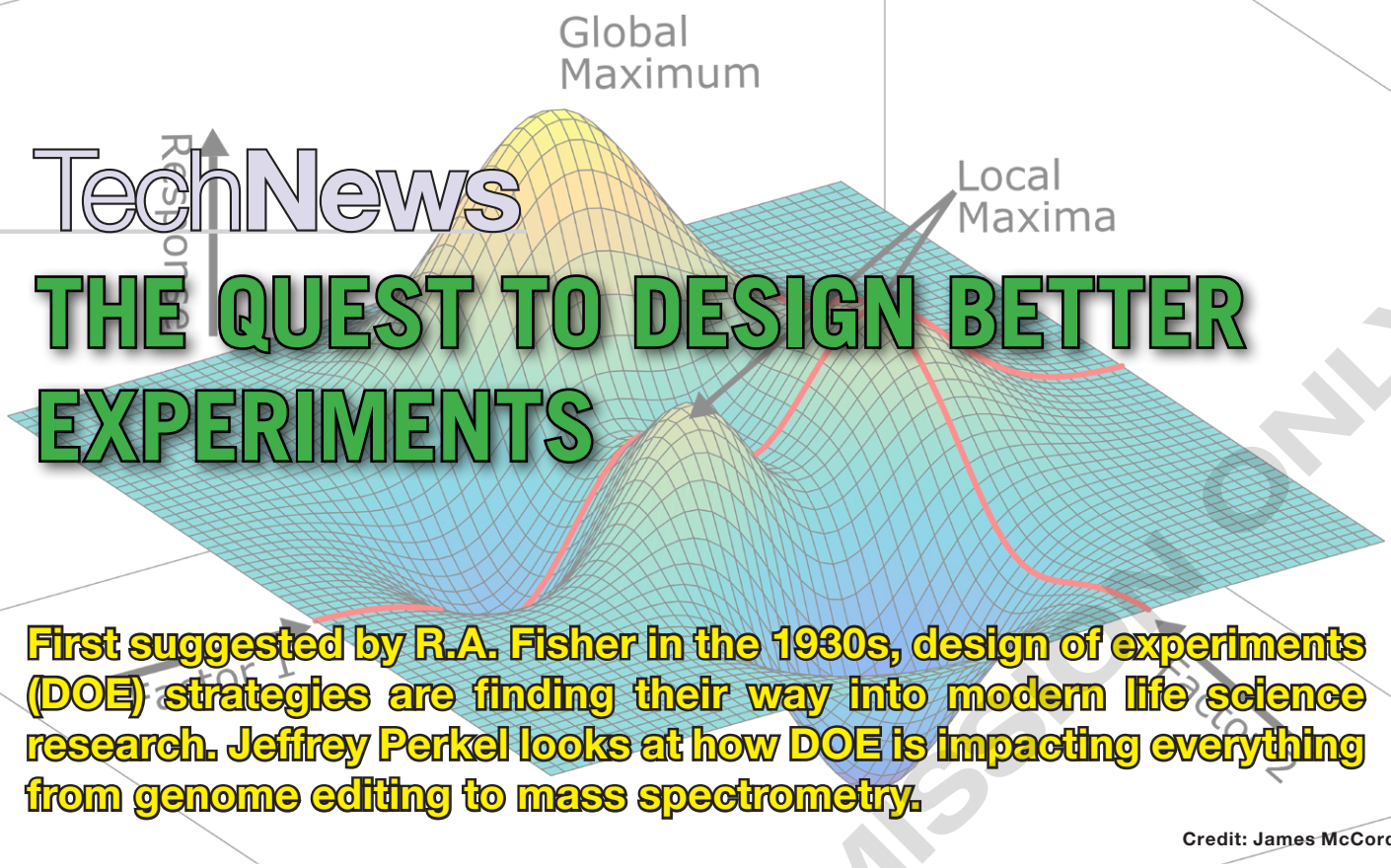
THE QUEST TO DESIGN BETTER EXPERIMENTS



First suggested by R.A. Fisher in the 1930s, design of experiments (DOE) strategies are finding their way into modern life science research. Jeffrey Perkel looks at how DOE is impacting everything from genome editing to mass spectrometry.

Credit: James McCord

Ramya Kumar was starting her doctoral research at the University of Michigan when she was tasked with a frustrating project. Kumar, a chemical engineering student, was working in the lab of professor Joerg Lahann. In 2010, Lahann's research group described a fully synthetic cell culture plate coating called PMEDSAH, which is capable of supporting indefinite growth and proliferation of human embryonic stem cells in defined media. Previously, stem cell biologists were forced to culture pluripotent stem cells on murine tumor extracts such as Matrigel or on murine embryonic fibroblast monolayers, both of which are chemically undefined and variable.

