PHIT Partnership Data Collaborative

Principles, Policies, and Procedures

This document has been drafted by the Doris Duke Charitable Foundation with input from the PHIT Partnerships and the Data Coordinator

The policies and guidelines for the Population Health Implementation and Training (PHIT) Partnership Data Collaborative (referred to below as the PHIT Collaborative) have been crafted by the Doris Duke Charitable Foundation with significant input from the PHIT Partnerships and the Data Coordinator. It respects the needs of each individual PHIT Partnership team while promoting cross-site analyses, shared learning, and dissemination of information to both the national governments where the projects take place and the larger global health community.

The following principles guide the operation of the PHIT Collaborative:

- **Respect** for data and sample ownership by the individual Partnerships;

- **Rapid** dissemination of results;

- **Rights** of the collaborating investigators, especially students, fellows and junior faculty involved in the design and execution of the project and in the collection and analysis of the data to publish their results; and

- **Recognition** that this is a multi-site initiative with a number of questions that can best be answered through cross-site analyses.

I.   Purpose of the PHIT Collaborative and its Activities

The Doris Duke Charitable Foundation (DDCF) African Health Initiative aims to help catalyze a shift from the current public health focus on single-disease programs to an emphasis on strengthening health systems to effectively deliver integrated primary care to underserved populations. Each PHIT Partnership funded by the African Health Initiative is working in a defined region to accomplish the following goals:

- Significantly and measurably improve health;

- Strengthen the health systems in the region; and

- Increase the knowledge for evidence-based health delivery and health systems planning by conducting implementation research.

Because generating new knowledge of global significance is a central goal of the African Health Initiative, DDCF has created the PHIT Collaborative to foster global learning through Partnership activities. The intention of the PHIT Collaborative is to provide a forum for sharing ideas and generating new knowledge for the field. In order to accomplish this it is expected that the PHIT Collaborative members will participate in the following activities:

- Attend annual meetings, participate in relevant working groups, and communicate regularly, both electronically and in teleconferences;

- Define, collect and share data from all Partnerships that have been jointly agreed upon as "core metrics" and "documentation", and possibly also "common metrics" (described in detail below);

- Participate in cross-site research projects and training activities when feasible and relevant;

- Work to facilitate cross-site data analysis, interpretation, dissemination and use in improving public health policies and programs;

- Publish jointly authored articles in the peer-reviewed literature when appropriate.

II.   PHIT Collaborative Members

Data Coordinator - DDCF has contracted with Robert Black and Jennifer Bryce from the Institute for International Programs at Johns Hopkins University to serve as the Data Coordinator for the PHIT Collaborative. In this role, Bob Black, Jennifer Bryce and their colleagues will work with the PHIT Partnership grantees to assure comparable, high quality core and common metrics[1] and documentation collected by each PHIT Partnership, to manage access to these data and to promote cross-team learning. The five specific functions of the Data Coordinator are to:

- Ensure standardization of cross-site data collection and analysis methods;

- Monitor data quality;

- Facilitate core data sharing;

- Promote cross-site learning; and

- Facilitate the dissemination of AHI findings.

 Note – The Data Coordinator's role is limited to cross-site activities and does not include responsibility for the quality or technical rigor of studies conducted on topics that do not address the PHIT Collaborative core metrics or documentation to be shared across PHIT partnerships.

The Data Coordinator's role in the PHIT Collaborative is a technical and consultative one, not a directive one. As a consultant to DDCF, they will work closely and cooperatively with DDCF staff and participate in the Collaborative Management Committee (see below) to ensure that the main programmatic goals of the Initiative are being met.

---

[1] The 'core metrics' are defined as a set of quantitative indicators that will be measured by each PHIT Partnership. These measurements will be maintained in PHIT Collaborative databases designed to support cross-site analysis and interpretation. A set of 'common metrics' that are measured across at least two but not all of the sites will also be developed and may or may not be incorporated into the PHIT Data Collaborative Database set. See "Cross-Site Data Collection and Scope of Work for the PHIT Data Collaborative" for more details.

