VALIDATION OF STATISTICAL PROGRAMMING

Executive Summary Report

CONTENTS

RESOURCES

The following resources were referenced.

- Good Clinical Practice Guide, Medicines and Healthcare products Regulatory Agency, TSO information and publishing solutions, 2012
- Computerised Systems validation in clinical research A practical guide, 2nd edition, Association for Clinical Data Management
- Guideline for good clinical practice E6 (R2) (EMA/CHMP/ICH/135/1995)
- Statistical Principles for Clinical Trials E9 (CPMP/ICH/363/96)

DISCLAIMER: This summary report is intended for NIHR (funder). The full report will be available on request. Please email C.Gamble@liverpool.ac.uk and/or Sharon.kean@glasgow.ac..uk.

PROJECT LEADS

Project leads:

Professor Carrol Gamble (Head of Statistics at the Clinical Trials Research Centre, the University of Liverpool, and Chair of UKCRC Statistics Operational Group subcommittee)

Sharon Kean (Director of Information Systems, Glasgow Clinical Trials Unit and Chair of UKCRC Information Systems Operational Group subcommittee)

This project was a collaboration between the UKCRC Statistics Operational Group and the UKCRC Information Systems Operational Group.

1. INTRODUCTION

The objective of examining data management and statistics functions during GCP systems inspections is to establish that the processes in place give assurance that the trial data are collected, managed and analysed to give accurate and credible trial results.[1] Following the introduction of Good Clinical Practice[2] in to UK law there has been increased focus on development of compliant systems with the introduction of risk-proportionate approaches seeing an increased role for statisticians throughout the clinical trial life cycle. However, there is little content within GCP that relates directly to the practice of statisticians and in particular translates into statistical programming responsibilities directly.

In October 2014, by invitation, an MHRA inspector attended a UKCRC registered CTU Statistician's Operational Group Meeting (statisticians network meeting). The presentation focused on aspects of an inspection relevant to statisticians highlighting key GCP inspection findings in non-commercial CTUs Box 1.

Box 1: Key areas of concern

a) Specification, production and control of the randomisation schedule/code
b) How statisticians obtain the data for analysis
c) Validation of the analysis programming
d) Insufficient documentation of statistical processes
e) Data security - considered excellent protection in data management, but complete lack of control in statistics
f) Unclear processes for data management end and statistical analysis commencement
g) Inadequate or poorly documented validation of statistical programming for tables, figures, listings, standard macros (validation).
h) No recommendations in place for "good programming practice"
i) No control over hard coding.
j) Inability to link output used in report/publication to programming output – reconstruction of process not possible
k) Overwriting output /datasets with subsequent runs or system updates

The presentation sparked concerns among senior CTU statisticians around the expectations of the inspectorate and resources required to meet them. The issues raised have become increasingly poignant with CTUs undergoing inspections reporting an increased focus on statistical programming and its reproducibility.

1.1. Objective

The key objective is to identify the core requirements and efficient implementation of the infrastructure required to facilitate regulatory compliant statistical programming practices.

1.2. Aims

1) To identify core requirements for statistical programming processes. These core requirements will be translated in to a set of risk proportionate standards to streamline programming, increase efficiency, aid cross cover between statisticians, and enable more efficient review and approval.
2) Identify options for implementation of the requirements considering costs of platforms and staff resources (both statistical and information systems) during trial setup and conduct outlining the advantages and disadvantages of each approach.
3) Develop guidance to support statistical programming that will support long term validation for transport between trials achieved by developing an agreed consistent program structure and methodology for performing common tasks
4) Promote development of code libraries within CTUs by ensuring code is written to allow validation to be maintained within application across trials. Determine the role for such libraries between CTUs.

1.3. Methods

The delivery of this project was centred on the engagement and collaboration of the UKCRC Statistical Operational Group and the UKCRC Information Systems Operational Group. The aim of the operational themed groups is to share best practice and to help develop standard approaches to common issues.

In April 2015 the Chair of the Information Systems Operational Group attended a statistician's network meeting to discuss statistical programming. Attendees interested in contributing to the project were asked to share relevant materials, including related inspection findings, Standard Operating Procedures (SOPs), statistical programming documents developed in-house or otherwise they felt were relevant, relevant publications.

The information from several meetings was then used to draft the guidance and a workshop held in Glasgow April 2018. Attendees for this small workshop (n=15) were selected to ensure representation from statisticians (n=7), information systems (n=5), and quality assurance managers (n=2), and a network

representative. Consideration was also given to ensuring variety in CTU selection (n=7) in terms of their trials portfolios.