PMEDSAH—more formally known as poly[2-(methacryloyloxy)ethyl dimethyl-(3-sulfopropyl)ammonium hydroxide]—promised to change the uncertainty of stem cell culture. Animal product- and pathogen-free and fully chemically defined, it could be precisely synthesized and indefinitely stored. But Lahann's team still needed to explore every facet of the polymer, including how best to synthesize it.

Over the next several years, they probed PMEDSAH's properties. "We found there is actually a very interesting relationship between some of the physicochemical properties [of the polymer] … and the way they expand pluripotent stem cells," Lahann says. What wasn't clear, however, was precisely how to create polymers with a given set of properties.

Enter Kumar. She started working with the team's cell biology collaborators, the ones who were actually using the coating to culture stem cells. They would report back on each batch's performance, asking for specific tweaks. But with no obvious map to guide her, Kumar was effectively making changes in the dark. "The way I was going about making coatings was very trial-and-error based," she recalls. "I would try a certain set of reaction conditions; sometimes it would work, sometimes it wouldn't. And I had to go through this very iterative process to get the properties I wanted." For every experiment that worked as planned, four or five didn't. "I felt there had to be a better way," she says.

As it turns out, there was: A century-old strategy known as "design of experiments" (DOE).

Optimizing experiments

For most life scientists, experimental optimization involves the systematic dialing in of control parameters, adjusting one lever at a time while holding the others constant until a desired outcome—maximal protein production, for instance—is achieved.

Suppose, explains David Muddiman, Jacob and Betty Belin professor of chemistry at North Carolina State University, who wrote a recent review on DOE, a researcher wanted to examine two factors that affect the outcome of an experiment: temperature and chromatographic flow rate. "What people will do is they will study temperature, and they'll do it very nicely. They'll take the temperature from room temperature up to 100 degrees Celsius, and they'll do it in steps of maybe 10 degrees. And they'll say, oh, the optimal temperature is 50 degrees."

Then, while holding that temperature constant, they adjust flow rate. "They'll take the flow from 100 nl/min to 5 µl/min, and they'll step that at maybe 500 nl/min increments. And they'll say, oh, the optimal flow rate is 1.2 µl/min."

There are two problems with this approach. First, many experiments depend on far more than two variables. It's relatively trivial to test three variables at each of three levels—that's just nine trials. But some studies require tweaking dozens of variables, the systematic assessment of which rapidly becomes impractical.

More significantly, however, the approach presupposes that each factor is entirely independent of every other one, and that's not always the case. It may turn out, for instance, that temperature and flow rate interact to produce unexpected outcomes, a so-called "factor–factor" interaction. In that case, the actual best performance may occur at a flow rate of 2 µl/min and a temperature of 70 degrees—a condition that would never be detected using a one-factor-at-a-time approach.

DOE is a blanket statistical approach where researchers design a matrix with all of the variables they wish to test, the ranges over which they want to test them, and the granularity of those parameters (for instance, whether they are relatively continuous—such as testing at 10, 20, 30, 40, and 50 degrees—or coarse—for instance testing only high and low flow rates. They plug those variables into any of a number of commercial statistical software packages, which then direct the researcher to conduct a series of experiments. The researcher inputs the data back into the software, and voila! The software computes a linear regression from which it can predict the conditions to achieve a desired outcome.

Researchers can use either of two designs, says Helene Cardasis, a senior research scientist at Thermo Fisher Scientific. In a "full factorial design," every possible combination of parameters is tested—a rigorous, if exhausting strategy. The alternative is a fractional design, which allows researchers to sample a diverse set of variables "without preconceived notions that a parameter is or is not important," says Muddiman. Cardasis, with Shan Randall, used that latter strategy during a project they conducted with Muddiman to optimize data acquisition on a quadrupole-Orbitrap hybrid mass spectrometer—identifying key variables users should consider in their own optimization work, such as maximum ion injection time.

As Muddiman explains it, imagine the universe of all possible reaction conditions for a given experiment as a rugged landscape of mountains, plains, hills, and valleys. Using one-factor-at-a-time, researchers might identify a local maximum, the highest elevation reachable from their current position. But that's different from the global maximum—the very best outcome possible. "You might be missing an order-of-magnitude gain in your experiment." DOE can help find that global maximum.

