PhUSE 2017

Paper DV03

REACTing to data: The Use of Data Visualisation within Early Clinical Statistical Programming at AZ

Shyamprasad Perisetla, AstraZeneca, Cambridge, United Kingdom
Suzanne Tautz, AstraZeneca, Cambridge, United Kingdom

ABSTRACT

REACT is an innovative data visualisation system that has been developed for AstraZeneca clinical trials to provide clinical study teams with real-time access to their study data in an interactive format. The system is used during the lifecycle of a clinical trial until reporting the study data and beyond, and is beneficial for both Early Clinical Development (ECD) and late phase trial functions at AstraZeneca.

In this paper, we describe how the REACT system is utilised by the AstraZeneca ECD statistical programming group in order to enhance the analysis and reporting of the clinical studies whilst evolving the traditional role of the statistical programmer. We will explore how the system improves the quality of SDTM and ADaM, builds value into vendor deliverables, and supports the study team in preparation for key decision-points in the clinical programme either to stop failing treatments early or to accelerate beneficial treatments to market.

INTRODUCTION

Data visualisation is a critical component of analytics in the clinical environment. With the standardisation of clinical study data into SDTM and ADaM, the opportunity for applications, such as visual analytics, is realised. As the saying goes “A picture is worth a thousand words”, especially when we aim to identify patterns in the study data comprising many thousands of records and variables. Reviewing data of such size is more efficient through a digital review solution, eliminating the need for unwieldy and time-consuming static outputs. Data visualisation can allow study teams to get closer to their data, providing the ability to identify safety signals, support data quality checks, and enabling fast drill-down from a programme to a study population to cohorts to individual subjects.

Early phase studies are designed for a different purpose to later phase studies; they are exploratory in nature, being designed for dose finding and investigating safety, tolerability and therapeutic efficacy. Data visualisation has potential to be particularly impactful for early phase studies to access the data regularly, transform it visually and understand it quickly. The ECD organisation at AstraZeneca conduct early phase (I up to IIb) clinical studies. ECD are focussed on understanding emerging study data, to either advance the program and adapt to signals, or make an early decision to stop the study. Intensive exploratory analysis is performed regularly within ECD to support this internal decision making, both in the study itself and within the wider clinical research programme for the disease-treatment. To support these activities, ECD utilise an internally developed suite of web-based data visualisation tools, called REACT (REal time Analytics for Clinical Trials), which is specifically designed for interactive exploration of clinical data. REACT helps the study teams in many ways to investigate a benefit/risk profile of a treatment, respond to and prevent the escalation of data issues, make decisions on dosing, and primarily ensuring the safety of patients in clinical trials.

The role of statistical programmers is to translate the raw study data into meaningful information for use by the project teams. Within ECD this statistical programming role is evolving from standard data manipulation and generation of Tables, Figures and Listings (TFLs), to “data scientists”, who not only help to create, develop and enhance visual analytics applications such as REACT, but also leverage them to support their translational activities. With their knowledge of underlying data structures and transformations, they are able respond to requests from different stakeholders regarding the clinical data, from providing expertise on the available data, to defining and creating ad hoc analysis to further data exploration. ECD statistical programmers are therefore an integral part of a clinical study team; providing essential support to statisticians, physicians, study managers, data managers and vendors.

Utilising REACT brings ECD statistical programmers closer to the study data, enhancing the traditional statistical programming deliverables, and improving their ways of working as well as ECD clinical programmes. The instantly available summaries in REACT enable ECD statistical programmers to respond to ad hoc requests more quickly, review vendor deliverables more thoroughly, and observe emerging data more efficiently. It also allows ECD statistical programmers to focus on more complex summaries and profiles to assist in the decision making alongside

