

Building Data Science Capacity to Make Rehabilitation Research More Rigorous, Reproducible, and Replicable

As rehabilitation research advances, researchers increasingly recognize the importance and the value of digital databases, standardized data processing, and reproducible, well-documented analyses. This often translates to a need to learn computer programming skills. However, knowing that programming could improve your research and knowing how to get started are two very different things.



This gap helped inspire Reproducible Rehabilitation ([ReproRehab](#)), a research education program designed to help rehabilitation researchers develop practical data science and programming skills while building a community around open, transparent, and reproducible science. And we're proud to share that a session at one of our Annual Meetings sparked conversations that helped drive the creation of ReproRehab.

In 2019, ASNR Members Dr. Sook-Lei Liew, Dr. Keith Lohse, and Dr. James Finley organized an ASNR symposium focused on reproducibility and reliability in neurorehabilitation research. They demonstrated how tools such as Python, MATLAB, and R could make research workflows more reproducible by moving analyses into code that can be rerun, shared, and applied consistently.

The response from attendees revealed an important challenge. As Dr. Liew recalled, people were enthusiastic about the possibilities, but some explained that they didn't really know where to start. This candid feedback got Drs. Liew, Lohse, and Finley thinking about what they could do to give others in the field the skills and the confidence to change how they do research.

Filling a Training Gap in Rehabilitation Research

Researchers in rehabilitation come from diverse backgrounds, including physical therapy, occupational therapy, speech therapy, kinesiology, neuroscience, neurology, physical medicine and rehabilitation, and other disciplines. Many receive extensive training in their areas of clinical or scientific expertise without necessarily receiving formal training in computer programming.

There was an opportunity, Dr. Liew and her colleagues realized, to provide training that was more hands-on than the usual online tutorials, books, or articles. Dr. Liew also drew inspiration from [ReproNim](#), an initiative focused on training people in reproducible principles for neuroimaging, led by ReproRehab team members Drs. David Kennedy and JB Poline. Her experience as a fellow in that program helped provide a model for a rehabilitation-focused program that could introduce similar principles to learners with different levels of programming experience.

ReproRehab launched in 2022 as an NIH-funded R25 Research Education Program (R25HD105583), led by Drs. Liew, Finley, Lohse, Kennedy, and Poline. Its goal is to improve reproducibility and replicability in neurorehabilitation research by building data science capacity. The program uses three complementary strategies: hands-on training for rehabilitation researchers, a train-the-trainer approach that develops researchers who can teach these skills to others, and broad dissemination of resources through a publicly accessible online platform.



Left to right: Dr. Sook-Lei Liew, Dr. David Kennedy, Dr. James Finley, Dr. JB Poline, and Dr. Keith Lohse.

“The big picture purpose of ReproRehab is to train people in data science and programming so they can apply these tools to their own research,” Dr. Liew said, giving researchers foundational skills they can apply to new computational tools and approaches as they emerge.

Meeting Researchers Where They Are

Although ReproRehab was initially envisioned to be largely for people who had never written any code before, the team quickly discovered that demand extended well beyond beginners. Learners span a wide range of experience levels. Some are programming for the first time, while others have used R, Python, or MATLAB before but want to expand their skills and confidence. Another subset of learners includes experienced programmers who are attracted to

ReproRehab because they want to connect with researchers who share an interest in open science, data sharing, and reproducibility.

That diversity is reflected in ReproRehab's pod-based training approach. Rather than giving every participant exactly the same programming curriculum, learners are grouped according to their goals and experience and curricula are tailored to learners' needs. Current course materials include introductory, intermediate, and advanced content in R, Python, and MATLAB, along with topics such as reproducible science, open data, data sharing, GitHub, data management, and data archiving.

Participants also come from all career stages. Dr. Liew noted that learners have included students, full professors, program chairs, and every stage in between.

The current program begins with an in-person summit, followed by 6-8 weeks of online learning. Learners work in small pods with peers who have similar goals and receive hands-on training from teaching assistants. They then move into four months of self-guided work, applying what they learned directly to their own research projects. The project-based component is important. Dr. Liew encourages applicants to arrive with a specific research project or dataset in mind so the skills they learn can immediately be put into practice.

More Than Just Learning to Code

Over time, ReproRehab's leaders have discovered that teaching people how to program is only part of what participants value. After the program's first fully virtual year, participants asked for more opportunities to meet and build relationships beyond their individual pods. ReproRehab responded by adding an in-person summit, which has expanded as the program has grown.

For some participants, the sense of community and confidence instilled have become just as important as the technical training. Dr. Liew said that participants have described ReproRehab as a place where they finally "found their people." The ability to build a community of researchers who share the same enthusiasm for data sharing, open-source tools, and reproducible research practices has been really meaningful.

The program has created lasting relationships and engagement that extend beyond the time spent in a one cohort. ReproRehab teaching assistants are rehabilitation researchers with data science experience who mentor learner pods while developing their own leadership and teaching skills. Many of the teaching assistants are former learners who return to help train the next group, and about half of the teaching assistants each year return to mentor future cohorts. Other teaching assistants join from outside the program but share its commitment to reproducible rehabilitation research.

Seeing learners bring what they learn back to their own laboratories and institutions has been rewarding for Dr. Liew. ReproRehab has supported the development of more than 150

research-related products from program participants, including papers, abstracts, grants, and dissertation work. Perhaps just as meaningful is the way people hear about the program. Dr. Liew said many new participants now apply because a colleague or lab member who completed ReproRehab recommended it to them.

Reproducibility in the Age of AI

The technological landscape has changed considerably since the ideas behind ReproRehab first took shape. Dr. Liew remembers a time when computer skills might have been viewed as helpful additions to a researcher's toolkit. Today, she sees the ability to work responsibly with digital, machine-readable data as increasingly central to research.

Artificial intelligence (AI) introduces another layer to that evolution. The landscape is rapidly changing, but for now, Dr. Liew and the ReproRehab team still see value in learning the fundamentals. Researchers need enough understanding to evaluate whether AI-generated code is correct, troubleshoot problems, validate their analyses, and manage and share their data responsibly.

Each year, ReproRehab continues to adapt. The 2026 Reproducible Rehabilitation Research Summit, for example, includes programming on how researchers can use AI to become better programmers. This will be taught alongside the usual sessions on reproducible science, data standardization, GitHub, and data analysis.

"We hopefully will just continue to shift and adapt," Dr. Liew said, "to train people in whatever skills they need to be able to use computers with their data responsibly."

Getting Involved

ReproRehab welcomes applicants across all career stages and levels of programming experience. There are separate applications for teaching assistants and learners, and cohorts aim to include a mix of individuals from different backgrounds, institutions, career stages, and experience levels. The program has grown over time, and interest continues to increase each year. As a result, there are a limited number of slots available. Applicants are encouraged to re-apply in the next cycle if they are not selected in a given year.

For anyone who wants to apply these practices and principles in their research, but cannot commit to the full program schedule, all of the ReproRehab training materials have been made publicly available. This includes course content and other resources accessible through its [website](#), [GitHub](#), and [YouTube channel](#).

For Dr. Liew, seeing the ReproRehab community grow has made this one of the most rewarding parts of her career. This endeavor goes beyond helping individual researchers learn a new

programming language or complete a particular analysis. She reflects on the potential for researchers who participate to bring the same practices into their labs, incorporate this knowledge in mentoring their trainees, and contribute to a broader shift toward more open, collaborative, and reproducible rehabilitation science.

You can learn more about ReproRehab, the course content, resources, past and present cohorts, and the application process by visiting the [ReproRehab website](#).