4. Take-home messages

- 4.1 Final cautionary notes
- 4.2 Should I use PRIMER or R? (in short)

4.1 Final cautionary notes

The purpose of this exposé has been to highlight some important pros and cons associated with using PRIMER and R in routine analytical work. It is clear that both R and PRIMER have great capabilities and using them both should be encouraged.

A genuine question about 'which one to use' really only arises when it is perceived that both PRIMER and R each have a specific routine that will (purportedly) do the same thing. For example, both adonis2 (in the vegan package) in R and PERMANOVA for PRIMER assert the ability to implement a dissimilarity-based permutational multivariate analysis of variance. At the current time, PERMANOVA in PRIMER has a far greater scope and capacity than adonis2 to achieve this, and (unlike adonis2) its results are correct and reliable for *any* design.

In Chapter 3 above, we compared the results of a PERMANOVA obtained using PRIMER *vs* R for a specific dataset. We showed that using an R routine outside its limits is a dangerous and flawed enterprise. It turns out there are a lot of other routines in R like the adonis2 function in this respect: they allegedly perform a certain analysis, but may in fact have an inherent weakness in their design, or limitations that are not obvious from a casual (or even a detailed) glance at the available documentation. It becomes clear upon inspection that a broad range of specialised methods available in PRIMER (such as PERMANOVA, PERMDISP, CAP, multi-way ANOSIM, BEST, etc.) are not able to be replicated using any available R packages at the present time.

A lot of R packages (or freely available R code) may look, on the face of it, to be able to do an analysis you want to do. Please bear in mind that there may be:

- problematic assumptions that can lead (unexpectedly) to incorrect results; or
- important limitations (not necessarily obvious) on their correct use.

When using a given R routine, here are some questions you should probably ask yourself:

- How will you know if the results are reliable?
- Can you check it by programming it independently yourself?
- What will it do for your particular use-case at this particular time?
- Are you prepared to re-check a given R routine again when you use it at a different time and for every new use-case you wish to throw at it?

4.2 Should I use PRIMER or R? (in short)

Use Both!

The take-home message here is: **use both**! Neither replaces the other. They are good at different things.

- R is a wonderful programming language and is very flexible and general. You can use it to do heaps of stuff, but it is difficult to use well. Depending on the package, you cannot always trust it without doing a lot of additional checks. It is easy to make mistakes - you need to debug your code and make sure you use existing packages within their scope. Correct functioning may depend on context and can also vary over time.
- PRIMER is an excellent software package with a narrower focus. What it does, it does exceptionally well. It specialises in performing a suite of robust multivariate methods, many of which are simply not available in R (or any other software). It is very easy to use and you can trust the results.

When do I use what?

In my own work, I use PRIMER first and foremost for all the stuff that it can do and is really good at, not just because it is easier (which is reason enough), but also because I know I can trust the results. I use R for most other things, and with few exceptions I program and de-bug my own R code. Breaking this down into some concrete recommendations:

Use PRIMER (with PERMANOVA+) for:

- Robust dissimilarity-based multivariate analysis in general
- Non-parametric methods
  - ANOSIM, BEST, RELATE, MDS, CLUSTER, ... etc.
- Semi-parametric methods
  - PERMANOVA, PERMDISP, DISTLM, dbRDA, CAP, ... etc.
- other methods unique to PRIMER (or better-implemented or easier to run in PRIMER)
  - Second-stage MDS, threshold-metric MDS, BVSTEP, LINKTREE, SIMPROF, custom-ordered shade plots, segmented bubble plots, biodiversity metrics, taxonomic distinctness, functional resemblance, ... etc.

Use R for:

- Running routine classical stats
- Programming bespoke data wrangling
- Tailoring graphics

- Creating and testing new statistical ideas or methods
- Running a specific analysis/method using a package that you know, trust and have checked thoroughly for your particular application or context.
- Running simulations, etc.

I sincerely hope that this contribution will help researchers get the most from their software tools for data analysis. The focus of this exposé has been exclusively on PRIMER and R, as the specific question 'Should I use PRIMER or R?' seems to keep bubbling up. There are clearly a large number of other statistical software options out there (with their own pros and cons) and I would encourage researchers to explore them as well, with an open mind.