



AccelerateR

Utilizing Supporting and Mentoring for Delivering a Statistical Package in R



Table of Contents

1. Introduction
2. Learning from the past
3. Embracing the Future
4. Making the Transition
5. Lessons Learned
6. Where we want to go

Introduction

Who we are and why we know what we are talking about



Ben Arancibia is a Director of Data Science with 12+ years of technology driven solutions experience in the public, private, and social sector. He has a track record of producing thoughtful, high-quality advanced analytical work on highly visible engagements that are typically subjected to scrutiny by C-suite executives, board members and major stakeholders. Considerable experience in implementing data science solutions, either in the form of transformational efforts in data science strategy, new market entry or new business model development; to more tactical changes, such as machine learning optimization, data visualization improvements, technology platform development, system integration, and operational risk analysis.



Alice Ehmann is an Associate Director of Clinical Programming with 20+ years of experience in statistical programming as well as data management programming. At GSK, in addition to leading study programming, she is involved with activities to support the integration of R within clinical programming including R pilot projects and package development. Prior to GSK, Alice worked in the CRO realm in various capacities. This has given her the opportunity to experience many system platforms and processes. Alice looks forward to continuing the R and open source code journey and integration in the pharmaceutical programming world.

GSK Biostatistics

Analysis of clinical data for our Assets

Primary Goal: Reporting Clinical Trial Data

Summarizing study population, safety, and efficacy data
(SDTM -> ADaM -> TLFs -> CRT package)

Provide accurate summaries of study data

Respond, communicate, and collaborate with regulatory agencies and internal stakeholders

GSK Biostatistics Organization

Talk about three groups (Stats, CP, and SDSIH) and how they operate together

GSK Biostatistics

Statistics

Clinical
Programming

Statistical Data Sciences
Innovation Hub (SDSIH)

Statistics and Clinical Programming

- Work on pipeline delivery of assets
- Traditionally SAS oriented

SDSIH

- Center of expertise for using R to implement in both Statistics and Clinical Programming Functions.
- Enable business verticals with training, support, mentoring, and business specific R technical toolkits.

GSK Biostats Makes a Strategic Decision

2020 - A Pivotal Year



Adopt Multilingual
Approach for Clinical
Delivery



Begin to move
codebase to R and
piloting QC activities
using R



Start to use R as a
core tool for delivery



Begin to train and
embed R as a core
data science tool

Learning From the Past

How Old Methods of Training Failed to Scale



Lessons Learned

Timing of Adoption is Crucial

Problem

Not all study teams adopt and transition to R at the same time. Requires retraining of the use R.

- Can't actively train users all the time
- Small agile team cannot scale to meet intensive in person training
- Study teams have different needs according to their timelines
- R experts sit outside business

Solution

Adopt a new training strategy focused on just in time training AND support and mentoring.

- Focus on creating training material that is relevant to study teams
- Make training available all the time in either recorded videos or documentation
- Create agile team of data scientists and statistical programmers to rotate through study teams and provide direct mentorship and support

Embracing the Future

Learning How to Adopt a New Training Strategy

Consolidate Training Resources

Internal R
Resources

External R
Resources

Identify and
create training
resources to
make R and
GitHub work
for clinical
reporting

AccelerateR Team Formation

Core Team
with Clinical
Programmers
to speak the
same
language

Focus on
mentoring and
support
through
training, code
reviews, and
drop in
sessions

Build
frameworks
and take
lessons
learned and
turn in
consumable
resources

AccelerateR Team

Principles and How the Team Works

Principles

Business Problem Focused: Team focused on helping to solve problems that impact the business

Radical Transparency: Team agrees to be transparent about all work and challenges performing the work

Agile: Prioritizing work and being able to adapt quickly based on changes or new needs from business

Knowledge: Create knowledge and document in consumable resources for broader adoption

