APESS: A Unified Methodology for the Development and Review of Quantitative Analysis SOPs

Anthony Rossini\(^1\), Anne Mounier\(^1\) and Andy Richardson\(^2\)

ABSTRACT

A simple but well aligned methodology for creating and reviewing standard operating procedures and the associated process documentation required to control the generation of quantitative analysis products (TFLs, datasets, modelling simulations etc.) requiring programming and code development is described.

The method is based on the recognition that five key steps can be identified in most, if not all, quantitative analysis SOPs. These are, an Acknowledgment step, - confirming that the work is required, a Planning step – detailing the terms of reference and overall scope and framing for the work, an Execution step – where the plan is turned into reality with the planned deliverables (outputs such as tables, listing, figures, or a complete report, or computer programs, packages, or libraries, are generated), a Storage step – ensuring the outputs are available for future use and that there is a reference location, and a Sharing step – distributing the results for further work (i.e. “APESS”).

The presentation describes the method in detail and shows how it can be used to create a unified set of SOPs supporting the generation of any quantitative analysis product. The value of the approach thereafter in guiding the development of case-specific quality and user manuals will be illustrated, together with its utility in ensuring seamless integration of these clinical study lifecycle operations into the overall SOP and quality manual framework.

INTRODUCTION

Operating in compliance with documented processes and procedures is a key pharmaceutical industry requirement; control, governance, and evidence demonstration is typically implemented using standard operating procedures (SOP) and related documentation. The importance of quantitative analysis deliverables to support decision-making at all phases of a medicine’s life-cycle requires all programming and related tasks are undertaken with a minimal level of quality, and this is established by following written procedures to ensure both compliance and appropriateness of decision-making surrounding the label and registration.

There is an extensive literature discussing both generally and specifically best practice for the structure and development of SOPs, but it focuses is often on compliance, with only a limited investment in addressing operational efficiency. The increasing importance and practical needs of data science groups undertaking the daily work of creating high quality quantitative analysis deliverables requires operational efficiency to be a cornerstone for processes and procedures that reduce repetitive and/or manual steps and focuses the coding effort on ensuing data accuracy and fit-for-purpose outputs. In practice, this often results in procedural documentation sets consisting of numerous individual documents of various types - policies, SOPs, Work Instructions/Practices, Guidelines etc. – with many versions reflecting departmental or individual nuances.

Similarly, there are many texts describing proven business process improvement methodologies which are regularly adopted to support new systems/application development and implementation within pharma. Example methodologies include six-sigma\(^1\), Lean\(^2\), BPMN\(^3\) etc. When used in support of statistical programming operations, these methods are usually employed to define processes to be implemented as new software functionality, and their outputs are often then incorporated (or re-presented) in terms of the adopted SDLC methodology (e.g. a business flowchart being re-stated as use cases).

We describe a methodology for creating quantitative analysis deliverables that supports meeting the dual data science procedural goals of regulatory compliance and operational efficiency, through the use of a using a unified
approach. Using this methodology, Novartis has streamlined its quantitative analysis documentation set, reduced the effort required to confirm compliance requirements (audit, inspection), and reduced the training burden, along with making compliance a natural by-product.

THE APESS FRAMEWORK

The APESS methodology was developed to reduce and rationalize the procedural documentation for creating the varied quantitative analysis deliverables generated as part of clinical trials, safety monitoring, trial design simulations and business ad hoc reporting. Quantitative analysis deliverables here include study tables, figures and listings, safety reports, dataset creation or re-structuring - essentially anything that involves the analysis/manipulation of quantitative measurements and calculation, often categorized as statistical analysis, data analysis, data mining, data visualization, computer programming etc.

The acronym “APESS” comes from the early recognition that the creation of most, if not all, quantitative analysis deliverables follow five common steps: Acknowledgment, Planning, Execution, Storage, and Sharing – hence “APESS”. The value of adopting this unified structure is described below, in practical terms, for assuring regulatory compliance, and in terms of business/operational efficiency. Some commonly met ‘variations on the theme’ are also described illustrating how each step might be important to the final outcome or require minimal or extensive effort.

