1.0 Prerequisites

Prior to verification, participants will consent to the study and complete their user profile. Active recruits will have the opportunity to begin their questionnaires after completing the user profile and prior to verification; passive recruits will not be able to proceed with any study activities between user profile creation and verification.

2.0 Purpose

The purpose of this document is to provide standard procedures for the verification of participants for the Connect for Cancer Prevention Study.

3.0 Scope

This SOP describes the data that will be used to verify active and passive recruits; the automated process by which recruits will be verified; how a successful match will be defined; the plan for cases in which automated matching fails; the protocol and timeline for manual matching; the validation data to be sent to NCI; and the timing of the validation reporting from the site to NCI.

4.0 Definitions

4.1 Active recruits:

Patients of the partnering health care system (Site) who have been determined eligible and invited to enroll in Connect using a unique token/PIN.

4.2 Passive recruits:

Individuals who enroll in Connect without having been pre-screened by the Site, determined eligible, or assigned a unique token/PIN prior to enrolling.

4.3 Successful match/match threshold:

- For the Active recruits, the participant NCI unique token, Last name and DOB should all match exactly for automated review. Otherwise, will be moved to manual review.
- For the Passive recruits, First name (full legal first name or preferred first name), Last name, and DOB should all match exactly.

4.4 Local Site CRM:

All the information on the current and future eligible population is stored into the local site CRM (REDCap) database from EMR/ Enterprise Data Warehouse for the study.

5.0 Procedures

5.1 Data required to verify recruits (Connect -> Site)

- First name (full legal first name or preferred first name) **only used for passive recruits*
- Last name
- DOB
- NCI participant token **only used for active recruits*

5.2 Automated verification process

5.2.1 System Pre-Processing

The data items defined for verification are imported from NCI through the getParticipants?type=notyetverified Connect API for participants who need identity verification. This information will be imported from NCI on-demand whenever [redacted] study staff open or refresh the Identity Verification report in the site REDCap CRM. The following pre-processing steps will be performed before the match verification process starts:

1. Trim white spaces
2. Convert unique code char to ASCII, if any
3. Text case conversion – to lower case letters
4. Duplicate participant record check - This process does two things and sets an additional data flag in REDCap, as needed:

    - Checks if REDCap has duplicate (more than one) record based on participant core demographic identity variable such as First Name, Last Name, DOB retrieved from NCI. This process helps identify possible duplicate records created part of Enterprise system and data flow into REDCap. The program sets a data flag (iv_local_site_crm_dups=1) if duplicate record(s) found in REDCap and will be taken care by 5.5 Resolving duplicate process.
    - Checks for any existing participant record found in the identity verified status population in REDCap based on participant core demographic identity variables: First Name, Last Name and DOB retrieved from NCI. This process helps to identify duplicate participant enrollment in the MyConnect App . The program sets a data flag (iv_nci_dups=1) if duplicate record(s) found in NCI web site which is limited to Site level data set check and will be taken care by 5.5 Resolving duplicate process.

5.2.2 Process for active recruits

The automated identity verification batch job will run every 15 mins which will import data from C4CPDS through getParticipants API and complete the pre-processing steps.

After pre-processing, the batch job performs automated match check process to verify the identity as defined as a successful match (Token, Last Name, DOB for active recruits & First Name, Last Name, DOB for passive recruits). If the successful match is found, the participant verification status is updated to 'Verified' (197316935) in REDCap and sent to NCI through the identifyParticipant API for the individual participant. No manual intervention is required. If a successful match is not found, it will fall into manual identity verification workflow queue and display in the custom site CRM dashboard for manual identity verification. See section 5.3

[redacted] will audit the automatic verification process using manual verification periodically throughout the study. We will do a qualitative check to determine if any changes to the automated verification processes are needed.

5.2.3 Process for passive recruits

The automated identity verification process for passive recruits is the same as active recruits, with the addition of updating the NCI assigned participant token in REDCap.