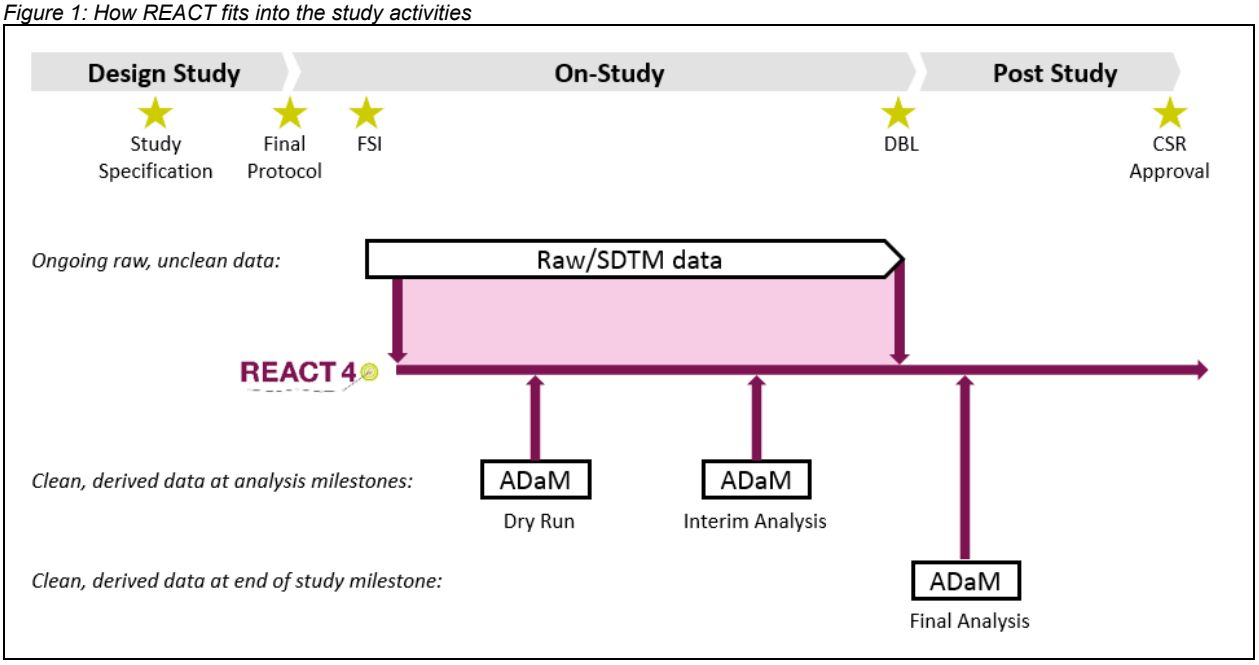
REACT. The purpose of this paper is to describe the benefits realised by ECD statistical programmers through having REACT in their analytical toolkit.

WHAT IS REACT?

REACT is a user friendly data visualisation tool, designed by AstraZeneca for their clinical trials. It is a web application that enables you to view and interact with clinical data as it emerges during a clinical study, or at specified data cut points and end of study reporting data. REACT provides data visualisations and easy-to-interpret graphical charts and tables which are interactive and can be drilled-down to individual patient level data. Whilst REACT produces data displays, it does not perform any derivations of endpoints and summarises the available data as it is in the database.

After first subject in (FSI), data for ongoing studies are read daily from files held in the AstraZeneca database system (see Figure 1). These files are either in raw CDASH-formatted or SDTM-formatted SAS file datasets, and are updated as frequently as agreed for the study. The ongoing data may be uncleaned and unvalidated. Whilst on-study and post database lock (DBL), REACT also supports dry runs, interims, and end of study data review and analysis. For these time points, data additionally come from ADaM-formatted SAS file datasets, making use of derived data types and analysis flags that are contained within those data files and used in corresponding programmed TFLs. These data are read into the REACT database, which in turn populates the graphs and tables within REACT.

Figure 1: How REACT fits into the study activities



All members of the study team can use REACT, being configured in such a way that no prior understanding of underlying data structures are required. The tool can display summaries of many of the collected data domains, including Adverse Events, Concomitant Medications, Labs, Biomarkers, Dosing, Exposure, Efficacy, and Medical History. Patient Profiles and single subject summaries are also available.

AstraZeneca have designed a best practice framework, which is a series of recommendations and guidelines informing study teams and additional stakeholders how the REACT tool should configured, deployed and used. This is integral to the tool to ensure REACT usage has no risk to study integrity and conduct.

REACT is not a GxP validated system, therefore any signals identified in REACT should initiate further review of the source data and any interpretations leading to conclusions and or decision making are always confirmed by a GxP validated method on cleaned data.