ACKNOWLEDGE

The first step in the creation of any quantitative analysis deliverable is to acknowledge that it is required, i.e. confirming the work is needed. Acknowledgement here can range from the associate simply admitting to themselves that they have taken on the activity with the intent of completing it, to the formal approval from a senior group member granting permission to undertake the activity. Data science groups who develop quantitative deliverables range from operational groups such as statistical programming to innovative and flexible departments including statistical methodology and quantitative modeling groups. This is the first step that should take place before any additional effort is placed into the activity.

PLAN

Once confirmed that work should proceed, the next step is to develop a plan for creating the output – i.e. detailing specification and terms of reference. The plan is essential for establishing a scope and framework for the activity, which is critical to ensure that the activity is time-constrained and for establishing feasibility of the proposal. The plan can be as formal as a Statistical Analysis Plan for a Clinical Study, to a simple paragraph describing at a high level, data sources, data analysis approach, and decisions to be supported. A basic plan requires no more than a few minutes to put together but is invaluable as a sanity check. To establish support and reproducibility, a formal statistical analysis plan document ensures and establishes a minimal level of quality.

EXECUTE

The execution step turns the plan into reality. This is the stage when the plans become computer programs which are executed in order to generate the deliverables which are at an appropriate level of quality to support the intended business process.

STORE

Once the required quantitative deliverables are created and confirmed they need to be put into storage, ensuring that the outputs are available for future use. This step is critical in order to assure reproducibility and discoverability of the work, both important controls that need to be managed and maintained.

SHARE

Dependent upon the detailed reason for creation the final step, which may or may not immediately follow the early steps, is to distribute the results for use within the wider organization or beyond. Ideally this is done by sharing the location, though some systems and procedures will support simple distribution through channels such as email or business-supporting social network and document management systems.
APESS IN PRACTICE

REASONS FOR ADOPTION

The critical importance of quantitative analysis deliverables at every phase of a medicines life-cycle requires that each table, figure, listing, analysis or report needs to be shown to (a) been created and used in compliance with written procedures, (b) are developed using validated tools and methods, and (c) have been confirmed as ‘fit-for-purpose’ (i.e. meeting the specification).

The adoption of the APESS methodology has enabled a consistency in programming and analysis documentation to be achieved when both developing new documentation and reviewing and revising existing documentation.

DE NOVO DOCUMENTATION DEVELOPMENT

The following example illustrates the use of the APESS method to develop new quantitative analysis procedural documentation. This was required for two key reasons: first, due to the introduction of a new statistical computing environment, existing documentation was needed to be withdrawn and replaced, and secondly, to use the opportunity to improve operational efficiency by adopting a smaller streamlined SOP set.

The principal objectives for a new quality manual (SOP) documentation set was:

- to remove all references to system or software from all documentation subject to audit or inspection
- to remove all references to job titles and replace with generic roles applicable to all quantitative analysis operations
- to assure regulatory compliance as the procedural documentation is followed
- to serve as step-by-step instructions for quantitative deliverable creation
- to significantly reduce documentation maintenance requirements
- to remain flexible and responsive to new or revised quantitative deliverable requirements
- to adhere to Novartis’ regulatory documentation policies and requirements

As is general within the industry, Novartis implements a regulatory document framework that recognizes the operational need for different document types, These separate key regulatory compliance documents (SOPs) from key working documents and guidelines. User manuals and training materials supplement these to support day-to-day operations. In appropriate cases, a generic SOP or business process operating procedure is employed to describe a required high-level over-arching process.

Figure 2 shows how the APESS methodology supports the regulated document structure. This grid serves to guide the development of specific documents with the broad expectation (but not mandated), that each document type will focus primarily on describing the APESS steps shown in yellow, in general at the top level and then with increasing granularity moving down. By combining this with the generic flow in Figure 1 a clear association between both the document types, their expected content, and the process steps is created.

This can be further strengthened by adopting a hierarchical referencing convention for procedure steps. Figure 3 shows, organized by APESS category, the ‘numbering’ between the SOP steps and an associated work instruction. The SOP procedure is following the convention outlined in of Figure 2, with an acknowledgement step (Step 3.1), planning step (Step 3.2) and an execution step (Step 3.3). The related work instruction reflects and extends the SOP numbering (Step 3.2 and Steps 3.2.1, 3.2.2, …), starting with step 3.2 since, in this example, it does not have any specific acknowledgement step. This approach establishes an unambiguous traceable compliance relationship between the documents and the steps and ensures that following the work instruction steps is de facto in compliance with all associated documentation, meeting the regulatory compliance and step-by-step objectives established earlier.

