

# Why Generic **R** Training Is Not Good Enough for Clinical Programmers

## Overview

There is a significant buzz in the Life Science industry around using R in the clinical space. Many companies are actively transitioning some of their clinical tasks to R. Some companies are only using R for QC. In contrast, others are starting production study programming in R, building shiny apps, or creating CDISC datasets.

All of these companies need a workforce capable of performing clinical tasks in the R programming language. Where can a life science organization find people who understand both R and the clinical space?

The answer for most companies is to retain and train their current workforce. The current programmers already know the clinical space, they already know their data, they already know how the company operates. So, just let them learn R—right?

Yet, for a clinical programmer working in SAS® for the last 20 years, learning R is a tall order. They have very different languages, capabilities, and mental models for how you get things done.

Further, the widely available R training is typically generic. It is not tailored to clinical tasks at all. There are several reasons why generic R training, such as the kind you can find easily (and cheaply) online, is inadequate. Let us explore some of them.

1

## Generic R Training Doesn't Explain How R Works

R is a vector-based language. Many functions take vectors as input and return vectors as output. Even the R data frame is constructed of vectors. To understand how the language works, an R learner needs to understand these basic facts; otherwise, they will never solve problems when things go wrong.

Yet, these basic facts are typically ignored by generic R training. Generic R training usually ignores Base R almost entirely. It skips right over Base R and starts the user out with Tidyverse functions, then moves on to ggplot2.

With this type of curriculum, the learner will not understand the foundational data structures that give R its power and flexibility. Instead of coding from a position of understanding, they would be coding from a place of ignorance. Instead of thinking their way through a problem, they would be resigned to Googling everything. Instead of writing the most efficient code for their use case, they would blindly copy code from the internet into their programs without understanding what that code is doing.

Imagine your workforce training solution is putting your people in that position. That is what happens when you use generic R training as your workforce enablement solution.

2

## Generic R Training Provides No Familiarity to the Clinical Programmer

Learning a new programming language is hard. But what makes it genuinely onerous is when the examples and exercises have no relevance to anything you are doing. Modeling the distribution of quail nests on the Alaskan tundra may be interesting to a wildlife biologist. However, it would leave clinical programmers scratching their heads and mystified trying to relate these examples to the routine tasks they perform every day.

The best type of training starts with something the learner already knows and understands and then leverages that to explain something new. Step by step, the learner moves from familiar concepts to unfamiliar ones.

By definition, however, it is impossible for generic R training to accomplish this progression. Generic R training cannot presume anything about what the learner already knows. In trying to create a “one-size-fits-all” type program, what they end up with is a program that doesn't fit anyone.

That is why people who take generic R training end up forgetting much of the training after a few months—because the learning was not anchored to something they already knew.



# Why Generic R Training Is Not Good Enough for Clinical Programmers

3

## Generic R Training Focuses on the Wrong Things

Many R packages are helpful to the clinical programmer. For instance, the janitor package has very nice functions for creating summary statistics and comparing data frames. The logr package allows users to create a traceable program log easily. And the haven package will enable users to get data back and forth between SAS® and R.

Unfortunately, generic R training does not teach any of these packages. Generic R training typically only focuses on the most popular R packages: dplyr, tidyr, and ggplot2. While these packages are indeed helpful for clinical programmers, there are many packages essential to clinical programming that generic R training does not address at all. Overall, the scope of generic R training is too narrow and not aligned to the types of activities that clinical programmers need to perform.

The best type of R training for the clinical programmer is one that has gone through the universe of R packages and pulled out those packages that a clinical programmer needs.

4

## Generic R Training Doesn't Teach Clinical Tasks

If the point of your R enablement program is to teach clinical programmers how to work in R, performing clinical tasks in R should be a part of the program. Yet, with generic R training, they are not. Generic R training teaches you how to perform functions in R. It is then up to the learner to figure out how to apply that knowledge to their clinical work.

For example, a generic R training program may teach programmers how to select variables from a data frame and filter the observations. But a generic R training program will not teach you precisely how to create SDTM datasets in R. It will be up to the user to cross that last mile and apply the generic learning to their particular use case. It is a question of direct learning versus indirect learning.

The best learning programs will teach you what you need to know directly. If your task is to create CDISC datasets with R, then the best learning program will teach you exactly how to accomplish this. If your task is to create TLFs in R, it will show you step by step how to achieve that task.

5

## Generic R Training Provides No Support

The final problem with generic R training is that it provides little or no support for the learner. If the learner is struggling, no one notices. If the learner needs help, there is no one to support them. If the learner misunderstands something, there is no one to correct them and get them on the right path.

A learning program would monitor a learner's progress and reach out to them when it sees them falling behind. Such a program would have experts available any time the learner needs help. More generally, it would shepherd both the learners and their managers along the R learning journey and make sure everyone gets to the desired destination.

## Conclusion

Generic R training makes some attractive promises. It is easy to sign up, flexible and has an affordable cost.

Unfortunately, these promises fail to achieve the primary goal of most organizations: get your team confident, competent, and comfortable in R. What usually happens is that everyone signs up, takes a few lessons, gets busy with other things, and then drops the program. After two years, no one has learned how to do clinical programming in R. Experis has seen this happen in multiple organizations. We have seen it happen in our organization. That is why we created the Accel2R Guided Learning Program.

If you want to achieve R competency in your Life Science organization, Accel2R is the best way to do it. It is tailored specifically for the clinical programmer and will give them the foundation they need for mastery of the language. It progresses them from what they already know and focuses directly on what they need. It will also provide the guidance and support your clinical programmers need to be happy and productive in R.



**Experis**  
ManpowerGroup

### About Accel2R by Experis® Solutions

Experis Solutions, part of ManpowerGroup, has been providing timely, high-quality, and cost-effective services to clinical organizations since 1982. Accel2R is an accelerated learning program developed by Experis Solutions that's designed for clinical organizations (i.e. pharma, medical devices companies, and customer research organizations) to upskill their teams to R. Accel2R uses advanced learning techniques, a clinical specific curriculum, learner AI, a 24/7 learning lab, and a unique SAS® to R learning approach to transition from SAS® to R in 4 to 6 months.

For more information about the Accel2R guided learning program, [please contact us](#).