BENEFITS TO ECD STATISTICAL PROGRAMMERS

ECD adopts an outsourcing model where the protocol specified statistical reporting activities are performed by Analysis and Reporting (A&R) vendors. ECD statistical programmers are accountable for the planned deliverables

from the A&R vendor at key milestones throughout a study, whilst also supporting the study teams internally with their requests for data access, exploration, and further analyses. ECD statistical programmers appreciate the benefits of using REACT to build quality and efficiency into their planned and unplanned study deliverables, whilst ultimately bringing them increased visibility on study data.

SDTM AND ADAM DATA QUALITY

In order to set up a study in REACT, firstly the source data must be mapped into the system. This mapping process is performed by the statistical programmer, as they understand the study data and underlying structures. Therefore they can ensure that the correct information is read into the REACT database and is displayed in the right way. Since ECD outsource the protocol specified analyses of studies to CROs, the process of mapping the SDTM and ADaM data into REACT provides an opportunity for ECD statistical programmers to get closer to the A&R vendor derived SDTM and ADaM datasets, in particular for those domains included in REACT, as per the table below.

REACT View	SDTM dataset	ADaM dataset
Population Summary	DM, DS, SC	ADSL
Dosing & Exposure	DS, EX	ADEX
Adverse events	AE	ADAE
Concomitant medications	CM	ADCM
Labs	LB	ADLB
Vital Signs	VS	ADVS
Cardiac function	CV, EG	ADEG
Liver function	LB	ADLB
Renal function	LB	ADLB
Tumour Response	RS, TU, TR	ADFAON, ADRESP, ADEFF, ADTR
Respiratory	CE, FA	ADCE, ADEXAC, ADEFFRE, ADRE, ADQS
Timeline	AE, CM, DM, DS, EG, EX, LB, SV, VS	ADAE, ADCM, ADEG, ADEX, ADLB, ADSL, ADVS
Single Subject	<all included domains>	<all included domains>

Once SDTM and ADaM data are mapped to REACT, the data quality of the source data and the A&R vendor deliverables can be reviewed visually. Outliers or unexpected results may highlight errors in the derivation rules, which can be investigated further by the ECD statistical programmer and fed back to the A&R vendor as necessary.

There is a challenge with the consistency of domains across a programme of studies, either in design or in the way data are captured, often from deviations to the standards used due to differing interpretations by A&R vendors. Reviewing SDTM and ADaM data during this mapping process early on in the study allows the deviations to be rectified and clarity to be added to the AstraZeneca standards. This ensures quality and ultimately saves time for the statistical programmer during the study process and beyond.

Including SDTM and ADaM data in REACT is therefore advantageous to ECD statistical programmers. This approach improves the quality, regularity and speed of the SDTM and ADaM deliverables, and it provides an opportunity for the ECD statistical programmer to get closer to the study data. This makes the ECD statistical programmer more efficient both with reviewing these data and when analysing them later for ad hoc work.