5.3 Manual verification process

5.3.1 Process for active recruits

The research study staff will use custom REDCap for identity verification, which lists all the participants with 'not yet verified' status and 'match not found' status. The Manual Verification Queue will be actively monitored and handled during standard working hours and any verification failures will be acknowledged within two business days. The research study staff will use the table below in order to make informed determination of participant verification.

Manual Verification Steps	Possible Outcomes
Step 1: The research study staff search for the participant in REDCap using any of the information contained in the User Profile, including address, email, phone number.	A. **VERIFICATION CANNOT BE CONFIRMED:** If an obvious error was made in First Name (for passive recruits), Last Name or DOB or if a match is suspected based on match of other demographic variables, then proceed to Step 2. Note: If the only difference between the first/last name in the User Profile and the participant record in REDCap is a hyphen or special character (i.e. tilde), but all other demographic variables match, Manual Verification will be completed without necessitating outreach to the individual. B. **INELIGIBLE**: Individuals who do not appear to be associated with [redacted] upon manual review (i.e. there are no possible matching records found in our REDCap

	database) will be marked as 'Cannot be verified' (219863910). Verification status will be sent to NCI through the identifyParticipant API.
Step 2: The research study staff will attempt to contact the participant with the contact information provided in their NCI user profile, per the Outreach Protocol outlined below. See section 5.3.3. The caller will utilize REDCap to assist with making determinations when communicating with the participant and log the contact outcome in REDCap.	A. **VERIFICATION CONFIRMED:** If research study staff is able to confirm the participant identity after the participant contact attempt, they will first update the information found to be incorrect in the User Profile through the NCI Connect Study Manager Dashboard. Then, the verification status will be updated to 'Verified' (197316935) in REDCap. Study staff will also update the intervention type to 'manual' in the identity verification audit log data entry field. Verification status will be sent to NCI through API on-demand when the verification status is saved in REDCap. B. **VERIFICATION CANNOT BE CONFIRMED** If research study staff is not able to confirm the participant identity upon discussion with the individual, participant status will be updated to 'Cannot be verified' (219863910) in REDCap. Verification status will be sent to NCI through the identifyParticipant API. C. **VERIFICATION OUTREACH TIMED OUT** If the individual is not able to be reached, two weeks after the last contact attempt, research study staff will mark as 'Verification Outreach Timed Out' (160161595) status in REDCap and the identity verification status will be sent to NCI through the identifyParticipant API.

5.3.2 Process for passive recruits

The manual verification process for passive recruits is the same as active recruits (above 5.3.1), in addition, the following process is in place:

The REDCap participant record will be updated with the NCI-assigned token if no token exists or if the existing token does not match and a possible identity is determined. If the existing token does not match and the possible identity is determined, a 'Duplicate' (922622075) status flag will be sent to NCI for the existing token. Please refer the attached Appendix 1 for the identity verification workflow diagram.

5.3.3 Outreach protocol

Individuals who are suspected to be associated with [redacted] based on manual review but whose token, First Name (for passive recruits), Last Name and/or DOB do not match with the local REDCap record will be contacted by research study staff though phone call (Appendix 2). Study staff will reach out to these individuals via the phone number indicated in the Connect User Profile record, making up to four contact attempts per individual with up to two voicemail messages. If contact is made and a match is determined, research study staff will mark as 'Verified' (197316935) in REDCap and the verification status will be sent to NCI through the identifyParticipant API. If the individual is not able to be reached, two weeks after the last contact attempt, research study staff will mark as 'Verification Outreach Timed Out' (160161595) status in the REDCap and the verification status will be sent to NCI through the identifyParticipant API. If an individual later responds to the outreach inquiries and a match is determined, their status will be changed from 'Verification Outreach Timed Out' (160161595) to 'Verified' (197316935) and the new identity verification status will be sent to NCI through the identifyParticipant API. If a match is not determined upon contact, the research study staff will mark as 'Cannot be verified' (219863910) status in REDCap and the verification status will be sent to NCI through the identifyParticipant API. The participant will be notified by NCI and can reach out to Connect Support Center if they think this is an error. See Appendix B for verification scripts.

