Principles of Data Sharing Checklist – v1.1 14FEB2024

National Institutes of Health (NIH) data sharing guidance (https://sharing.nih.gov/data-management-and-sharing-policy/about-data-management-and-sharing-policies/data-management-and-sharing-policy-overview) and the NIH RFA for the Rare Diseases Clinical Research Network (RDCRN) (RFA-TR-18-020) emphasize the need for data sharing that is not limited to members of a consortia, but broadly available to the community. This checklist outlines the basic elements that need to be incorporated in the informed consent language, the contractual language established between the consortium administrative core and its sites, and the consortium policies that govern data sharing, use, and publication within the consortium and with external users. The principles stated in this checklist are incorporated and further elaborated in the RDCRN Guidance for the Development of Data Sharing Policies by Individual Consortia.

1. Informed Consent for every protocol should include language that allows broad sharing of the data while protecting the confidentiality of and minimizing the risk for re-identification of the participant. Example language can be found in the RDCRN Guidance for the Development of Data Sharing Policies by Individual Consortia. The Informed Consent should also describe:
    a. Sharing data with the Administrative Core of the RDCRC as a limited data set,
    b. Sharing data with the Data Management and Coordinating Center (DMCC),
    c. Sharing data with other researchers for future studies,
    d. Allowing the transfer of data, without identifiers, to a federal data repository maintained by the NIH,
    e. Options for restricting data sharing that the investigators deem necessary to offer the participant (e.g., as mandated by the IRB).
2. A subcontract between an Administrative Core (AC) and clinical sites must include language that:
    a. Makes explicit that each site agrees to enter data in a data capture system maintained by the DMCC unless otherwise specified by the sponsoring NIH institute, acknowledging that the data entered into the DMCC may meet the definition of a limited data set,
    b. Allows the AC to collate participant data in the form of a limited data set from all protocols that the site participates in,
    c. Gives the AC the authority to distribute and share data with the DMCC and the NIH,
    d. Gives the AC the authority to share the data with other third parties as governed by specific data use agreements,
    e. Gives the AC the authority to transfer the data, without direct identifiers, to a Federal data repository maintained by the NIH.
    f. References the Data Management & Sharing Plan as written in the grant application and any Data Sharing Policy between consortium sites and the administrative core.
    g. *Suggested language*: Pursuant to this agreement, as outlined in the Data Management & Sharing Plan of award [NIH award number] [Site] will enter [RDCRC] data in a data capture system hosted in the NCATS cloud and maintained by the DMCC on behalf of the [RDCRC] unless otherwise specified by the sponsoring NIH institute. The Administrative Core of [RDCRC] has the authority to collect information from [Site] in the form of a limited data set for the purpose of combining the data across the sites of each protocol. The Administrative Core has the authority to manage the data with the

support of the RDCRN DMCC and further share the data with third parties as governed by data use agreements that the Administrative Core will prepare and execute according to the policies of the [RDCRC]. Further, the Administrative Core has the authority to transfer [RDCRC] data to Federal data repositories to fulfill their obligation to the sponsor(s), in doing so, the Administrative Core will utilize the services of the RDCRN DMCC.

3. A Data Sharing Policy between all participating sites of an RDCRC should be consistent with the RDCRN Guidance for the Development of Data Sharing Policies by Individual Consortia and policies published by the RDCRN and the NIH. This policy should facilitate data sharing and outline the mechanisms of data sharing within the RDCRC and with external partners. The policy should clearly state the intention to transfer patient-level information, stripped of identifiers, to a federal data repository.
4. A Data Use Policy, consistent with the guidelines published by the RDCRN, that includes a publication policy outlining how data collected in RDCRC studies is to be used, disseminated, and attributed.
5. To the extent that the DMCC is going to support the RDCRC data sharing efforts and prepare RDCRC datasets for submission to a Federal data repository(ies), ensure that appropriate legal agreements are in place to allow the DMCC to use and further share the data (execution of Data Use Agreement).
6. Each research participant's informed consent selections pertaining to data sharing and future data use should be tracked in the research database in order to facilitate the designation of an appropriate consent group when the data is submitted to a Federal data repository.