SUPPORTING ON-STUDY ACTIVITIES

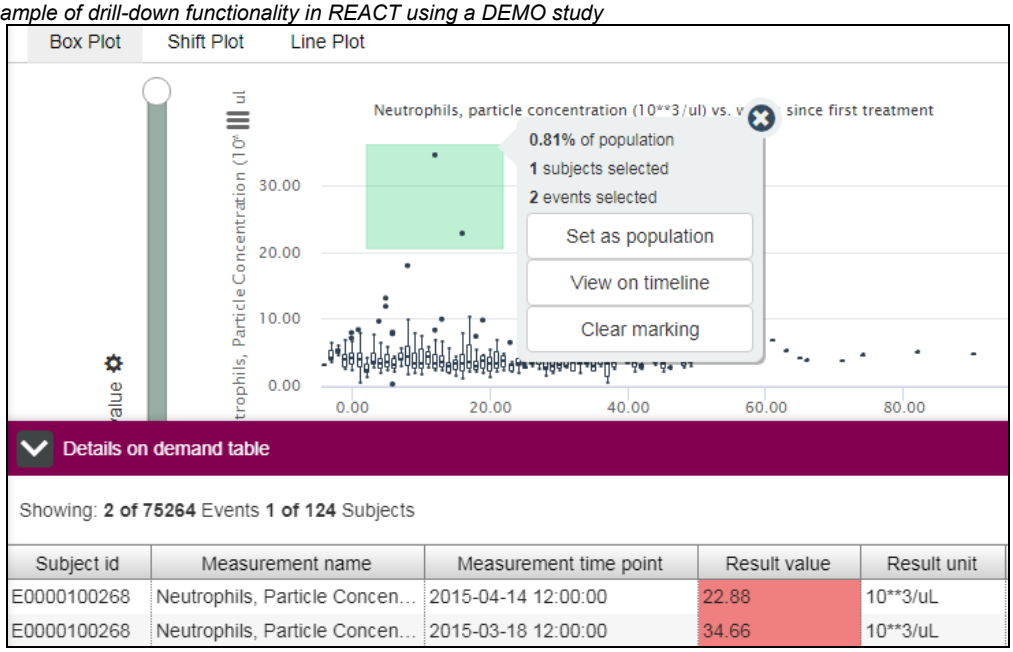
On-study analysis may take many forms, from data and/or analysis queries to delivering ad hoc TFLs. These are usually considered unplanned requests and often require production at short notice, they also require the ECD statistical programmer to be agile as the request may evolve during study team discussions. For these reasons, they are often difficult to outsource to A&R vendors, and therefore are supported by ECD statistical programmers throughout the study. REACT forms part of the support of these on-study analysis activities, benefitting ECD statistical programmers in a number of ways.

Study teams are increasingly engaged in emerging data within ECD, and ECD statistical programmers educate study teams about the possibilities in REACT. Some available REACT views have and will continue to reduce the need for ECD statistical programmers to create static outputs or run them regularly at pre-specified time points for study team review. Review is enhanced through REACT, as study teams can do further investigations with instant drill-down functionality. This enables more issues to be identified, enhancing overall data quality. Since study teams are more engaged in the data, they also capitalise on the potential of data, by identifying issues or realising trends earlier. This engagement drives further analytics and helps ECD studies evolve.

Many standard ad hoc requests from study teams are readily available as visualisations in REACT, or they can be created through filtering and modifications within REACT. The ECD statistical programmer therefore does not need to spend time developing similar outputs in a static manner. Instead, the ECD statistical programmer can delve deeper into any requests, validating the REACT results or providing supporting analyses through data and derivations not available in REACT.

When the study team has an ad hoc request that cannot be obtained through REACT, the ECD statistical programmer may first use REACT to visualise and verify the data in question, before producing the ad hoc output. This can reduce the time taken to deliver ad hoc requests, since it can be quicker to explore data first through the readily available visuals in REACT than writing new programs to do so.

