



## Standard Operating Procedures

### 5.5.1 Electronic Data Handling ©2007

History			
Version	Date	Author	Reason
1.1	18 <sup>th</sup> July 2007	B Fazekas	New procedure
1.2	18 <sup>th</sup> August 2007	B Fazekas	Changes ratified by MAB
1.3	16 <sup>th</sup> October 2007	B Fazekas	Update after David Currow review

Approval				
Version	Author	Signature	Approval Name	Approval Signature
1.3	B Fazekas	<i>B. Fazekas</i>	D Currow (CI)	<i>David Currow</i>

#### Scheduled review

**Date** August 2009

**Responsible person** PaCCSC Project Officer

## 5.5.1 Electronic Data Handling

### Purpose

The correct management of research data and files is crucial to the integrity of the final results.

Correct management enables researchers to accurately substantiate publication results, and also meet reporting and auditing requirements. This SOP describes the procedure for data entry and error resolution. A comprehensive manual for CRF completion and data entry will be available for each site to refer to for specific routine site data management.

---

### Other related SOPs

Essential Documents  
Data transfer

---

### Attachments

Data Report Form  
Study Staff Database Access form

---

### Other files that apply

File Master Index.xls  
Master Patient Index.xls  
Forms Tracking Index.xls

---

### References

Note for Guidance on Good Clinical Practice (CPMP/ICH/135/95). Annotated with TGA comments 2000 (accessed 250207)

<http://www.tga.gov.au/docs/pdf/euguide/ich/ich13595.pdf>

NHMRC National Statement on Ethical Conduct in Research Involving Humans 1999 (accessed 250207)  
<http://www.nhmrc.gov.au/publications/files/e35.pdf>

Joint NHMRC/AVCC Statement and Guidelines on research practice 1997 (accessed 250207) <http://www.nhmrc.gov.au/funding/policy/researchprac.htm>

King DW, Lashley R. A quantifiable alternative to double data entry. *Controlled Clinical Trials* 2000; 21(2):94-102.

---

## Definitions

### Source data

Source Data is all information in original records (or certified copies) of clinical findings, observations, or other activities in a clinical trial necessary for the reconstruction and evaluation of the trial. This can include:

- pathology reports
- clinical case notes
- observations sheets
- x-rays etc.

Source data also includes the data recorded within case report forms (CRF's) or contained on audiotapes if these data forms clinical data from which analysis is conducted and not contained within other source documents. For example, if clinical observations are recorded within the CRF and used as trial data, this is then source data.

### Error coding

Data errors can be grouped as being Data Entry errors, or Form Entry errors.

- data entry - where the error is related to the entry of data into the data base
- form entry - where the error is related to completion of the paper CRF (including wrong form use)

In addition the error at data points will be coded according to the impact of an error to the objectives of the study:

#### *Outcome*

Those data points that will relate to the study primary and secondary outcome reporting.

#### *Reporting/Key Performance Indicators*

Data points used in reporting functions such as demographics, and baseline scores, and data collected for maintaining Key Performance Indicators.

#### *Eligibility*

Data to determine the eligibility of the patient for the study from the inclusion and exclusion criteria, but excluding the fields related to establishing the criteria.

#### *Safety*

Data that indicates safety monitoring and the establishment of safety parameters to proceed.

#### *Non-outcome*

This includes all other data collected during the completion of the CRFs including questionnaire data

Each data error may also be identified if the query was related to

1. other data for that patient within the same form or between forms, and

2. where the data in one CRF is consistent with corresponding data within another CRF (for example, if the subject proceeds to participation in the study (as evidenced by further CRF completion), but the inclusion criteria is ticked as not assessed this would be coded as being inconsistent as all criteria have to be met in order to participate).
3. Judgement will be required to determine how one data point relates to other data presented.

Sample Only

**Procedure**

**Data collection**

The points at which specific data are collected is to be specified within the study protocols. All investigations, forms, questionnaires and all other data are to be included, including if these data are described as source data. An example table is given below.

	Time point 1	Time point 2	Time point 3	Time point 4	Time point 5	Time point 6	Study Exit
<b>Investigations</b>							
Haemoglobin							
Blood gas analysis							
Vital signs							
Pulse oximetry							
<b>Forms</b>							
CRF's							
Consent form							
<b>Questionnaires</b>							
Safety monitoring							
Side effects							

**Electronic recording**

Study data will be recorded in a number of files for both the administration of the study and collection of subject data.

