study objectives, not to build data structures beyond what is needed, when it is needed.

5.4.2 DEVICE-RELATED DATA

The benefits of scale provided by a distributed data network need to focus on exposure to a particular device, or subset of devices, under study. Ideally, the RWD sources would contain a Unique Device Identifier (UDI) or UDI Device Identifier (UDI-DI), akin to a drug's National Drug Code (NDC) unique identifier. Unfortunately, use of UDIs is not ubiquitous at this point in time, an obstacle that poses challenges to querying all patients exposed to a specific device or augmenting the data with device characteristics from the Global Unique Device Identification Database (GUDID) hosted by FDA. The NESTcc team is exploring approaches to improve the identification of a specific device within the Data Partner environments.

5.4.3 DATA ELEMENTS EXPECTED

As listed in **Table 4**, test and validation data used during the Phase I implementation are derived from EHR systems and EHR-like synthetic data. The team expects to capture for all research activities a broad set of data elements that span demographics, vitals, clinical history, therapeutic details and outcomes. A specific and limited set of data elements, listed in **Table 4**, are required to be handled during Phase I/II, with an as-needed expansion of fields as future projects warrant. NESTcc plans to expand the sources of data surveilled in Phase II and beyond.

Table 4. Expected data elements in NESTcc Cloud system.

	Phase I	Phase II
Data Source		
EHR	✓	✓
Registry		✓
Claims		✓
Patient-generated health data (PGHD)		✓
Data Element		
Demographics	✓	✓
Comorbidities	✓	✓
Vital signs, assessments (e.g., pain, physical therapy, mobility)	✓	✓
Device exposure, characteristics	✓	✓
Medications (orders and inpatient administration)	✓	✓
Labs (relevant orders and results)	✓	✓
Diagnostic procedures	✓	✓
Surgical procedures	✓	✓
Supplies	✓	✓
Outcomes: Mortality, re-hospitalization, re-operation	✓	✓
Outcomes: Project-specific		✓
Other: As needed by use cases		✓

5.5 System Testing and Validation

5.5.1 PHASE I TESTS

In Phase I there are two datasets for test and validation purposes: one using synthetic data and one based on real-world, EHR-derived cardiovascular stent data as detailed in ***Figure 7***. A collaborator from Vanderbilt University Medical Center, not otherwise directly involved in the NESTcc active surveillance project, provided synthetic data in the form of a flat table, with approximately 10,000 records, delivered in separate files, to represent multiple submissions over time. The data have limited covariate complexity (n=20), but include hidden associations among the data elements, including a certain percentage of missing data and a non-zero adverse event rate (a signal). The purpose of this file was to test the system, including data ingestion within DELTA system, user and methods verification testing, as well as DELTA results extraction.

Following the successful system tests with the synthetic data, testing proceeded with two validation use cases containing data on patients treated with cardiovascular stents. The datasets were comprised of EHR data augmented with unique device identifier (UDI) information from inventory systems and FDA's Global UDI Database (GUDID), mortality data from the Social Security Death Index (SSDI) and other complementary data sources. The first validation use case replicated[1] previously published research using data from a single data warehouse (Drozda, Zeringue, Dummitt, Yount, & Resnic, 2020), validating the NESTcc Cloud system with DELTA 3.7 embedded. The second validation use case utilized new data provided by two collaborators to simulate an integrated network of real-world health care data sources. The objective was to validate the system by performing a prospective safety analysis using both inverse probability treatment weighting (IPTW) and propensity score-matched survival analysis on everolimus versus zotarolimus drug-eluting stents. Successful completion of these validation use cases is a principal exit criterion for the Phase I build.

5.5.2 PHASE II TESTS

The Phase II validation use case selection process is underway and has not yet been finalized. Options discussed have included the use of both synthetic EHR data and "real" EHR data. The Phase II test data will be transformed into the NESTcc CDM, with additional pre-determined processing steps executed.

If a synthetic dataset is used, it should have advanced covariate complexity with hidden associations among the data elements, including a percentage of missing data and a non-zero adverse event rate within a relational data structure. To add realism, the data set will include an order of magnitude more records than in the Phase I synthetic data set (i.e., >100,000) and represent submissions from multiple collaborators with both data entry errors and mapping errors that vary by simulated source. Finally, a means to simulate data updates and refreshes over time could also prove useful. NESTcc believes that having robust synthetic data sets in the NESTcc Cloud will be valuable in supporting the development and refinement of the data curation and data quality checking scripts and algorithms, feasibility and

[1] We expect to find small differences in the propensity matched analyses since there is some random sampling to find 1:1 matches. If results are found to be significantly different from those previously published, analysis will be iterated upon to obtain a distribution of results which will be the basis for comparison.

study queries, analytic routines, and reporting dashboards, without the risk of handling real patient data. Providing synthetic data will not be a requirement for becoming a data partner.