Figure 2: Example of drill-down functionality in REACT using a DEMO study



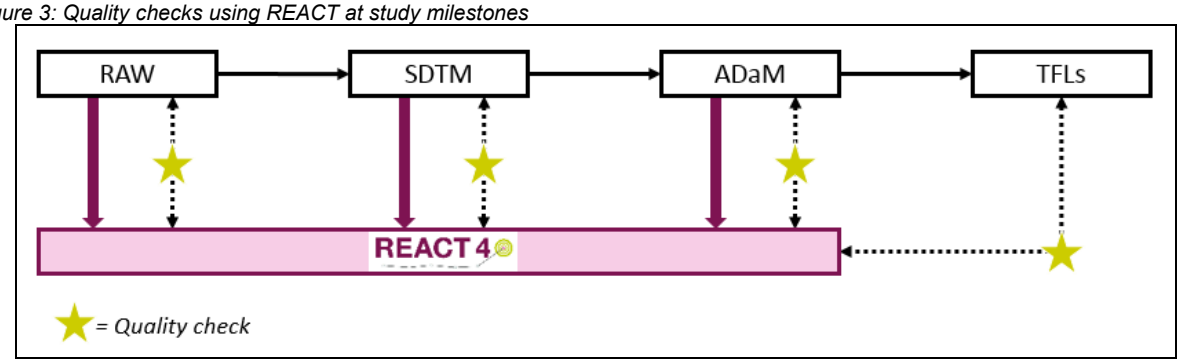
ENHANCING MILESTONE DELIVERABLES

The milestone deliverables at interim and/or end of study analyses, are produced by the A&R vendor, including any dry runs for the analyses. It is the responsibility of the ECD statistical programmer to review such deliveries in partnership with the A&R vendor. REACT assists with this review, as it can be used to cross check the A&R vendor outputs. Raw data from the specified data cut, or SDTM/ADaM data delivered from the A&R vendor at such milestones can be loaded into REACT separately to the daily data views. In this way, the data available in REACT can match exactly that used by the A&R vendor outputs.

Dry runs are often the first opportunity to receive and review ADaM datasets from the A&R vendors, and therefore also the first opportunity to map the derived endpoints into REACT. Setting up ADaMs in REACT allows for a closer review of the derivations, during the mapping process into REACT. Reviewing ADaMs in REACT at the dry run can also highlight where additional programming rules may need to be added to the ADaM derivations, whilst also enabling a cross-check of the end-to-end transformations through the REACT views. Traceability checks between raw data to SDTM to ADaM can be performed by comparing the key visualisations in REACT. Once set up at the dry run, ADaMs are ready to launch in REACT at end of study, shortly after receiving the final data.

The review of the A&R vendor deliverables by the ECD statistical programmer comprises various checks, including independent review of the key primary, secondary and safety endpoints. This independent review is in addition to the validation checks performed by the A&R vendor. Since many of the REACT visualisations display data in a similar manner to the planned outputs from the A&R vendor, they can be used to cross-check the results. If the numbers from REACT outputs do not match those numbers in the A&R vendor TFLs, this will be investigated to find the source of the discrepancy. In this way, the need to develop programs from scratch to assist with the review is eliminated. This saves the ECD statistical programmer time and enables them to do a more detailed assessment of the outputs which cannot also be visualised through REACT, thus improving overall quality of the deliverable.

Figure 3: Quality checks using REACT at study milestones



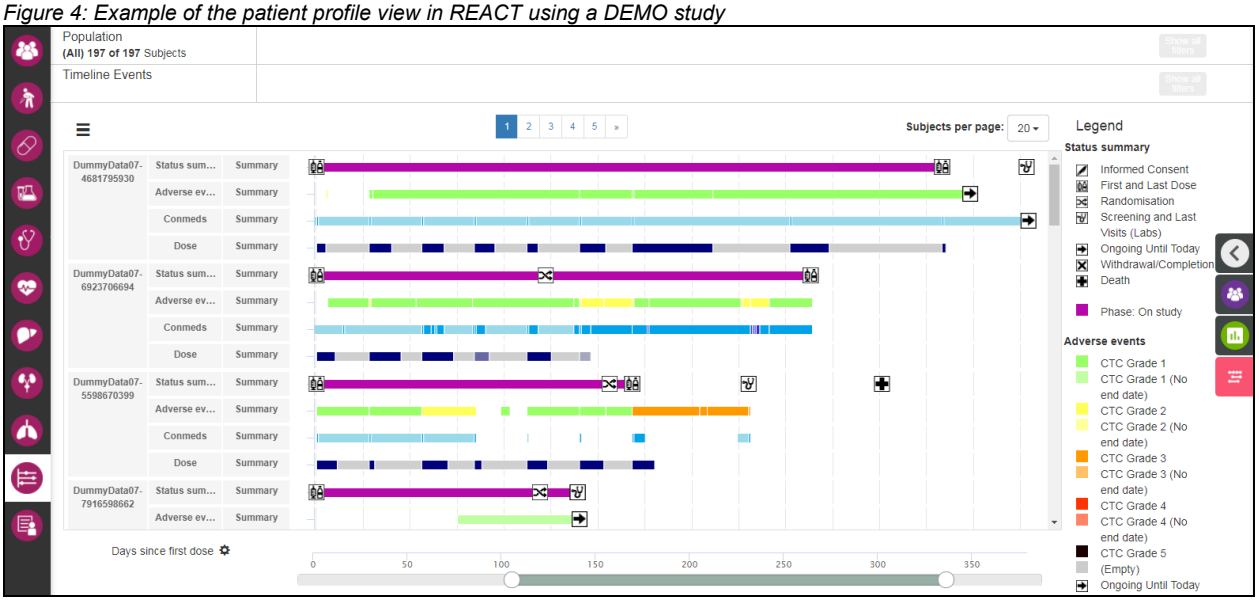
Using REACT in this way throughout the study and at the key milestones, means that it is less likely to have surprises in the data requiring lengthy analysis at the end of the study, and the validated outputs programmed by the ECD statistical programmer required to support decisions, submissions or presentations can be targeted.