PHIT Partnerships – Each PHIT Partnership team has a co-Principal Investigator (co-PI) at the sponsoring institution and a co-PI in each of the countries in which they work. It is expected that all of these individuals will participate in the leadership of the PHIT Collaborative activities. In addition, each PHIT Partnership team is composed of many outstanding individuals with diverse expertise in the conduct and analyses of research and other relevant areas and it is expected that, whenever practical, those individuals also will participate in PHIT Collaborative work.

DDCF – The Director of the African Health Initiative and other foundation staff will participate in most PHIT Collaborative activities as observers, but the DDCF will not be a voting member of the Collaborative Management Committee (see below).

III.   Governance and Management

A Collaborative Management Committee (CMC) will be formed and consist of a representative from each PHIT Partnership and a representative of the Data Coordinator, each of whom will have one vote. It is hoped that decisions regarding the CMC's recommendations to the foundation will be made largely by consensus. When possible, co-PIs should participate in the discussions of the CMC. Such participation is meant to assure that country perspectives are well represented, as the co-PIs from the sponsoring institutions are all US nationals. Although the CMC voting members will number five, CMC meeting participants may number as many as 10, depending on the additional participants from each PHIT Partnership. It is expected that the CMC will elect a Chair from one of the PHIT Partnerships that will rotate after a term agreed upon by the CMC.

The CMC will make recommendations to the foundation regarding a range of issues involving cross-site activities, including:

- The release of core data to other members of the PHIT Collaborative to pursue jointly agreed upon analyses;

- The public release of baseline and final core data sets as well as documentation and common metrics where appropriate;

- The process and procedures for the public release of data;

- The rank order of the priority cross-partnership research questions proposed by each PHIT partnership that will be pursued by the PHIT Collaborative;

- Cross-partnership training opportunities and opportunities for program-based/shared learning;

- Allocation of extra DDCF funds (expected to be about $150,000/year beginning in 2010) to support cross-team activities;

- Establishing the guidelines for cross-partnership publications, authorship, and acknowledgements.

It is expected that, when there is a consensus, the foundation will follow the recommendations of the CMC. However, if there is a disagreement among CMC members, the foundation will exercise its right to be the final arbiter. Details about how the CMC review and recommendation processes will be carried out will be worked out by the CMC together with the foundation. The CMC may decide, for efficiency purposes, to form sub-committees, which may include additional members from the PHIT Partnerships with appropriate expertise, to deal with separate issues such as training opportunities or publications. In the case of urgent matters that may arise, the Chair of the CMC may convene an emergency teleconference. A member of the Data Coordinator team will facilitate the CMC's work by ensuring that information is shared in a timely fashion, meetings are convened and meeting agendas are created with input from CMC members. It is expected that the CMC will meet at least 3 times a year (once in person, scheduled to coincide with the Annual PHIT Partnership meeting, and the other two or more times[2] times by conference call).

IV.   Processes and Guidelines for Shared Data

A. Establishment and Management of the Collaborative Database

A data repository will be created to house de-identified data from all of the Partnership sites. The core metrics and documentation data collected by Partnership teams will be provided to the Data Coordinator *at least annually*, and stored in a set of computer-based data files maintained either by the Coordinator or a specialized technical group contracted for this purpose. This latter decision will depend upon the volume and complexity of the data bases and the anticipated needs of the PHIT Collaborative.

Individual level data records from Partnership sites will be kept in site-specific files and maintained using a single file format with full documentation. Additional quality assurance checks and audits will be performed by the Data Coordinator. Any questions will be addressed through consultation with the PHIT Collaborative members from the individual Partnership sites. Each Partnership will be able to download its data from the central databases.

When summary measures or core metrics are provided, the accompanying documentation will include sufficient information so that the metrics could in theory be reproduced. Documentation should include a brief description of the analytic approach, the name and date of creation of the source data file(s), and computer code that computes the metrics or calls standard available routines that perform the computations.