NESTcc is actively working to determine which data validation use cases will be pursued in Phase II. This decision has two dependencies: (1) the capability of the to-be-determined analytical tool that NESTcc is in the process of identifying and (2) existing and new to-be-determined collaborator clinical areas of expertise and available data. In 2020, the Active Surveillance Task Force was asked to provide potential validation use case areas. FDA reviewed these use case areas and highlighted focal points of interest. The medical devices for which Phase II validation use case(s) will be designed include:

1. Ventilators
2. Robotic surgery systems
3. Breast implants
4. Fully disposable duodenoscopes

6 RESEARCH METHODS AND DATA QUALITY

6.1 Research Methods

6.1.1 GENERAL

The NESTcc Research Methods Framework is an efficient and comprehensive guidebook for prospective study design in the medical device ecosystem. The document outlines both general and specific principles to adhere to according to the product life cycle stage of the device/technology. The NESTcc Data Quality Framework provides guidance on data governance, characteristics, capture, transformation and curation. The Active Surveillance Methodology WG has also been assembled to specifically identify and characterize current methods useful for active surveillance studies, particularly on the four possible Phase II validation use case areas.

NESTcc research scientists will work with validation project teams to examine protocols, results and other study materials to ensure alignment with the NESTcc Frameworks. Taking a page from Checklist Manifesto (Gawande, 2010), NESTcc reviewers will vet study materials via a “do-confirm” checklist aligned with the NESTcc Frameworks. The protocol reviews will be conducted by a NESTcc research scientist prior to their submission to the applicable institutional review boards (IRBs). In the case of FDA-requested studies, the NESTcc research scientist will collaborate with and defer to the FDA-assigned research scientist.

Bottlenecks in the protocol review process may surface as study volume increases. Two concepts for mitigation are outlined below, and while these suggestions are under consideration, no work is currently being done to cultivate these capabilities:

- It has been suggested that NESTcc consider establishing a team of reviewers to augment the current capacity of NESTcc internal staff to review protocols and methods. This has not been implemented to date, but it is a possibility if and when study volume warrants such an approach. These independent, scientific reviewers would be briefed on NESTcc methodologies and assigned studies according to relevant expertise and absence of conflicts of interest.

- Governance for these studies will follow standard IRB protocols in place at collaborating institutions. The concept of NESTcc establishing a "Central IRB" has been discussed as a potential future improvement to address the known constraints and challenges associated with coordinating across multiple collaborator IRBs.

6.1.2 ACTIVE SURVEILLANCE METHODOLOGIES

The design of the NESTcc Cloud is intended to support straightforward distribution of analytic tools to research teams through containers or other means, as necessary for particular collaborators. For collaborators with cloud-based data warehouse environments, containers allow software or tools to easily and consistently be deployed without concern for application dependencies (as these are resolved within the container). NESTcc is deploying one containerized, analytic software package, DELTA, during Phase I, and will be adding one to-be-determined) containerized analytic in Phase II. The team believes deployment of an analytic workbench tool would also be beneficial so that research teams could explore and develop their own code in, for example, Python or R. Another option would be to deploy a licensed product, such as SAS.

In Phase I, a containerized version of DELTA 3.7x will be deployed, offering frequentist inference methods, which are designed to predict truths based solely on data in the current experiment. To date no Bayesian inference methods, which rely on prior known probability to predict the current truth, have been included. The specific methods available from DELTA 3.7x are summarized in **Table 5**.

Table 5. Risk-adjusted inference methods in NESTcc Cloud's instance of DELTA 3.7x.

Frequentist, Risk-adjusted Inference Methods	
Propensity matching	Sequential probability ratio test (SPRT)
Inverse probability treatment weighting (IPTW)	Weighting-by-odds
Survival analysis Propensity-matched survival analysis	Logistic regression

Selection of analytic tools is a priority, as this decision is a prerequisite for completing Phase II technical requirements. The Active Surveillance Methodologies WG was enlisted by NESTcc to better inform the decision in Q3 2020. To date, a final selection of the tools has not yet been made, but a combination of statistical (e.g., SAS) and open-source data science tools using Python and R packages is under consideration. It is important to maintain flexibility in the analytic capability of the system and accommodate the needs of the studies as well as the preferences of the federation Data Partners. Inclusion of DELTA in Phase II will require some updates to function within the federated environment.

