Terms of Reference for PulseNet USA

1. SCOPE:

   The National Center for Emerging and Zoonotic Infectious Diseases (NCEZID), Centers for Disease Control and Prevention (CDC) and the participating PulseNet USA (PulseNet) state, county, and local laboratories agree that it is in their mutual interest to develop a joint cooperative program to perform standardized DNA fingerprinting of foodborne disease-causing bacteria so that the PulseNet participants will be able to rapidly compare the whole genome sequencing (WGS) DNA fingerprints, of infectious agents from any laboratory participating in the PulseNet network. This will facilitate rapid and early identification of disease clusters and outbreaks at the national level and, when coordinated with PulseNet International networks, will serve as an effective global early alert system for foodborne disease outbreaks. PulseNet may be extended to cover surveillance for other pathogens and share real-time sequence analyses with systems such as the National Antimicrobial Resistance Monitoring System (NARMS) and the Antibiotic Resistance Laboratory Network (ARLN).

2. PURPOSE:

   This Terms of Reference document (TOR) will define the rights and responsibilities of each participant in the PulseNet network.

3. DEFINITIONS:

The following terms/abbreviations will be used throughout the document:

3.1 APHL: Association of Public Health Laboratories

3.2 CDC/ATSDR: Centers for Disease Control and Prevention/Agency for Toxic Substances and Disease Registry

3.3 DDBJ: DNA Data Bank of Japan

3.4 DNA: Deoxyribonucleic Acid

3.5 DNA fingerprint – A subtype result obtained by a standardized DNA molecular method, e.g., WGS, which has been adopted by the PulseNet network

3.6 EDLB: Enteric Diseases Laboratory Branch

3.7 ENA: European Nucleotide Archive

3.8 FDA/CFSAN: Food and Drug Administration/Center for Food Safety and Applied Nutrition

3.9 GenBank: Open access sequence database produced and maintained by NCBI. Contains all publicly available annotated sequence assemblies and their protein translations.

3.10 MDR: Multi-drug Resistant

3.11 Metadata: a set of data that describes and gives information about other data

3.12 MOU: Memorandum of Understanding

3.13 NCBI: the National Center for Biotechnology Information at the National Institutes of Health

3.14 Participant: An individual at a Participating Laboratory who engages in PulseNet activities

3.15 Participating Laboratory: A federal, state, county, local public health, state agricultural, veterinary or environmental laboratory that engages in PulseNet activities.

3.16 PFGE: Pulsed-field Gel Electrophoresis

3.17 PulseNet Steering Committee – A committee made up of selected PulseNet participants and the Association of Public Health Laboratories (APHL)

3.18 SOP: Standard Operating Procedure

3.19 SRA: the Sequence Read Archive at NCBI

3.20 TOR: Terms of Reference

3.21 USDA/FSIS/OPHS: United States Department of Agriculture/Food Safety and Inspection Service/Office of Public Health Science

3.22 WGS: Whole genome sequencing

4. BACKGROUND:

Foodborne infections are a national and global problem. Trade of raw and processed foods across borders within and between different regions of the world and international travel make it possible for the source of a foodborne infection to be located in a different country, state, or region from where the illnesses are observed. International trade and travel have grown rapidly

during the past decades resulting in foodborne infections appearing increasingly in parts of the world that differ from their origin.

A critical component in the investigation of foodborne outbreaks is the DNA fingerprinting of the causative organisms and comparison of the DNA fingerprints of strains isolated from ill persons who are considered part of a common source outbreak and the possible sources throughout the food chain. CDC began setting up the PulseNet network in 1996 to facilitate rapid, standardized PFGE DNA fingerprinting of foodborne pathogens by state and local public health, and food regulatory laboratories in the United States with data submission to a central electronic database at CDC for purposes of making the database of fingerprints available to all PulseNet participants. In 2019, WGS replaced PFGE as the primary method for comparison of most bacterial foodborne pathogens; WGS is now the “gold standard” for active laboratory-based surveillance of all organisms tracked by PulseNet. PulseNet’s transition from PFGE to WGS was a collaborative effort involving many public health laboratories that provided insight on how to best standardize and implement protocols and tools to analyze WGS data. For example, the *Listeria monocytogenes* pilot study initiated in 2013 provided valuable insight on the development of feasible WGS implementation strategies for the organisms tracked by PulseNet: Shiga toxin-producing *Escherichia coli* (STEC), *Shigella* species*, Salmonella* species*, Listeria monocytogenes, Vibrio* species*, Clostridium perfringens, Clostridium botulinum, Yersinia* (non-pestis) species, *Campylobacter* species and

Cronobacter species. Along with the full implementation of WGS, an updated data sharing policy was implemented to ensure that public health scientists within and outside the network could use the data effectively while protecting patient confidentiality and avoiding inappropriate use of preliminary or potentially sensitive public health data. This policy includes the public release of some PulseNet data. The primary drivers and benefits to the public release of this subset of data include:

1) To conform with the President's Open Government Directive and the CDC/ATSDR Policy on Releasing and Sharing Data.
2) To provide basic information to the food industry on the genotypes and frequency of foodborne pathogens infecting humans, so the information could be used in combination with industry's own testing data to assess the possible risk posed by their products.
3) To provide researchers in microbial phylogenetics, molecular epidemiology and food microbiology access to data useful in their research endeavors.