Quantitative analysis tasks rely on computing systems and software. Meeting the goals of having procedural documentation that is system agnostic, and recognizes only functional roles, rather than job titles relies on

(a) replacing any reference to specific software or systems by a generic term, and
(b) developing a set of organizational functional roles (as small as possible) that are recognized by all users as the core set required for quantitative analysis work.

The first is relatively easy to achieve – replace any specific system name with ‘validated computing environment’. This comes however with a definition and training burden that must be addressed (e.g. a ‘validated computing environment’ is defined as, “any authorized software or system used for the specification and creation of quantitative analysis deliverables”; the training burden being to explain that ‘yes, your system falls into this category’).
Developing a core set of functional quantitative analysis roles can be more organizationally challenging. Whilst standard industry practice may suggest roles such as ‘programmer’, analysis of the fundamental functional roles contributing to quantitative analysis deliverables delivery suggests that four are required. These are shown in Table 1 together with their procedural function and some examples of typically associated job titles. Formal and informal use case “testing” using these roles has not revealed any situation where additional roles are required.

<table>
<thead>
<tr>
<th>APESS Functional Role</th>
<th>Description</th>
<th>Typical Associated Job Titles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quantitative Analysis Lead</td>
<td>Principal role commissioning work and/or responsible thereafter for deliverables quality and fitness-for-purpose. (APESS:APELLSS)</td>
<td>Lead Statistician, Study Principal Programmer, Epidemiologist, Modeler, Quantitative Safety Scientist</td>
</tr>
<tr>
<td>Quantitative Analysis Contributor</td>
<td>Main role undertaking the planning, creation, confirmation and storage of the required quantitative analysis deliverables (APESS:-PES-)</td>
<td>Trial Programmer, Study Statistician Database Programmer, Data Analyst</td>
</tr>
<tr>
<td>Quantitative Analysis Collaborator</td>
<td>Individuals or groups not directly involved with the creation of quantitative analysis deliverables but provide key, often expert, inputs (APESS:-PE- -)</td>
<td>Clinical Scientist, Data Manager, Pharmacologist, Subject Matter Expert</td>
</tr>
<tr>
<td>Quantitative Analysis Consumer</td>
<td>Individuals or groups requiring the quantitative analysis results for use in further work. (APESS:A---S)</td>
<td>Project Team, Medical Director, Clinical Monitor, External Consultant, Regulatory Agency</td>
</tr>
</tbody>
</table>

Table 1: APESS Functional Roles

This establishes a unified procedural ‘infrastructure’ that meets all the documentation goals above and brings with it considerable savings for document maintenance, internal consistency and compliance. However, it still fails to adequately address the users (programmer, statisticians…) daily operational task – that of how to go about generating a quantitative analysis product. This is principally because the replacement of any direct references to specific software or systems now leaves the user without the step-by-step guidance required to create any specific quantitative analysis deliverable.

Using the same APESS approach discussed earlier (Figure 2), ‘joining’ the procedural infrastructure to the practical tasks is now achieved through user or business manuals. These documents focus on describing the specific ‘how-to’ and will now include direct reference to systems or software. Arriving at the minimum set of business manuals required can be optimally achieved by identifying and analyzing the quantitative analysis use cases. This exercise, when undertaken systematically, highlights both the range of outputs of interest and their associated roles (often recorded as job-titles). This set is then used to identify the primary use cases which become the subject of each user manual. This approach also offers an unambiguous system agnostic entry point to support systems implementation or evaluation (as a primary input to a SDLC methodology)

Figure 4 shows the components of a completed APESS document set, together with examples of how this unified approach offers different but consistent views of the procedural infrastructure to the business functions who are responsible for it and need to work with it.

EXISTING DOCUMENTATION REVIEW AND REVISION

The method can also be applied to the review and revision of existing quantitative analysis documentation. Starting with an existing process document, the approach is as follows:

1. Determine what the document under review is supposed to be – a SOP? a Work Instruction? a User Manual? It is not uncommon, for example, to find one type masquerading as another – for example,
regulatory and compliance focused work instructions often contain detailed execution elements that are more appropriate for user-focused business manuals.