6.1.3 DATA QUALITY

Strategies and operating procedures for producing quality data leverage the practices developed by NESTcc for traditional RWD-based research projects will be applied to active surveillance studies. **Table 6** outlines practices planned for active surveillance studies. NESTcc intends to continuously improve this list based on lessons learned during future operations and from entities like Sentinel working in the safety surveillance space.

Table 6. Standards impacting data quality.

Procedural Standards impacting Data Quality	
Protocols	**Project Management**
Written prior to study initiation with approvals following standard IRB review process (as necessary)	Engage key stakeholders early
Updated as needed, with version control and re-review	Communicate frequently with stakeholders
Pre-defined analysis plan, including sensitivity levels, definition of a signal, scheduled interim reviews	Disclose and address conflicts of interest
Pre-defined signal validation plan	Plan for contingencies for smaller-than-expected data sets, mitigate long term through the expansion of the Research Network
Pre-defined signal communications plan (including means for patient traceback, regulatory review and reporting)	Establish publication and results dissemination plan
Data	**Cloud Access and Security**
Documentation of data quality and defects	Technical safeguards (e.g., multi-factor authentication, VPN perimeter)
Documentation of how missing data is handled	Independent security roles to system features and datasets
Documentation of how subject "dropouts" and re-appearances are handled	Distributed, de-identified, study-relevant, patient-level data saved in collaborator-managed Clouds, with no sharing or pooling of data (exception: see Section 7.2)
Documentation of data value mapping for multi-sourced datasets	Automated detection/rejection of PII, with downstream reviews to prevent re-identification
Metadata tracking data provenance throughout processing	System-wide logging enabled
Data usage documented in Data Use Agreements	Rigorous access authorization process
Signal Management	
Capture multi-stakeholder requirements and recommendations to document steps and timelines (like service level agreements) in the event of a signal alert, confirm alignment with FDA	
Monitor developments on FDA expectations for MDR reporting in RWE setting	

7 SIGNAL MANAGEMENT AND COMMUNICATIONS

7.1 Signal Detection

The objective of surveillance is to monitor data for a "signal," which the FDA defines as new information that suggests an association between a device and one or more adverse events and warrants further evaluation and/or action (U.S. Food and Drug Administration, 2019). An "emerging signal" connotes a signal of which there is early awareness, and an initial determination has been made by reviewers that there is potential for an adjustment of the existing benefit-risk profile of the device (US Food and Drug Administration, 2016). As mentioned in **3.3 Scope of NESTcc Active Surveillance**, NESTcc's efforts will focus on signal refinement.

Trustworthy results arise from thorough and pre-defined outcomes and thresholds, robust analytical methodologies, and careful scrutiny of results and factors that may influence them. A signal refinement study will begin with the disclosure of a signal – a known or new adverse event – associated with a medical device.

Signal

New information that suggests an association between a device and one or more adverse events and warrants further evaluation.

Alert

A message triggered by a data analysis algorithm when a threshold established for an outcome of interest is exceeded. Handled as sensitive information.

Figure 10. Definitions of key terms.

Initially, active surveillance projects will be managed according to the same processes as used and validated for NESTcc Test-Cases. A NESTcc project manager and research scientist will be assigned to the study and will collect as many requirements as known to issue a participation request to the Research Network to opt into the work. Rough inclusion/exclusion criteria will be shared so that feasibility-level patient exposure counts may be summed. For studies deemed feasible, activities will proceed according to **Figure 5**'s depiction of the NESTcc study process flow. In collaboration with the study sponsor[2] (entity requesting the investigation) and committees participating in governance and oversight, the research team will define thresholds on study outcomes of interest and alert logic[3] will be configured in the NESTcc Cloud. The team will also confirm and document the procedure for communications of alerts. This procedure will include establishing a regular cadence of communications in the event of an absence of alerts. NESTcc plans to document and manage changes to its approach to signal management in a standard operating procedure (SOP) as part of its quality management alignment to ISO 13485.

Once a study is in progress with the data curation pipeline established and its outputs validated, the data analysis task will be disseminated to participating collaborators and run locally by each collaborator on the curated study-relevant datasets. Collaborators will submit the results of the locally run analysis, not patient-level data, to the NESTcc Cloud, where the various results sets are aggregated, and alerts are triggered where applicable.