The Partnership data managers will have primary responsibility for quality control, trouble shooting, data collection, and cleaning and uploading the data to the central databases. Any questions concerning quality control or technical assistance in these tasks will be addressed through regular telephone contact and the annual visits to each site by the Data Coordinator.

The Data Coordinator will seek to ensure that the data provided by each Partnership meet minimum quality criteria for cross-site analyses. The Data Coordinator is not responsible for the quality or technical rigor of studies conducted on topics that do not address the PHIT Collaborative core metrics or documentation.

---

[2] When issues arise, unscheduled meetings may be requested by the CMC, the Data Coordinator or the foundation.

In order to stimulate cross site sharing and analysis, the CMC (the PIs and the Data Coordinator) will both generate and review concepts for cross site analyses. At the direction (in writing) of the CMC chair, the Data Coordinator will create needed files of relevant de-identified data and provide these to the designated analysts. Procedures for review, data analysis, and drafting Collaborative manuscripts will be developed by the CMC.

B. Privacy and Confidentiality

All PHIT Partnerships as well as the PHIT Collaborative are expected to assure the privacy of individuals from whom data have been collected, whether community residents or health workers. Thus, it is expected that prior to uploading any data to the PHIT Collaborative, all identifying personal data will be removed and every effort will be made to avoid the inadvertent identification or likely potential for identification of individual study subjects, health workers, or facilities. Individual level data will be provided where privacy can be assured. As part of assuring privacy, the smallest geographic unit identified in the repository will probably be the district (rather than sub-district or village), although sub-district data may be collected and used by the individual teams in their analyses. The Data Coordinator will not have access to the identifying information or the key that can link the de-identified data sets to its personal identifiers.

C. Data Sharing, Release and Access

Members of the PHIT Collaborative acknowledge the need to engage closely with regulatory and research institutions in countries where each of the Partnerships work. Priority is placed on respect for national guidance on matters of data access. Further, the PHIT Collaborative will work to assure that there is maximum data access and dissemination of findings in the countries and settings where the data collection occurred.

It is expected that the PHIT Collaborative will make their results publicly available, largely but not exclusively to the wider scientific community through peer-reviewed publications in high-impact peer-reviewed journals.

In addition, prior to receiving DDCF PHIT Partnership grants each of the Partnerships agreed to abide by the "Policy on Data Management and Sharing for the Population Health Implementation and Training (PHIT) Partnerships", reproduced in Annex A for easy reference. This policy requires that each PHIT Partnership submit core metrics along with the relevant documentation to enable usage of the data from a central data base. Further, the agreement outlines a commitment to public access to these data within an agreed timeframe. This timeframe is intended to:

- Ensure the teams that collected the data have sufficient time to analyze, write-up and obtain acceptance for publication of their main study findings;

- Provide the PHIT Collaborative with sufficient time to analyze, write up and obtain acceptance for publication agreed-upon cross site analyses;

- Assure that the larger public health community- in countries where data were collected, Africa more generally, and elsewhere- has timely access to these data so that further

analyses can advance understanding of ways to strengthen health systems for the delivery of comprehensive primary care.

Individual site and cross-site publications are expected to adhere to basic principles of reproducible research.[3] In the case of core and common metrics, the Data Coordinator will facilitate this by maintaining careful records of data collection methods submitted by the PHIT Partnerships so that a fully documented data base will be available for public release. For both the individual teams and the PHIT Collaborative, the data collected and its documentation should be sufficient to reproduce the main published analyses, and include a minimal data set, coding manual, and human-readable computer code.

Baseline data collected by each Partnership will be the first data available for cross-site analysis in the Collaborative Database. Assuming that baseline data collection is completed in August 2010, the expectation is that these data will be finalized within the following 6 months. After a review of these data by the Data Coordinator, the data will be declared cleaned and closed. It is expected that the final data will be shared with the national governments prior to their public release. During the next two-year period, cross-site data analysis as agreed upon by the Collaborative Management Committee (CMC) will be pursued as well as any site-specific analyses. At the end of this two year period, estimated to be February 2013, the baseline data sets from all the PHIT partnerships will be made available to the public. The table below describes the expected timeline for the release of the baseline data.