Following discussion at this workshop the guidance was developed in further detail. The second larger scale workshop (n=47) was held in June 2018 in Leeds. The workshop aimed to gain wider perspective and discussion on the guidance and its implications for delivery. This workshop supported attendance from a larger number of CTUs (n=35) and importantly engaged with the MHRA inspectorate to create dialogue around the issues identified. Attendees at this meeting also included the UKCRC registered CTU Quality Assurance Operational Group Chair, the Chair of the Data Sharing task and finish group, a member of the UKCRC registered CTU network secretariat and an IS consultant leading an NIHR project across the CTU network with previous experience of IS management within a large UKCRC registered CTU. The guidance was then supplemented with the discussions and suggestions made.

2. STATISTICIANS AS PROGRAMMERS

Historically statisticians are generally self-taught programmers. Their educational background focuses on the theory and application of statistical methods. While statistical software is used the focus is on the correct application of the statistical methods and the output rather than the programming used within its application. As a consequence the structure and code used to produce analyses between statisticians varies greatly even when using the same programming language. When Quality Control procedures are used, for example independent programming by another statistician, the emphasis is on obtaining the same result rather than on producing a validated program.

2.1. Risk proportionate approach

While the MHRA state that 'it is essential to apply QC checks to the statistical analysis process to ensure that the output is accurate', it is also stated that 'it is acceptable for the QC checks to be undertaken on a risk-based approach with the detail and level of checking varying depending on the item being checked.'[4]

An ‘in-process’ assessment of performance and compliance can be considered an appropriate approach with this real-time assessment building quality into processes rather than relying solely on an end-product assessment. Statisticians often build in checks on data management/derivation coding, that is while writing the program they check the output to ensure it is performing as expected. However, there is often a lack of documentation to evidence these checks.

Documentation of QC activities is a requirement[5,6] and may take various forms, with the QC activity, *where possible*, being performed by a second individual who is independent of initial activities. The level and type of QC and acceptable forms of documentation should be considered within the risk assessment of the trial. Elements relevant to statistical programming include: use of bespoke vs off-the-shelf products, electronic transfer of data, and statistical analysis software and programs.

In considering validation of statistical programming the following factors should be considered:

Do the programs need to be flexible and fluid? For example, is it likely that emerging issues, and additional checks may need to be added in as time goes on? Is there a high risk that the programs will need to evolve with the trial for example in response to changes to the database, or protocol amendments?

What is the frequency with which the programs purpose needs to be repeated? For example, complex checks and recruitment monitoring may need to be undertaken monthly while other programs may only be required annually or only planned for a single time point.

Is the content fixed and predetermined? Most statistics programs are predetermined in response to the risk assessment, monitoring and analysis plan or in discussion with the Data Monitoring Committee. However, during the life cycle of the trial unanticipated analyses may be required or requested in response to those previously undertaken.

How likely is there to be a need to replicate program output? Some programs may be at higher risk of a replication requirement during an audit. For example programs used to generate DMC reports or final analyses.

What is the complexity of the programming and data management steps involved? Where determination of an event requires derivation from multiple variables across repeated time points the programming required will have additional complexity. However, simple derivations may be low risk but their simplicity means that the QC activity would be particularly quick to undertake.

What is the impact of an error? Errors within the randomisation programs or on the primary outcome or safety maybe considered to have greatest impact.

The above considerations may start the formulation of a risk proportionate approach to the validation of statistical programming.

2.2. Validation

A risk assessment should determine the extent of validation required. If more complex or bespoke programming is to be used then a more extensive approach may be required.

If external vendors are to be used then they need to be assessed. Reputation based selection will not be sufficient and vendors should be willing and able to provide documentation to support appropriate validation.

2.3. Capacity implications

Considerations around randomisation programming and access control imply that multiple statisticians per trial are required. The key recommendations driving this are:

- Require a formal process to ensure that the statistician responsible for analysis of a blinded trial does not have access to the randomisation schedule until after analysis populations had been decided (assessment of protocol/GCP deviations).
- Recommended that all interim unblinded analysis reports are produced by a separate team of statisticians and programmers from those who undertake the final analysis[1] (Source MHRA grey book)

There are a number of statistical roles that need to be fulfilled within any clinical trial and consideration needs to be given to their supervision and checking. The number of statisticians required to conduct a clinical trial may be reduced by:

1) the involvement of programmers rather than statisticians, although statistical oversight would still be required.
2) the timing of activities such as whether they can be conducted and concluded prior to access to the randomisation schedule for the purpose of an interim analysis
3) whether the activity can reasonably be conducted using a dummy randomisation list.