In the event an alert is triggered at time zero (T-0), the (Phase II) system will issue an automated, non-specific message to the study's principal investigator and NESTcc project manager. If FDA is a project sponsor or key project stakeholder, the FDA's designated point of contact will be included in this initial notification.

When an alert is triggered, a preliminary yet comprehensive review of the study will be conducted by knowledgeable members of the project team. Reviews of technical configurations, data sets, metadata, queries, processing, locally contributed results sets and aggregated results should be conducted in preparation for briefing the full project team and project sponsor within five (5) business days. The intent of this Preliminary Evaluation Briefing is to rule out errors in the analysis that could have erroneously triggered the alert, ensure clear and open communications with key project stakeholders, and set a plan and timeline for next steps. This work is **not** to be confused with a Signal Evaluation effort (see definition provided in **Table 1**), which is outside the initial scope for NESTcc. If FDA is a project sponsor or on the project team, the agency's point of contact would be included in this briefing. If

[2] The study sponsor may be FDA or any other public or private institution materially supporting the investigation of RWD related to medical devices. NESTcc retains the right to decline research requests for ethical or for other reasons that do not support the mission and vision of the organization.

[3] See the Appendix for additional expectations and considerations on the subject of alerts.

questions are raised in the Preliminary Evaluation Briefing, the project team will make every effort to obtain answers and will support any regulatory reporting as required by law. A notional timeline of signal management activities is illustrated in **Figure 11**.

The NESTcc stakeholder community has recognized challenges with the use of RWE and meeting medical device reporting requirements, and NESTcc has a role to play in this evolving landscape. As such, for manufacturer-sponsored NESTcc active surveillance activities, NESTcc will work with study sponsors to facilitate their meeting of applicable regulatory timeframes for reportable adverse events. NESTcc will also continue to closely monitor developments and incorporate updates into active surveillance study procedures to ensure compliance with FDA policies on adverse event reporting as they are made available.

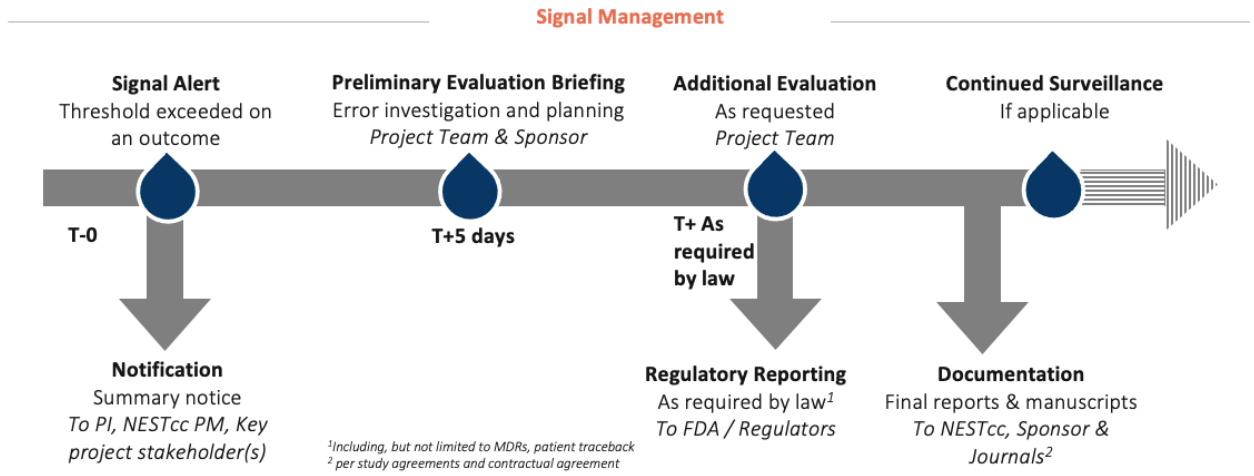

Figure 11. Flow of activities following a signal alert with notional timeline.

7.2 Patient-level Data Sharing for FDA Regulatory Purposes

The design of the NESTcc Cloud's federated analytics architecture is intended to protect and preserve patient privacy while allowing active surveillance studies on medical devices across distributed data sources. FDA has an occasional need to receive patient-level data that may contain protected health information (PHI) from active surveillance activities. Currently, NESTcc is exploring technical approaches on how to facilitate data partners' disclosing of patient-level health information containing PHI to the FDA with approval of the data partner(s).