2. Examine each procedural step and categorize as A, P, E, S or S. This exercise highlights the major detailed type of activity the process is describing. For example, heavy attention to acknowledgement steps related to decisions about if and when work is required should be associated with supervisory roles; the absence of any reference to a storage step suggests the work may not be traceable in the future.

3. Examining the resulting detailed step-by-step breakdown is now be used to identify first the core execution steps – those that actually create the required quantitative analysis deliverable – and secondly the required planning steps. This separates the principal reason for the control document from the ‘wrapper’ that oversees the business requirement.

4. With the planning and execution steps now at the core of procedure, the acknowledge, storage and sharing can now be added back to create a draft revised APESS consistent document. This in turn can now be tested to remove redundant steps, or to reorganize into a more consistent and operationally focused document that can go forward for inclusion in the quality management system.

In practice when undertaking this exercise, it is helpful consider the APESS steps as ‘phases’ rather than steps in themselves. Many SOPs and Work Instructions, particularly in the planning and execution phases often require reviews of interim quantitative analysis deliverables or confirmation (acknowledgement) of particular technical planning steps, for example, when new statistical methodology is involved.

DISCUSSION

We have developed and deployed the APESS methodology as part of the generation of a quality and user documentation for a clinical data processing system. This method has reduced the amount of paper work and the difficulty of generating and reading the corresponding documentation through the use of a common methodological framework, which allows the writer or reader to know the high-level activity being described and its purpose. This general approach, which is based upon best-practices from data science practices for transparency and reproducibility, has been found to be comparatively easy to understand and grasp, when compared with similar documentation which is generated without such as unifying structure.

REFERENCES

1 Six-Sigma: https://en.wikipedia.org/wiki/Six_Sigma
2 Lean: https://www.lean.org/WhatsLean/Principles.cfm
3 BPMN: http://www.bpmn.org/

CONTACT INFORMATION

1 Anthony Rossini, Anne Mounier, Novartis Campus, Postfach CH-4002, Basel Switzerland
2 Andy Richardson d-Wise, Suite 7A, Manchester One, 53 Portland Street, Manchester, M1 3LD
Figure 1: Generic APESS Procedural Documentation Flowchart. The 5 APESS steps are shown along the major axis together with the associated principal decision points. This serves as a starting point for developing new procedural documentation and reviewing and revising existing procedures.

Figure 2: Relationship between the regulatory document framework (rows) and the APESS steps (columns). The APESS steps in yellow are expected to be referenced/described in each document type, whilst the grey APESS steps are not expected to appear routinely in that document type. This top row of the grid is expected to describe the process generally but at a high level, and then with increasing level of detail moving down. Using this as a ‘test’ it is expected that a user manual would discuss principally execution steps, a work instruction details of how to plan, execute and store a quantitative analysis deliverable, but without referencing specific software to be used, and the associated SOP to detail the circumstances required to start the process (acknowledge the requirement) and, following creation, with whom it may be shared.
Figure 3: Hierarchical referencing. Using the APESS step categorization in conjunction with a hierarchical step ‘numbering’ convention traceable compliance through all related procedural documentation can be established unambiguously. This improves the ‘auditability’ whilst at the same time reducing the user training burden; in the example following work instruction steps is guaranteed to ensure compliance with the SOP.

Figure 4: APESS Quantitative Analysis Procedural Document Set. The diagram shows the shape of a completed procedural document set. The document cascade from SOP to business manuals (in yellow – describing individual use cases – see text) is internally consistent and can be followed forward or backwards to confirm compliance. The outer lines show the different, but consistent, views on the document set that various parts of the organization require and includes (dotted lines) the system and functional views that link the procedural documents to the IT infrastructure. The red lines show the compliance routes through the documentation set to create two different quantitative analysis deliverables. WI-002 is an example illustrating the creation a study table or listing – UC-004 creates the required datasets, UC-003 the final output. Reporting PK parameters is shown by WI-004 – dataset preparation is the same using UC-003, analysis and reporting is different (UC-005) and a different computing environment used. The different overall APESS requirements for these two outputs is recognized through the different SOPs that control their creation.