In addition, the large amount of raw sequence data produced by PulseNet requires long-term storage within high-capacity data repositories in the public domain. PulseNet sequence data is stored in the SRA and in the GenBank database at NCBI with copies in the ENA and the DDBJ. The metadata to be shared with the WGS data is presented in more detail in Section 6.B.5 and Appendix A. The signatories of this ToR are expected to upload their sequence data and the associated metadata to NCBI.

The PulseNet network has enabled CDC and its partners to detect and investigate outbreaks of foodborne infections in humans across the United States and internationally in collaboration with public health colleagues within PulseNet International.

The objective of PulseNet is to use DNA fingerprinting to characterize foodborne pathogens in order to reduce the number of national and global foodborne outbreaks and to perform surveillance of different foodborne pathogens in humans and throughout the food chain. An ultimate goal is to help decision makers in establishing policies for safer food on the national and global level.

5. GOVERNANCE:

5.1 Activities of the PulseNet network are coordinated by the PulseNet USA Steering Committee. The PulseNet Steering Committee is chaired by the Chief of the CDC EDLB or a designee and is comprised of participants from state and local PulseNet laboratories, state agricultural, veterinary and environmental laboratories and federal regulatory (USDA/FSIS/OPHS, FDA/CFSAN) laboratories, APHL and CDC.

5.2 Membership on the PulseNet Steering Committee is by invitation only; terms on the Committee are open-ended. The Committee meets via conference call as necessary, generally once per quarter. The structure and membership of the committee is subject to change upon agreement of the current members.

5.3 Decisions regarding topics brought before the committee are reached by consensus.

6. ROLES AND RESPONSIBILITIES OF PARTICIPANTS:

6.A. Coordination and collaboration relative to public health activities

It is mutually agreed that:

6.A.1 Each participating laboratory will utilize the expertise, resources, and relationships of the network in order to increase capability and readiness to respond to outbreaks as outlined in the PulseNet SOPs. In addition, each participating laboratory will designate central point(s) of contact for communications dealing with matters covered by this agreement.

6.A.2 One or more participants from each participating laboratory will attend periodic joint national and regional conference calls, and in-person meetings and conferences to promote better communication and understanding of regulations, policies, and responsibilities and to address questions and issues that may arise through routine or critical network operations.

6.A.3 One or more participants from each participating laboratory will maintain knowledge of improvements to the communication methods, data utilization tools, and methodological developments within the network.

6.A.4 Each participating laboratory will notify and collaborate with the other participating laboratories as soon as possible when issues or investigations of mutual concern become evident. Such collaboration may include providing alerts to the other participating laboratories regarding disease outbreaks

encountered as part of its activities, providing technical advice in areas of recognized expertise, providing results of analysis, and exchanging information, including identification at the state or local level of genetically highly related DNA fingerprints, development of a subtyping method that potentially may be implemented in the network or legal changes at the local level which may have consequences for the network as a whole.

6.A.5 This agreement does not preclude the PulseNet participating laboratories and/or participants from entering into other agreements which may set forth procedures for special programs which can be handled more efficiently and expertly by other agreements.

6.B. Principles and Procedures for the Exchange of Information

It is mutually agreed that:

6.B.1 Although there is no legal requirement for exchange of information, the PulseNet participating laboratories/participants agree in principle that there should be a presumption in favor of full and free sharing of information between them. The participants recognize and acknowledge, however, that it is essential that any confidential information that is shared must be protected from unauthorized public disclosure. Safeguards are important in order to protect the interests of all parties, trade secrets and confidential commercial information. Patient identifiers, other protected health information, privileged and/or pre-decisional agency records, and information protected for national security reasons shall not be shared without authorization. Such safeguards help guarantee the participant's compliance with applicable laws and

regulations. A participant should only decline to share information if credible information exists indicating that the requesting participant may not be able to comply with applicable laws or regulations governing the protection of non-public information or with the principles/procedures set forth in this TOR.

6.B.2 Each participating laboratory will, to the extent permitted by law, keep all information received from other participants confidential unless the information is already publicly available or written consent for distribution has been received from the originating laboratory.

6.B.3 All participants must implement appropriate data and information security systems. To facilitate the sharing of information, the participants must implement procedures to ensure, at a minimum, that such sharing of information is indeed appropriate and that the recipient protects the confidentiality of all non-public information received. Document control procedures should also be implemented for the storage of any hard copy PulseNet data.