7.3 Public Communications by NESTcc

It is NESTcc's position that research studies in which NESTcc is serving as a contracted service provider are fulfilled within contractual guidelines that include clauses related to protection of client-proprietary information, limitations on information dissemination and non-disclosure agreements. As such, NESTcc

will not disclose or disseminate information on proposed or active studies, their sponsors or the signal being validated outside of the project team, project sponsor and agreed-upon project stakeholders in accordance with contractual obligations.

In order of importance:

- NESTcc will follow the letter and the spirit of the law, which includes both regulatory and contractual guidelines
- NESTcc supports work that empowers people to make informed medical choices and enable them to live their lives to the fullest extent possible, in accordance with NESTcc's vision
- NESTcc promotes the publication of scientific research that advances the safety profiles of medical devices

8 APPENDIX

8.1 Signal Alerts

It is expected that alert thresholds and acceptable false-positive rates may vary by outcome, with more grave outcomes warranting lower thresholds so as to trigger an alert earlier. For example, a side effect of nausea may trigger an alert when a difference >10% with α=0.05 is observed, while mortality may trigger an alert when a difference of >5% with α=0.10 is observed. Pre-defining criteria that trigger further investigation of unconfirmed signals is the first critical element of a usable surveillance system.

9 WORKS CITED

Drozda, J., Zeringue, A., Dummitt, B., Yount, B., & Resnic, F. (2020). How real-world evidence can really deliver: a case study of data source development and use. *BMJ Surgery, Interventions & Health Technologies*.

Gawande, A. (2010). *The Checklist Manifesto: How to Get Things Right.* New York, NY: Metropolitan Books.

Gottlieb, S., & Shuren, J. E. (2018, November 20). *Statement from FDA Commissioner Scott Gottlieb, M.D. and Jeff Shuren, M.D., Director of the Center for Devices and Radiological Health, on FDA's updates to Medical Device Safety Action Plan to enhance post-market safety.* Retrieved from https://www.fda.gov/news-events/press-announcements/statement-fda-commissioner-scott-gottlieb-md-and-jeff-shuren-md-director-center-devices-and-2

International Epidemiological Association. (2008). *Dictionary of epidemiology* (5th ed.). (M. Porta, Ed.) New York, NY: Oxford University Press.

Knepper, T. C. (2018, May 9). When will clinical trials finally reflect diversity? *Nature,* (557), 157-159.

McClure, D. L., Raebel, M. A., Yih, W. K., Shoaibi, A., Mullersman, J. E., Anderson-Smits, C., & Glanz, J. M. (2014). Mini-Sentinel methods: framework for assessment of positive results from signal refinement. *Pharmacoepidemiology and Drug Safety*, 3-8.

Robb, M. A., Racoosin, J. A., Sherman, R. E., Gross, T. P., Ball, R., Reichman, M. E., . . . Woodcock, J. (2012). The US Food and Drug Administration's Sentinel Initiative: Expanding the horizons of medical product safety. *Pharmacoepidemiology and Drug Safety*, 9-11.

Ross, J. S. (2015). Strengthening Medical Device Postmarket Safety Surveillance. *JAMA Internal Medicine*, 1350-1351.

U.S. Food and Drug Administration. (2015, April 13). *Balancing Premarket and Postmarket Data Collection for Devices Subject to Premarket Approval.* Retrieved from https://www.fda.gov/media/88381/download

U.S. Food and Drug Administration. (2016, December 14). *Public Notification of Emerging Postmarket Medical Device Signals ("Emerging Signals").* Retrieved from https://www.fda.gov/media/95125/download

U.S. Food and Drug Administration. (2019, June). *Clinical Trial Populations - Eligibility Criteria, Enrollment Practices, and Trial Designs Guidance for Industry.* Retrieved from https://www.fda.gov/media/127712/download

U.S. Food and Drug Administration. (2019, July 8). *Medical Device Reporting (MDR): How to Report Medical Device Problems*. Retrieved from https://www.fda.gov/medical-devices/medical-device-safety/medical-device-reporting-mdr-how-report-medical-device-problems#overview

U.S. Food and Drug Administration. (2019, September 18). *Medical Device Safety Action Plan: Protecting Patients, Promoting Public Health.* Retrieved from https://www.fda.gov/about-fda/cdrh-reports/medical-device-safety-action-plan-protecting-patients-promoting-public-health

CONTACT INFORMATION

For more information, please contact NESTcc at nestcc@mdic.org

NEST
Coordinating Center
An initiative of MDIC

www.nestcc.org 202-559-2938 nestcc@mdic.org