TASK	**TIME LINE**
Baseline data collection complete	Month 0
Partnerships clean and check their data; review data with government authorities per national protocols	Month 0-6*
Baseline data submitted to Data Collaborative	Month 7
Data reviewed by Data Coordinator	Month 8 – 9
Data available for cross-Partnership analyses	Month 13-30
Baseline data shared with public	Month 31

* It is acknowledged that timeline for the availability of data from DHS surveys depends on when each country completes the cleaning and release of their DHS data.

A similar timeline will apply to the uploading of mid-term data to the Database. However, it is expected that mid-term data will be released to the public along with the final data set after the end of the PHIT Partnership intervention. Six months will be allowed for cleaning the data at the site and closing the Collaborative Database, followed by 2 years for analysis by the PHIT Partners, both individual and collaboratively. It is likely that none of the intervention and control

---

[3] Peng RD, Dominici F, Zeger SL. Reproducible epidemiologic research. Am J Epidemiol. 2006;163(9):783-9.

site data for a PHIT partnership will be released to the public until two years after the completion of the partnership's intervention.

In preparing data for public release, the following processes will be used:

- The primary data records, including codes that would allow individual identification, will be collected, owned, controlled, and stored by the individual PHIT Partnership teams. The Data Coordinator will not have access to individual identifiers or the key that can link the de-identified data sets to its personal identifiers.

- Partnerships will upload data on core metrics and documentation to the Collaborative databases in accordance with the standardized data collection and management methods agreed upon by the PHIT Collaborative, in order to support multi-site data analyses in accordance with the specific aims of the AHI.

- Specific guidelines on publication, authorship and acknowledgement will be drafted by the Data Coordinator and shared for review, becoming final upon approval by the entire CMC.

- Following review for intellectual property, confidentiality and privacy considerations, and following the embargo period, the PHIT Collaborative will make data publicly available[4].

The Data Coordinator will not release any data, including to other Partnerships, without prior review and written approval by the CMC.

D. Cross-Partnership Publications and Presentations

Cross-site publications will be reviewed by the CMC, which may wish to convene a separate Publication Committee for this purpose. Four publication and presentation types are anticipated from the studies conducted by the Partnerships:

Type	Description
a	General overview of the cross-site framework and the set of core and common metrics to be included in the Collaborative databases. (DDCF expects that a synopsis of the "white paper" will be an early outcome of the foundation's initiative.)
b	Results from the comparative analysis of Partnership data (involving at least two Partnerships and addressing selected priority research questions).
c	"Add-on" analyses (other than those addressing the priority research questions) drawing on data from two or more Partnership sites.
d	Results based on data from a single Partnership.

---

[4] Publicly available data are defined as data which are made available by the CMC to individuals who are not members of the PHIT Collaborative through a process to be specified by the PHIT Collaborative working with the Doris Duke Charitable Foundation.

Publication Guidelines for PHIT Partnerships:

1. All publications and presentations at major international meetings resulting from the work of any PHIT Partnership should be shared with the local and national governments in advance of publication or presentation and copies of the communications should be sent to DDCF and the Data Coordinator.

2. Authorship of all publications and abstracts should reflect the genuine contribution of the collaborating investigators in accordance with best academic publication standards and accepted practices. The PHIT Partnerships aim to increase capacity for research and its publication in the countries in which they work, and it is expected that the authorship of publications will reflect this goal.

4. Manuscripts, abstracts and public presentations that fall within types a through c above should be shared with the Collaborative. For example, if two sites plan and publish or present a collaborative study, it is preferable that that Collaborative is notified of the pending publication in advance. In that way, any concerns about pending or actual publications that fall within types a through c can be considered and resolved by the CMC.

5. The CMC is not responsible for coordinating or overseeing type d or Partnership-specific studies. It is desirable that the Collaborative share its data analysis plans, so that the Collaborative members are aware of similar or parallel analyses underway among the sites, thus stimulating identification of additional areas of cross-site collaboration.

