Traditional healers working with primary care and mental health for early intervention in psychosis in adolescent: intervention development and feasibility study

1. Description of the data

1.1 Type of study

This is a mixed method study. The study will include an ethnographic and qualitative methods (focus group discussions and semi-structured interviews) and a feasibility cluster RCT.

1.2 Types of data

The following types of data will be collected;
(1) Quantitative data will include data on outcome measures included in the cluster RCT, and participants' knowledge and satisfaction with training(pre and post completion training activities) which is part of process evaluation
(2) Qualitative data from participant interviews and focus group discussions
(3) Participant-observation data, from focussed ethnography studies

1.3 Format and scale of the data

Data will be stored in multiple formats to provide redundancy and resilience of data and facilitate data sharing. These include: CSV exports, SQL Script, SQL Server Databases, Statistical software (e.g. Stata/MP14.1 for Windows), QSR NVivo 12

2. Data collection / generation

2.1 Methodologies for data collection / generation

Data on outcome measures for the cluster RCT will be collected by experienced research assistants at baseline and 12 weeks in a health centre which is not part of study, to maintain the blindness of assessors to the allocation status. The data will be entered into a specifically designed study database.

Qualitative data will be collected during in-depth face-to-face interviews and focussed group discussions. Interviews will be audiotaped and transcribed verbatim.
Ethnographic data will be collected by experienced ethnographers who will participate in daily communal life as well as observing interactions with traditional healers, healthcare workers, patients and carers in everyday encounters and settings. Observation data, ethnographers' field diaries and interview transcripts will be imported into the qualitative data-analysis software QSR NVivo 12.

2.2 Data quality and standards

The outcome data for cluster RCT will be collected with the help of previously validated instruments, which we have used in our ongoing research in Peshawar.
Data will be double entered by specially trained staff into a database specifically designed for this study.
Qualitative data will be collected by experienced social science researchers who will have received additional training to meet the requirements of this study and will be regularly supervised on site by a local co applicant (Ali) and through video conference by UK co-applicant (Dikomitis).
Data quality will be continuously monitored with feedback provided to interviewers to maintain data quality. A data monitoring committee will be established to provide governance over data collection, handling and analysis processes.

3. Data management, documentation and curation

3.1 Managing, storing and curating data.

Data will be stored on a secure server in Khyber Medical University (KMU), which is designed to hold clinical study data and is being used in another MRC funded study on the site. The data is backed up regularly. The access to the server is strictly password protected. A data dictionary will be created to technically define the databases; its tables and structure. Combined with annotated and coded case report forms this will provide a curated set of data.
A Data and Safety Monitoring Board (DSMB) will be established for oversight and monitoring of the conduct of the trial. DSMB will meet three times a year to ensure the safety of participants and the validity and integrity of the data. All adverse reactions and serious adverse events (SAEs) reported spontaneously by the subject or observed by the investigators or other staff members will be recorded by the research team and reported to the DSMB. SAE requiring urgent attention will be reported to research clinician in first instance, who will ensure participants safety.

3.2 Metadata standards and data documentation

Along with the data dictionary that will be created, which will define the schema of the data. Coded/Validated CRFs and PROMS will also to be stored and available for other researchers to use and working in line with the MRC's Good Practice Principles for Sharing Individual Participant Data From Publicly Funded Clinical Trials.

3.3 Data preservation strategy and standards

All anonymized electronic data files relating to the study will be stored in a secure location in Khyber Medial University secure server and held for 10 years. Electronic document and data files (including files relating to data analysis) will be stored as read-only files in a dedicated archive folder with restricted access on a dedicated location within the secure KMU site. Data in the Masterfile will be stored as CVS and SQL Script to make they are kept in a format which is independent of software or platform.

4. Data security and confidentiality of potentially disclosive information

4.1 Formal information/data security standards

KMU maintains high standards of data security. This will be regularly monitored by the Keele University and we will aim to maintain compliance with Information Governance polices and standards at Keele.

4.2 Main risks to data security

Data will be stored in an anonymised format to ensure participants cannot be identified. Participant identity will be protected via the use of a unique study code to pseudonymise their data. If personal data is to be collected this will be stored in a logically separate location from CRF/PROMS data with limited access for the site PI or the persons authorised by the PI and all such access will be logged. Such as data will be deleted as soon as feasible. Access to the data will use two factor authentication (ensuring that people accessing data are individually identified and approved). Data will have audits applied, which logs who, what and when changes were made in addition to the value before it was changed.

5. Data sharing and access

Once data collection has been completed, all data will be maintained in such a form that they cannot be linked with identifiable participants. The trial statistician at Keele will have the access to the anonymised data.

We will follow the best practice in data sharing which is currently in practice at the School of Primary, Community and Social Care, Keele University. The practice is in line with the MRC's Good Practice Principles for Sharing Individual Participant Data From Publicly Funded Clinical Trials.

5.1 Suitability for sharing

The data collected in this study is suitable for sharing and the study team is committed to maximising the impact of this work through data sharing and future collaboration. We will collect data in three key domains; socio-demographic, health status and clinical

measurement. Data will be entered and securely stored in electronic format making it suitable for secure transfer to other research groups with allied and complementary interests and expertise.

5.2 Discovery by potential users of the research data

We will promote discovery by potential users of the research data in a number of ways. We will publish an open access version of the research protocol at the start of the study, providing details about the design and conduct of the study and full details about the data collected. This will contain specific detail on how to collaborate with the research team. We will further promote the use of the data through social media (dedicated Facebook page and Twitter feed), at conferences and through further peer-reviewed publications.

5.3 Governance of access

We will establish an access committee to consider all formal requests for access within an appropriate timescale (4 weeks), including new uses by the study team itself (beyond the funder- approved programme). The committee will consist of independent global health researchers, primary care researcher, biostatisticians and social scientists, with other expertise drawn on as needed. The committee will be advised by the PI (Farooq) but the study team will not be formal members.

5.4 The study team's exclusive use of the data

The study team, working with key stakeholders in Pakistan (including clinicians, patients and healthcare providers) will prioritise research questions raised at a multidisciplinary workshop. Analyses and proposed publications resulting from these will be pre-planned and published in the study protocol. These analyses will fit with research expertise of the study team. We will encourage other potential users to contact the team during the data collection stage to promote and support early data sharing. Data sharing will be supported as soon as data is ready for analysis.

5.5 Restrictions or delays to sharing, with planned actions to limit such restrictions

As part of the consent process, proposed procedures for data sharing will be set out clearly and current and potential future risks associated with this explained to research participants. The consent for further contact will be sought to enable future intervention or follow up studies. There will be no time restriction on releasing anonymised data.

5.6 Regulation of responsibilities of users

External users will sign a data sharing agreement, detailing the responsibilities of the study team and external user. This will follow MRC data sharing guidance.

6. Responsibilities

The PI (Farooq) will have overall responsibility for the study. Within the study partnership Lewis (quantitative) and Dikomitis (qualitative) will have responsibility for study wide data management, data security, quality assurance of data and metadata creation.

7. Relevant institutional, departmental or study policies on data sharing and data security

Data Management Policy & Procedures

Data Security Policy https://www.keele.ac.uk/it/itpoliciesandprocedures/Keele%20University%20Information%20Security%20Policy%20version%201.0.pdf

Data Sharing Policy https://www.keele.ac.uk/pchs/keelectu/

Institutional Information

Policy

https://www.keele.ac.uk/researchsupport/researchdatamanagement/

8. Author of this Data Management Plan (Name) and, if different to that of the Principal Investigator, their **telephone & email contact details**

Professor Saeed Farooq