If the only difference between the first/last name in the User Profile and the participant record in REDCap is a hyphen or special character (i.e. tilde), but all other demographic variables match, Manual Verification will be completed without necessitating outreach to the individual.

5.3.4 Outreach timeline

Initial contact will be attempted within 2 business days of identifying an unmatched sample. Contact will be attempted every 2-4 days until all four contacts are made or the case is resolved. Up to two voicemails will be left as necessary.

5.4 Roles and Responsibilities

Table 1.

Personnel	Responsibility
Study staff (Research Techs or Research Specialists)	Check variants of names in site CRM
Study staff (Research Techs or Research Specialists)	Contact participants to verify information

Study IT Support Staff - Programmer	Manage record creation, merging and removal in Site CRM

5.5 Resolving duplicates

The list below of possible causes of duplication needs to be taken into account:

- Enterprise Record Duplication (two different studyIds with same demographic detail)
- Participant duplicate enrollment into the MyConnect app (e.g., participant enrolls once using phone and once using email)

5.5.1. The deduplication process

The custom REDCap report identifies the participant population who has duplicate data flags ((iv_local_site_crm_dups=1) or (iv_nci_dups=1)) set by the system pre-processing program based on patient match criteria, excluding token and PIN data items, to identify possible duplicate participant records. The participants with duplicate flags will not be included as part of the automated or manual identity confirmation process conducted by study staff. Instead the following procedure will be followed:

5.5.1.1 If a duplicate record is found in the Enterprise Record, a case by case manual process will be handled by IT Support Staff to fix the duplication where the same person would be assigned more than one unique study ID. A final decision will be made for record retention and the duplicate record will be marked with a flag for deletion (soft delete). As part of soft deletion, the following variables are updated in REDCap:

- flag_elig = 0 (not eligible)
- leavetype = 6 (Duplicate Record Removed from EHR)
- when_dups_updated = m-d-y
- dups_original_studyid

5.5.1.2 If a duplicate record is found in the C4CPDS , for example, if a participant created a MyConnect app account with email and also another account using phone number, the second entry will be flagged by IT Support Staff as 'Duplicate' (922622075) status and sent to NCI through the identifyParticipant API.

To ensure a complete audit log and data loss prevention, the deduplication process will be using the soft delete data flag and handled through [redacted] IT Support staff.

5.6 User Profile Updates Prior to Verification

Connect Support Center (CSC) may receive some requests from participants to update their user profile details (e.g., name, DOB, preferred email or phone number, etc.). When the CSC receives these requests, they will create a case and reassign it to [redacted]. If the request is made prior to verification, study staff will make the change(s) in the study manager dashboard. This participant's updated information will fall into the next verification batch job. Post verification, user profile updates will be handled by NCI via a prod manual change and [redacted] staff will be notified.

5.7 Verification status reporting (Site -> Connect)

5.7.1 Reporting requirements

The following match variables will be sent to NCI as part of the verification table through the submitParticipantData API:

- First Name Match (for passive recruits) – ConceptId 147176963
- Last Name Match - ConceptId 557461333
- Token Match (for active recruits)- ConceptID 679832994
- DOB Match – ConceptId 725929722
- Site Match – ConceptId 570452130
- Cancer Status Match – ConceptId 547895941
- Age Match - ConceptId 629484663
- Report verification method for each recruit
    - Automated verification method used - ConceptId 444699761
    - Manual verification method used - ConceptId 953614051
- Duplicate type - ConceptId 148197146
- Outreach required for manual verification - ConceptId 188797763
- Update recruit type - ConceptId 793822265

For passive recruits who log in directly to the MyConnect App, the de-identified demographic information, as described in the Recruitment SOP, will be sent along with the verification table, through the submitParticipantData API. For passive recruits that contact [redacted] directly, the de-identified demographic data will be sent in the next batch job following the force generation of the PIN/token.