1. A master index will contain the confidential subject contact information and will be the only link between individual subjects and the ID number. This will be an Excel spreadsheet (Master patient index.xls). Each site will generate a site specific master index for subjects recruited at that site.
2. The forms tracking index will be identified by ID number only. It will be used to track the data collection forms for each participant for auditing of data collection. It will contain the dates each form is due, entered and finalised. (Forms tracking index.xls). This index will be generated and maintained at the coordinating site.

3. The Data file will be held and administered in the coordinating site, and will contain the subject data as downloaded from the Electronic Data Capture System. This data will then be transferred to the data set for analysis.

### Data entry

Data will be entered from each site into a web-based interface specifically developed for each study. This password protected interface is protected behind a 'Ciskopix' firewall to reduce the risk of unauthorised access. No personal identifying information will be entered on this interface at any time. The coordinating site will download and store the data on a regular basis as a text file. The data can then be transferred to the study statistician using the Data Transfer SOP for conversion into other file types for analysis.

The CRF's will be printed from the on-line entry system for field data collection. Data entered onto the paper forms will be transferred into the web-based data entry form. On completion of data entry for each form, the study site will 'submit' the data, any further changes to the data will initiate an email alert to be automatically sent to both the person entering the data and the Data Manager as part of the monitoring process.

A copy of the original CRF will be forwarded to the coordinating site to enable verification of data entry and filing.

### Caresearch.com.au

This website contains the data for the case report forms and a range of data collection forms specific for each study. Access is allowed via the Project Officer. There are a number of levels of access to the data management system;

1. Administrator access – where very limited number of people have password access to the whole website and its contents, essentially those people who are involved in website design and maintenance
2. Manager access – where users can edit data entered by others, with access to the forms provided by the administrators. This level of access enables data corrections, reporting and download functions
3. project officer access – users are able to access a restricted number of forms (allocated by the manager) and are able to view the CRF's for printing and direct data entry only.

Nominated study staff will have access to enter data only. This will require password access. **This access must not be given to anyone else through the sharing of passwords.**

The National Project Officer will maintain a register of those provided with access to the data entry system (Study Staff Database Access form), along with the date of access and the date access ceased. As a safeguard, this is also stored within the data entry system.

Separate instructions will be developed for the data entry for individual studies and circulated prior to beginning recruitment for each of the studies undertaken by PaCCSC.

### Data verification

The coordinating site will systematically conduct a manual check of the data recorded on the CRF's against the data recorded online to ensure completeness and accuracy of data entry. This will be undertaken;

- For all the files of the first 5 subjects enrolled at each site or until complete data accuracy is achieved for each CRF
- Every fifth CRF thereafter until recruitment ceases
- If **any** data entry errors are encountered during the check of the fifth form, checking will revert back to a check of the next five forms until there are 5 consecutive error free forms (King, 2000).

Any discrepancies between the paper CRF and the electronic data system will be handled using the following review procedure.

### Review procedure

All data queries will be reviewed and coded according to the source of the error, and the likely impact on study data analysis as defined earlier. The codes will then be classified according to the appropriate site of responsibility.

The data source, level of data, the code for each error and the site of resolution is summarised in the following table.

Source	Level	Code	Resolution
Data entry	Outcome	1	Study site
Data entry	Non Outcome	2	Coordinating site
Data entry	Reporting/KPI	3	Coordinating site
Data entry	Eligibility	4 (consistent)	Coordinating site
Data entry	Eligibility	4 (inconsistent)	Study site
Data entry	Safety	5	Coordinating site
Form entry	Outcome	6	Study site
Form entry	Non Outcome	7	Coordinating site
Form entry	Reporting/KPI	8	Coordinating site
Form entry	Eligibility	9 (consistent)	Coordinating site
Form entry	Eligibility	9 (inconsistent)	Study site
Form entry	Safety	10	Study site

The coordinating site will complete a data report for each error or inconsistency:

- The errors requiring correction by the coordinating site will have the data report form completed by the National Project Officer, the data base will be corrected according to the report form, and the form will be filed in the participant folder
- The errors requiring correction by the study sites will be faxed or emailed.
  - On receipt of the error report, the study site will complete or correct the missing data on the original data form, sign and date the change, and fax the corrected page back to the coordinating site.
    - If the error is related to data entry the required change to the data base will be described by the study site on the report,

- if related to a form error, the CRF will be corrected and the report form along with the corrected CRF page will be faxed to the National Project Officer.
- The National Project Officer will make the changes to the data as per the faxed correction or instruction.
- The data will then be logged as complete.
- This correction will generate an automatic email notice of the correction for tracking purposes.