6. All publications and presentations will use the standard acknowledgement statement provided in the PHIT grant agreement: "This work was supported by the African Health Initiative of the Doris Duke Charitable Foundation."

7. The PHIT Collaborative will adhere to the principle and policy of public access of all publications. The policy ensures that the public and broader scientific community has access to the published results of the funded research. It requires scientists to submit journal articles that arise from the AHI to the digital archive PubMed Central (http://www.pubmedcentral.nih.gov/). The policy requires that these articles be accessible to the public on PubMed Central to help advance science and improve human health. This policy is consistent with the policy of the US National Institutes of Health which is available at http://publicaccess.nih.gov/.
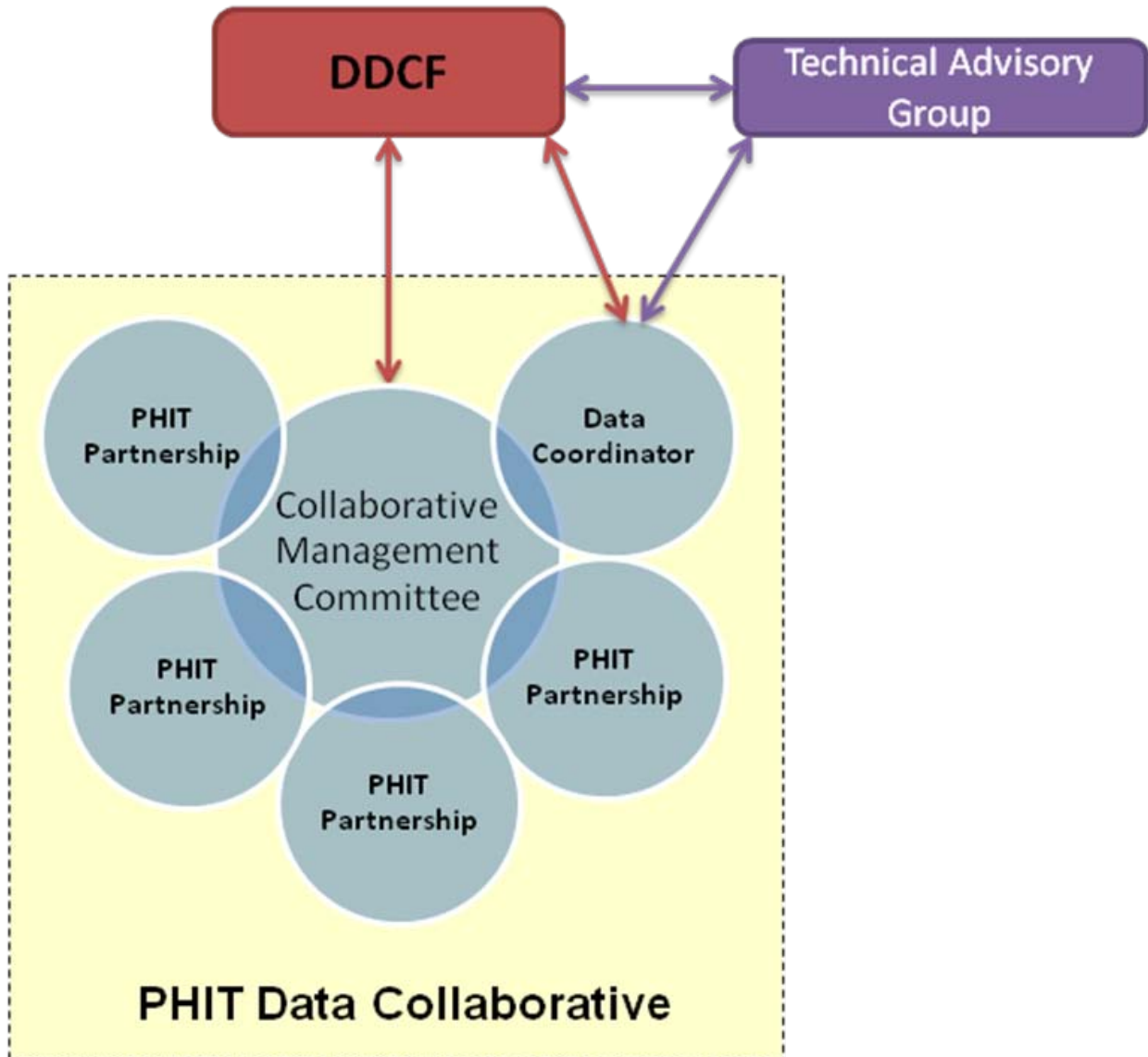
V.   Role of Technical Advisory Group