How it works

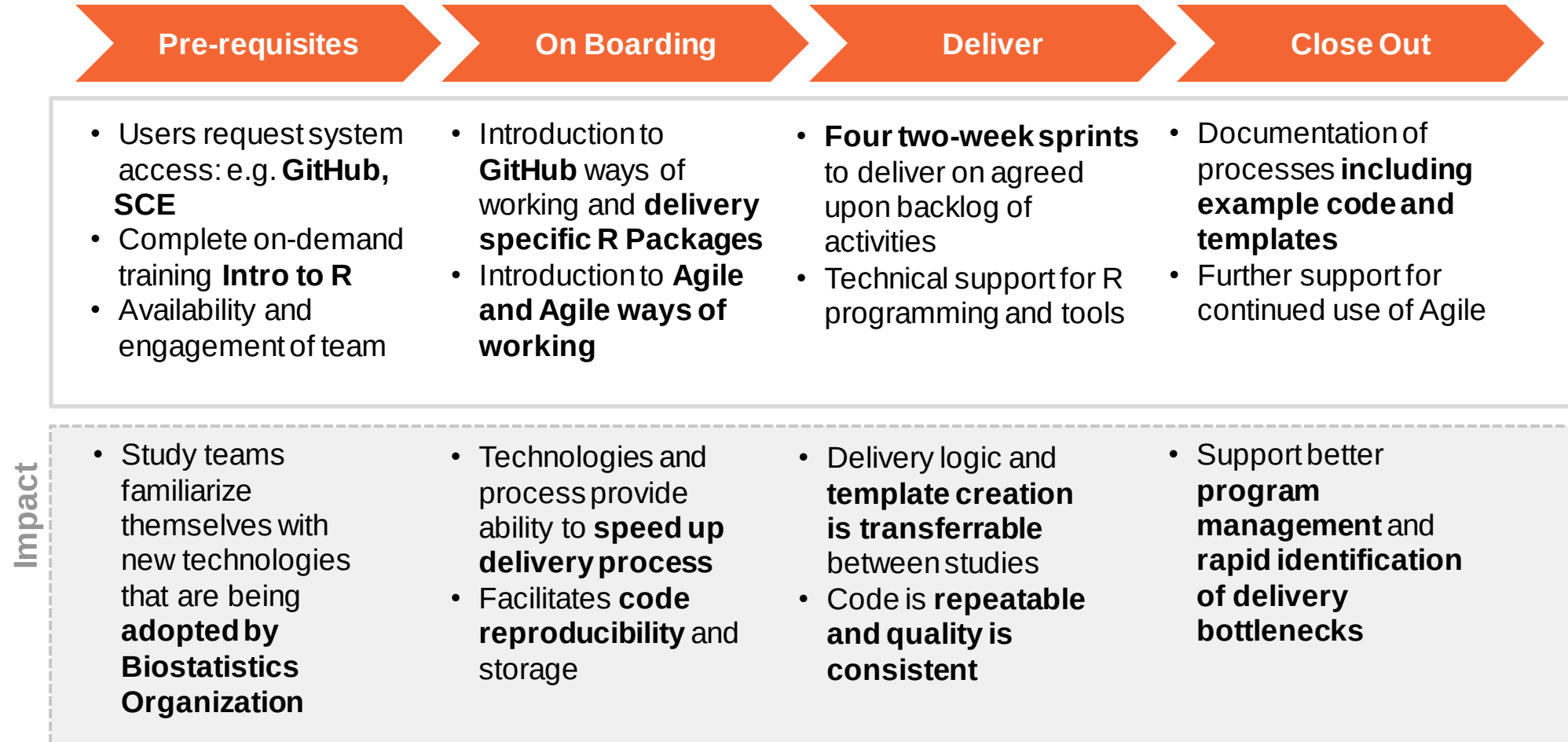
SMEs supporting and mentoring specific project teams and individuals working as an agile pod to rotating through study teams to put learnings into practice

AccelerateR rotates through teams in 9 week intervals.

- Initial system access help
- First week is focused on team specific training needs
- Four two week sprints to deliver on backlog items
- Close out – setting the team up for continued success

Typical AccelerateR Engagement

How AccelerateR Utilises Tools, Technologies, and Processes



28 March 2023

Making the Transition

Implementing the Strategy

What we stopped

Stopped large in person training courses – not worth the time and most attendees will need to be retrained

What we started

Spending more time in the business groups
Building more specific code templates and example use cases for clinical reporting

What we still need to do

Build consumable resources so that study teams do not have to consistently start from scratch when start delivery in R

Lessons Learned through AccelerateR

R, git/GitHub, Agile - Training

- Up front training and regular drop –in sessions on GitHub and R were well received
- Agile Training on top of the technology training is a lot, goal should be an introduction to agile not full embedding
- Having study level functions and providing template code eased onboarding

Timing of Support

Timing support with AccelerateR is crucial for success. Study can't be too far in advance because resources won't be available for training, but also study can't be too soon because of complexity of teaching new technology and ways of working for delivery

R Environment

There is tradeoff that study teams need to navigate in terms of using the frozen R4GxP environment and new packages that would be beneficial not available in the frozen R4GxP environment

Ongoing Problems AccelerateR looking to solve

Git and GitHub

- Steep learning curve (commits, branching, pull requests)
- Version control traditionally treated as a checkbox requirement or quick backup rather than a tool to aid in code development and management
- Limited support and historical implementation in clinical programming

Programming Team Management

- Continual upskilling of programmers given teams change over time (planning and assessing team makeup is critical)
- Encourage acceptance of existing packages and functions
- Accept what we did today may not be what we do tomorrow as open-source technologies evolve quickly

Code base (R vs. SAS)

- Packages vs. Macros – it does not need to be a 1:1 relationship
- Library of readily available template code is small but continues to grow
- Template vs. package – which is better?
- Educating programmers on how to utilize package documentation as well as package GitHub repositories

Code Review

- Historically a resistance to code review in clinical programming
- Needs to be implemented in a constructive, non-judgmental, and collaborative manner
- Timing is everything (early in the process)
- Time overhead associated with code review and needs to be planned ahead

Achievements

Partnership with Business

- Business is looking for support in adopting R, GitHub, and agile ways of working – many study teams don't know where to start or best practices
- Providing template code and having individuals to answer questions has helped technical team understand business needs and craft solutions for them

Technology Solutions

- New technology solutions to enable clinical programmers deliver while using packages that are not “qualified”
- Internal R wrappers and other code to create GSK specific outputs

GitHub Code Repo

- Identifying common needs that are repeatable across studies
- Turning code solutions of standard practices into reusable functions and storing code in easily access GitHub Repo – potential to turn this into start up package

The Future of AccelerateR

**Scale AccelerateR
Team to more than 1
agile pod**

**Engage with
statistics function**

**More robust code
repository**

**Turn suggestions in
backlog of items to
develop**

**Support adoption of
R by continuing to
expand capabilities**

**Guide users to open
source solutions and
projects that they
might want to
contribute to**

GSK