The coordinating site will maintain a register of data checking for auditing purposes. The register will record the date of data entry and checking, the date of return to the study site for correction, the date of return or correction, and the date of resolution. A log will be maintained detailing the corrections required for each data form. The completed query forms will be filed with the participant data.

All unresolved data queries will be completed during site closure visits on completion of the study recruitment.

### Rules

1. Study sites will not be required to correct data related to non-outcome or reporting/KPI. This should reduce the amount of data queries sent to study sites.
2. The coordinating site will **assume that the CRF data is correct** when there is any difference between the CRF and on line versions. This is based on the assumption that data collected and recorded close to data collection is more accurate than data entered some time later. This means that data entered onto the data base that is NOT recorded in the CRF will be deleted where possible on checking by the coordinating site, and will result in inaccurate reporting for that site. This specifically applies to the checklists, which may in turn effect KPI's for that site.
3. Study sites have a responsibility to ensure that the data entered on to the data base matches the CRF to avoid discrepancies as outlined above.
4. The National Project Officer will report error rates to each site and benchmark error rates between sites. This will enable each site to determine the origin of data errors and initiate plans for correction. Ongoing errors will be discussed directly between the site and the Project Officer.
5. All Email records of data entry are to be kept, either in an electronic folder or as printed versions filed in a secure cabinet.

### Data safety

The online data entry system is held within an off-site server. The files are electronically backed-up to tape each night. Tapes are held at the server site for 7 days and then transferred for storage for a period of 7 weeks at a third site.

Data will be downloaded by the coordinating site on the first Monday of each month to a password protected, network backup, desk top computer to further safeguard data. In addition, this data will be copied to disc for storage in a secure environment.

### **Data storage**

All data collected at each site for each participant will be kept in a subject file (identified by ID number only) which will contain the case report forms, any corrected and amended data, copies of adverse event reports etc. All data will be stored at each study site in a locked filing cabinet with all identifying information removed, and independent from the administrative files for the study. All study files will be stored in accordance with the Trial Master File index (Trial Master File index.xls).

All identifiable data (consent forms, pathology reports, etc) will be kept in a separate locked file, away from the study documents during the recruitment period. On completion of the study, these data will be archived with the study material, in a separate file with no link between these data and the study materials.

At completion of the study, all case report forms are to be prepared for collation and archiving by each site. All subject files will be reconciled and stored along with all study materials – both hard copy and electronic – consistent with the jurisdictional regulations regarding the retention and disposal of patient records.



# Data Report Form

Site	
Data entered by	
PID	
Source of Data (CRF)	
Data point	
Question Number	
Question Text:	

Data Entered onto database	
Data Entered onto paper CRF	

### Details of DATA QUERY

Source	Level	Code		Responsibility
Data <input type="checkbox"/>	Outcome <input type="checkbox"/>	1 <input type="checkbox"/>	2 <input type="checkbox"/>	Study Site <input type="checkbox"/>
Form <input type="checkbox"/>	Non-Outcome <input type="checkbox"/>	3 <input type="checkbox"/>	4 <input type="checkbox"/>	Co-ordinating Site <input type="checkbox"/>
	Report <input type="checkbox"/>	5 <input type="checkbox"/>	6 <input type="checkbox"/>	
	Eligibility <input type="checkbox"/>	7 <input type="checkbox"/>	8 <input type="checkbox"/>	
	Safety <input type="checkbox"/>	9 <input type="checkbox"/>	10 <input type="checkbox"/>	

### Action required by study site:

Nil <input type="checkbox"/>	Update original CRF sign and date – fax back <input type="checkbox"/>
Instruct co-ordinating site <input type="checkbox"/>	Other: <input type="checkbox"/>

### Action required by co-ordinating site:

Update database <input type="checkbox"/>	Update CRF <input type="checkbox"/>
Other: _____	Other: _____

Resolved: Sign \_\_\_\_\_

Date \_\_\_\_\_

**Please make correction and fax confirmation back to co-ordinating site.**





**PaCCSC**  
Palliative Care Clinical Studies Collaborative

## Study staff data base access

### **State**

#### **Site**

Name  
Study manual supplied –  
Data base access provided –  
Data base access ceased –

#### **Site**

Name  
Study manual supplied –  
Data base access provided –  
Data base access ceased –

### **State**

#### **Site**

Name  
Study manual supplied –  
Data base access provided –  
Data base access ceased –

#### **Site**

Name  
Study manual supplied –  
Data base access provided –  
Data base access ceased –

Study manual supplied –  
Data base access provided –  
Data base access ceased –