6.B.4 All PulseNet participating laboratories/participants shall limit the dissemination of shared information only to those participants, internal agency offices, and/or individuals that reasonably require the information and will be responsible for ensuring that the information is not disclosed to other recipients.

6.B.5 Sharing of WGS data and associated metadata optimally involves public disclosure of as-detailed-as-possible data while protecting patient confidentiality and avoids inappropriate use of preliminary or potentially

sensitive public health data. An anonymized isolate identifier, a taxonomic description of the isolate (genus, species), project name (PulseNet or GenomeTrakr), the serotype of the organism (if applicable), the source type (i.e., clinical, food, or environmental), the year of isolation, and the country of origin along with the sequence information will be released to NCBI Biosample (metadata) and SRA (sequence metadata and reads) repositories for all isolates as soon as WGS data is available (see Appendix A). For non-clinical (food, environmental, animal) isolates, additional metadata to be released include the exact source site, the host type (animal isolates only), isolate name alias (PulseNet key), the state of origin, the month of isolation, and the name of the submitting laboratory. PulseNet participants may choose, subject to local restrictions, to release additional metadata for their locally sequenced isolates to NCBI BioSample.

6.C. Standardized Protocols, Reference Strains and Quality Assurance

6.C.1 Each participant agrees to follow the standardized PulseNet SOPs for subtyping of foodborne pathogen isolates, as distributed and/or published by CDC. Laboratory-specific changes to PulseNet SOPs (e.g., use of automation) are allowed as long as the laboratory validates the change against the PulseNet SOP to demonstrate that the sequence data quality and downstream analyses are comparable.

6.C.2 Each participant agrees to utilize the customized PulseNet data management and analysis SOPs for WGS and other PulseNet subtyping data to be uploaded to the PulseNet national databases, utilizing the

recommended version of the analysis platform and the analysis guidelines distributed by CDC.

6.C.3 Each participating laboratory should establish a culture collection containing strains of specific interest to their laboratory, e.g., specific serotypes, outbreak strains or MDR isolates. The participants agree to share these strains with each other for network purposes at no cost or at the cost of shipping and handling.

6.C.4 Each participating laboratory agrees to participate in the PulseNet certification and proficiency testing programs designed to ensure comparability of data generated by all participants as outlined by the PulseNet SOPs.

7. PULSENET NAME AND LOGO:

7.1 The PulseNet name and logo are trademarks owned by CDC. Signatories to this TOR may use the PulseNet name and logo as needed in PulseNet activities such as conference presentations.

7.2 When a participating laboratory exercises its option to terminate the PulseNet MOU, that participating laboratory immediately loses the right to use the PulseNet name and logo.

8. SYSTEMS AND DATA SECURITY MEASURES

8.1 All participants will install and configure system equipment capable of running the latest recommended version of the PulseNet Customized analysis

platform. The participating laboratories will maintain data security following the direction of the PulseNet Technical Steward for information technology and database security, including such mechanisms as storage in a locked room or locked file cabinet, shredding of discarded documents, and memory erasure upon replacement of hard drives. Participants will not share logon identities or passwords, that have been assigned by CDC.

8.2 All participants who have access to the PulseNet listserv (SharePoint or the most current communication mechanism) will sign a non-disclosure statement assuring that all information will be treated as confidential and only shared with appropriate public health personnel or others as may be required by law.

Appendix A

Minimal set of Metadata to be released immediately to NCBI's GenBank or SRA with a WGS:

Metadata released immediately with the WGS:

Clinical (human) isolates:

- Strain: Unique sample ID (anonymized WGS_ID)
  - Created specifically for the purpose of NOT including information about isolation date or state of origin
- Organism Group: genus and species (e.g., *Listeria monocytogenes*)
- Serovar: Predicted serotype (if relevant)
- Project name: Submission to NCBI on behalf of (PulseNet or GenomeTrakr)
- Source name/type: Organism source (e.g., human, food, environmental)
- Geographical origin: Country of origin (USA)
- Collected by: Identity of the laboratory submitting DNA sequence ('CDC' if the participating laboratory cannot disclose its identity)
- Collection Date: Year of collection of the sample from which the isolate was obtained

Additional metadata required for non-clinical (food, environmental, animal) isolates

- Isolate name alias: PulseNet Key
- Isolation Source: Exact source site (type of food or environmental sample, sample site for animal sample)
- Geographical origin: formatted USA:State (e.g., USA:GA)
- Host: Animal type for animal samples
- Sequenced by: Identity of the sequencing laboratory
- Collection Date: Year and month of collection of the sample from which the isolate was obtained

This set of minimal metadata should be reviewed internally and may be expanded based on each state's data release policies.

The following PulseNet metadata should never be released to NCBI:

- Cluster codes
- Allele codes

If a state regulation or policy does not allow for the disclosing of all the agreed upon metadata presented here, the state point of contact should discuss other options with CDC/EDLB at pulsenet@cdc.gov.

_____

Signature Date

_____

Name, Affiliation, PulseNet laboratory