5.7.1.1 Report verification status for each recruit:

The verification status information is recorded as part of the participant record in REDCap and stored as NCI defined status categories:

197316935 = Verified

219863910 = Cannot be Verified

922622075 = Duplicate

160161595 = Outreach Timed Out

The verification status is reported in real-time through the identifyParticipant API when the verification process is complete.

5.7.1.2 See example verification table here
https://nih.app.box.com/folder/71984684825

5.7.2 Verification status reporting timeline

The verification status will be sent to NCI through the identifyParticipant API in real-time when the study staff record the status in REDCap. This real-time method will ensure that the Site Manager Dashboard is always up to date and in sync with the site REDCap system.

5.8 In-PersonVerification Support

To support on-demand/in-person visit scenarios for verification, the REDCap verification dashboard pulls the participant list (not yet verified) from NCI through the API whenever the web page is refreshed, which includes real-time data. The study staff is able to perform automated or manual verification through REDCap as needed and proceed to further study activity, such as biospecimen collection.

5.9 Verification Status Corrections

After the verification status is changed from Not Yet Verified to Verified, Cannot be Verified, or Duplicate, [redacted] will no longer be able to update the verification status. Only the CCC can change the verification status from Verified, Cannot be Verified, or Duplicate to something else. If [redacted] finds that they made an error in the verification status, they should jointly email the CCC and NORC (NCICohortSupport NCICohortSupport@norc.org) with the Connect ID, an explanation of the error and the proposed resolution of the error.

6.0 Help

If during the course of verification there is an unusual problem or occurrence that is not covered in the procedure, please stop and contact the Connect Coordinating Center:

Email: ConnectCC@nih.gov
Phone: (240) 276-5800

7.0 References to other SOPs

7.1 Connect Data System User Manual

7.2 [redacted] Recruitment SOP

8.0 Revision history

Version	Date	Summary of change
1.0	9/16/2021	Document created
2.0		

Working version: https://app.box.com/file/860023472726

If the SOP needs to be revised, please contact either:

[redacted]

Approval

Digital signature of approver

Appendices

Appendix 1: [redacted] Verification Workflow Diagram

Appendix 2: Identity Verification Scripts

Appendix 1: [redacted] Verification Workflow Diagram

The [redacted] Verification Workflow Diagram can be found here: https://app.box.com/file/835832212152

Appendix 2: Identity Verification Scripts

Initial voicemail AND Final voicemail

Hi, this is message is for [first_name]. This is [name] from [redacted] Institute with the Connect for Cancer Prevention Study. We need to confirm some information with you before you are able to continue in the study. Please call us back at [redacted].

Thank you and have a great day.

Script for phone call [to confirm First Name, Last Name, and/or DOB]

A. Good **[morning/afternoon/evening].** Is **[first name] [last name]** available?

> If no, schedule a call back and end call.
>
> If you have established that you are speaking to the correct participant, say:

B. Hi, this is [name] from [redacted] Institute with the Connect for Cancer Prevention Study. I see that you signed up for the study, but it looks like your [name and/or date of birth] does not match what we have on file. Can I confirm that we have the correct information on file?

> If no, Okay, you will need to confirm your information before you are able to continue in the study. Is there a better time to call back to confirm your information?
>
> If yes, continue to part C.

C. Thank you. [Could you please spell your [first/last] name?/ What is your date of birth?]

> If verification successful, continue to part D.
>
> If verification fail, Okay, according to your information and our records, you are not eligible to participate in this study. We appreciate your interest in this important work. Thank you so much for your time.

D. Great! I was able to find your information in our database. You are all set to continue participating in Connect. You will receive a confirmation email from the study within a day or so with instructions to start your first survey on the MyConnect app, if you have not done so already. [If active recruit] If you have time, you can go ahead and log into the MyConnect app and begin your first survey now.

Do you have any questions for me? We appreciate your participation in this important study.

Thanks and have a great day!