3. VALIDATION OF IT INFRASTRUCTURE IN A STATISTICAL ENVIRONMENT

3.1. Installation Qualification (IQ), Operational Qualification (OQ) & Performance Qualification (PQ)

The concepts and requirements for IQ, OQ, PQ in relation to computerised systems for clinical trials have seemed daunting to academic clinical trials units regarding their implementation and the level of documentation required in order to comply with regulations and to satisfy regulatory inspections. In general, it is considered that the IS team within most trials unit has the skillset and resource to perform these validation steps.

For CTUs to be compliant and have evidence of IQ, OQ, PQ the processes to be followed must be contained in a standard operating procedure (SOP). During an inspection, for each system, evidence must be provided. For statistical packages installed at server or a desktop level the manufacturer provided tools should be used.

It is our recommendation that the statistical analytical platform is centrally controlled by the IS group within the clinical trials unit.

It is advisable prior to each main analysis to state the last time validation (IQ, OQ, PQ) was run and to document when the last system update was undertaken as part of the statistical documentation.

4. VALIDATION OF STATISTICAL PROGRAMMING

4.1. What does validation mean to a statistician?

A statisticians understanding of 'validation' may be different to that applied by a programmer or software developer. The focus of statistical education and experience is the understanding and application of the statistical methodology with the ultimate goal being to ensure the result obtained is 'correct' considering potential for bias and methodological assumptions.

This result focus may also be evident in the approach to validation such that the actual approach to validation is centred on validation of the result rather than validation of the program used.

There is a fundamental expectation that the programs will need amending at each run and therefore require validation of each usage. However, this isn't how IS would approach it or CTUs with more commercial portfolios and increased efficiency can be achieved if we validate programs such that they can be used again without requiring revalidation.

4.2. Validation of programming

It is essential to apply the QC checks to the statistical analysis process to ensure the output is accurate (source: MHRA grey book) and there are multiple approaches which may be acceptable. QC checks can be undertaken on a risk-based approach, with the detail and level of checking varying depending on the item being checked (source: MHRA grey book).

When using independent programming the starting point of the validation needs to be clear. For example, does this start from extracting raw data and determining membership of the analysis populations or from a later point such as being passed an analysis set? Either may be appropriate depending on the use of previously validated programs prior to the start point.

It is also important to consider the programming outputs and how they generate the report. It may be considered gold standard to use programming that folds out the pre-specified statistical analysis tables straight in to the report removing the possibility of introducing errors from copying and pasting and typos. This approach also eases the ability to demonstrate reproducibility of the analysis. Alternatives are to generate the tables in the programs output which are then copied and pasted in to the report, or finally the values required are identified from the raw analysis output with each value being copied or typed in to the report.

All are acceptable but where there is an increased risk of introducing errors greater care needs to be taken that this is recognised in the QC process, for example comparison of raw output with the populated report, and documented accordingly. While this increases time required and needs to be applied at each run of the program it may be more efficient than folding out tables straight in to the report, particularly if the report is not to be repeated

While CTUs may wish to implement best practice across all trials in their portfolio there may be benefit in considering the nature of the trial interventions and intentions with respect to regulatory requirements, inspections, and marketing authorisation applications.

We recommend a risk-based approach to determining the QC checks required for statistical programming and suggest the following criteria be considered

- Purpose/type of programming
- Impact and likelihood of an error
- Flexible and fluid or fixed and predetermined
- Need to be repeated
- Need to be replicated

4.3. REPRODUCABILITY

It is expected that statistical results can be reproduced. This means that if asked to replicate a result contained in a report that it is possible to do so within a reasonable timeframe. This may be required during a MHRA inspection.

Increased risks of a breakdown in the ability to replicate results have been highlighted by statisticians in response to requests to amend reports with tight deadlines or from changes requested during the peer review process of an article submitted for journal publication.

As a starting point it is important to identify reports that may be used in an inspection for this purpose and whether the potential for the presence of those increased risks.

5. SUMMARY

Statisticians need to be working on an IS infrastructure that fulfils computer systems validation regulatory requirements. CTUs need to acknowledge the cross disciplinary knowledge exchange that compliance requires and provide the resources to support the activity. Procedures should be in place to ensure that there is evidence of this good practice.

There are a number of statistical roles that need to be fulfilled within any clinical trial and consideration needs to be given to their supervision and checking. The number of statisticians required to conduct a clinical trial may be reduced by involvement of programmers; timing of activities regarding access to the randomisation schedule; or the use of dummy randomisation lists.

A risk proportionate approach to validation is recommended where the risks associated with each statistical program and the impact considered are documented and mitigated where possible.

The use of shared codes and collaborations between clinical trials units and the introduction to standardised data models and formats would improve efficiency across CTUs for both stats and IS and should support data sharing. Few CTUs currently implement these methods and the barriers, including the expense of introducing them, need to be addressed.