The role of the Technical Advisory Group (TAG) is to provide independent, high-level, technical feedback to the foundation and the Data Coordinator on issues relating to the evaluation and impact assessment of the Population Health Implementation and Training (PHIT) Partnership interventions. Technical issues of concern to the foundation include rigor and suitability of the research design, quality and validity of data collected, alignment of the approaches used with national needs as well as global health priorities. The TAG will be composed of up to 6 members to serve 3 year terms or longer if possible. Figure 1 below illustrates how the TAG will interface with DDCF and the Data Coordinator.

The TAG will provide high level advice to DDCF, as well as to the Data Coordinator on technical matters, including, but not limited to:

- PHIT Partnership study design and implementation;

- Relevance of the core and common metrics being collected by the PHIT partnerships;

- Adequacy of documentation;

- Approaches to measuring cost effectiveness of the PHIT interventions;

- How to best assess the impact of each PHIT partnership and the initiative as a whole;

- How to align the AHI with other global health activities.

Figure 1: Data Oversight & Management Structure

Annex A: Policy on Data Management and Sharing for the Population Health Implementation and Training (PHIT) Partnerships

The African Health Initiative seeks to catalyze significant advances in strengthening health systems by supporting partnerships that will design, implement and evaluate large-scale models of care that link implementation research and workforce training directly to the delivery of integrated primary healthcare. One important aspect of the African Health Initiative is to ensure that information collected by the PHIT Partnerships, and by the team managing a program-wide evaluation, is managed in a way that maximizes public benefit. In support of this, this data sharing policy document has been developed to articulate the goals and the fundamental philosophy of the role and the use of data in the PHIT Partnerships.

All of the information which is regularly collected area-wide by all of the PHIT Partnerships (referred to as core data) will be submitted in a timely manner to a data repository and made available to all other PHIT Partnership members. That core data should include, but is not limited to demographic, health status, and health care quality data. All data held at the repository will fully protect individual privacy—of patients, community members and providers. Where appropriate, data may be aggregated or locations concealed to avoid inadvertent identification of families or communities. It is the position of the Foundation that the release of individual provider-specific data (such as mortality rates by clinical officer) is a government policy decision which should not be undertaken independently by the PHIT Partnerships.

Not all information collected and maintained by a PHIT Partnership need be submitted to the repository. For example, information used to monitor and improve primary care sites can be of great value locally but of little use to other researchers in the absence of a detailed description of the context. Even here, however, researchers are encouraged to submit the data to the repository for sharing once a report on the effort has been published.

We expect that most information will be made generally available to potential users including other researchers, individuals designing programs, and those involved in policy development, after a set period of time (~12 to 18 months). Some types of sensitive information may be held for longer periods. Deadlines for data submission will be an integral part of PHIT Partnership grants. The final proposals should indicate the individual PHIT Partnership's plans and time lines for cleaning data before submission. Where data is used which is also reported to local government, a reasonable and pre-agreed upon period for government review should also be included in the timeline and included in an MOU between the partnership and the public authority. These time lines are expected be developed and incorporated into the workplan that is developed during the planning grant phase. An adequate data sharing plan and the timely submission of data will be an important aspect of both the grantee selection process and continued funding.

In support of this data sharing policy, the Foundation will develop a data collaborative that will include representation from all grantees and from expert evaluators to address policy questions as they arise.