DELIVERING POST STUDY ANALYSIS

Enabling the ADaM data to be available in REACT as soon as possible at the end of the study, is vital to help the study teams get up to speed with the results quickly, through viewing the REACT visualisations using the same data source as the TFLs. Reviewing data from this common reference point, and having the analysis data available to interrogate interactively, can reduce the post-hoc burden and make the best use of ECD statistical programming time during this stage.

REACT can also be used to support exploratory investigation of regulatory questions. Since data can be interrogated quickly in REACT, it allows ECD statistical programmers to focus on what is relevant and identify if a new specific TFL output is required to be programmed in order to address the question from the regulator.

Figure 4: Example of the patient profile view in REACT using a DEMO study



BENEFITS TO A&R VENDORS

ECD A&R vendors are part of the clinical team and therefore able to access REACT and enjoy the same benefits the data visualisation system brings to the internal study team. The A&R vendors are able to identify data and quality issues through REACT, whilst also allowing for transparency of ECD and their views of SDTM and ADaM data. In this way internal and external study team members are kept aligned on understanding the study data and any issues identified.

FUTURE

The REACT functionality is continually evolving, as updates and new visualisations are requested from REACT users. ECD statistical programmers submit requests to the REACT team, since they are very familiar with the differing types of ad hoc requests from their study teams which could benefit all users if they existed as views in REACT. ECD statistical programmers track all ad hoc requests so they can gain metrics on the requested ad hoc output, and the most requested ad hoc outputs are submitted to the REACT team to become future REACT visualisations.

Much of the REACT visualisations exist for safety data however, more detailed efficacy visualisations are being derived. This often requires support from ECD statistical programmers who may contribute code to the REACT developers for more complex derivations, or be involved in the validation of such activities in REACT.

AstraZeneca are developing a Safety Review tool which will interact directly with REACT such that it will be possible to filter on data which have been medically reviewed. This will extend the capabilities of REACT further.

Beyond these activities within AstraZeneca, we feel the type of outputs delivered in regulatory submissions will change in the future to embrace interactivity. We feel it would be beneficial to a regulatory reviewer to interact with and cross-examine the results they are presented with. Static reporting will always have a place in science. But newer technologies using interactive visualisation of data have the potential to transform our ability to report on very complex data. To fully understand the relationship between all of the subject-state, time-dependent, and covariate relationships in data that we gather, we need a new way of reporting. Visual analytics tools like REACT provide end-users of the study data with the technology to reflect these relationships. Just as advances in genetics, chemistry, and even engineering have improved the development of new compounds and therapies, advances in the technology we use to analyse clinical study data provide the opportunity to improve our ability to learn from the data.

CONCLUSION

REACT enhances access to data for all, aiding communication and collaboration and empowering ECD study teams to review and improve their understanding of the study data. ECD statistical programmers are embracing such new technologies and using them to evolve as data scientists, to enhance their standard deliverables as well as the support they provide to the study team. In this way ECD clinical trials are better equipped to respond to the emerging data, be more informed at key decision-points, and are therefore able to run more productively, to ultimately benefit the patients.

ACKNOWLEDGMENTS

We would like to thank the REACT team at AstraZeneca.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Shyamprasad Perisetla
AstraZeneca
Melbourn Science Park
Royston
SG8 6HB
United Kingdom
Work Phone: +441763 269134
Email: shyamprasad.perisetla@astrazeneca.com

Suzanne Tautz
AstraZeneca
Melbourn Science Park
Royston
SG8 6HB
United Kingdom
Work Phone: +441763 269133
Email: suzanne.tautz@astrazeneca.com