Roy Goodacre, professor of biological chemistry at the University of Manchester, used a DOE strategy to optimize his SERS studies. Credit: R. Goodacre.

Sometimes, researchers can even see it directly—or at least, graphically. Roy Goodacre, professor of biological chemistry at the University of Manchester, UK, harnesses DOE to optimize surface-enhanced Raman scattering (SERS) studies. As part of that work, he builds Pareto graphs of signal strength versus reproducibility—which resembles an arc on which each experimental outcome is plotted. "You can decide where on the so-called 'Pareto front' you want to be," Goodacre explains.

In 2012, Goodacre's team used that strategy to optimize SERS detection of a beta-blocker called propranolol. The problem, he explains, is in the design of the surface itself—there are many ways to create the colloid that forms the basis of the SERS detection. In total, they considered 8000 possible permutations, of which they ran just 315—a time commitment of just 7 days, rather than 8 months. Still, it was enough to hone in on the best conditions. "You can find spectra that are absolutely beautiful in terms of both reproducibility and signal intensity," he says.

Experimental heresy

For biologists trained in the philosophy of one experiment, one parameter, DOE can seem downright heretical, says Todd Sandrin, professor of microbiology at Arizona State University. Sandrin is an environmental microbiologist who investigates methods to rapidly identify and discriminate bacteria in the environment and the clinic.

Suppose, for instance, that testing revealed the presence of *E. coli* in the water. "Did it come from a wildlife source, such as avian sources, or did it come perhaps from a wastewater treatment plant? Obviously, responses to each of those might be very different."

Sandrin has spent the better part of a decade developing and optimizing methods to draw those distinctions, using high-speed, high-throughput matrix-assisted laser-desorption ionization time-of-flight (MALDI-TOF) mass spectrometry as his technology platform of choice.

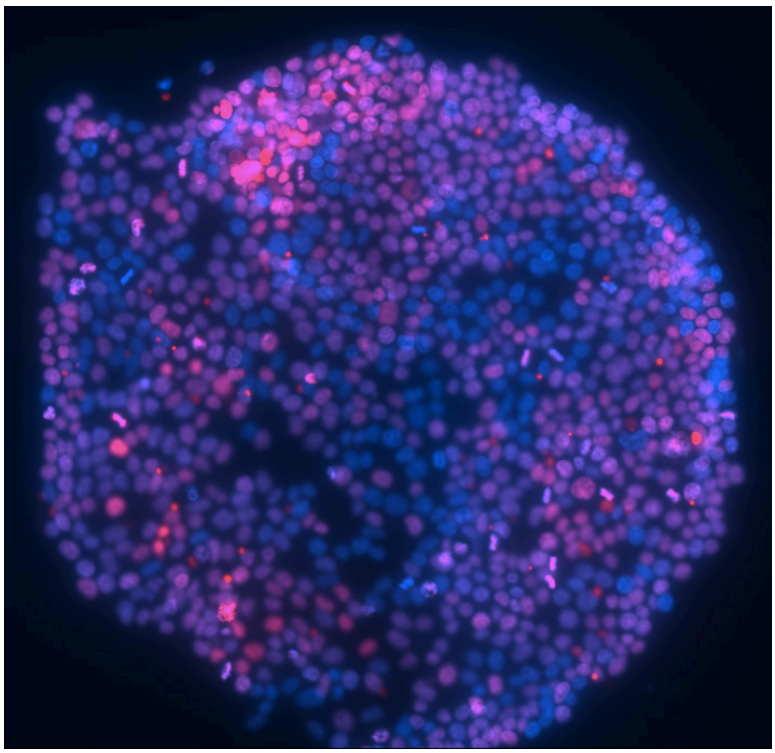
Automation is key. "You could take 100 or 200 or 300 protein extracts of different bugs, put them in the mass spectrometer, hit 'go,' and then come back a couple of hours later and have fingerprints of all those different microbes. And so automation was wonderful," Sandrin says. And yet, paradoxically, that very same automation seemed to negatively impact reproducibility. Sandrin and his team wanted to know why.

At first, they assumed that some preset variable was a little bit off, so his team set out to rectify the problem systematically. But no matter which variables they tweaked—such as laser intensity or number of laser shots per sample—nothing worked. Then Connie Borror, a statistician and close colleague at ASU, suggested DOE.

"I kind of raised an eyebrow, I think, and said, that's almost heresy," Sandrin recalls. "As a microbiologist, we adjust one variable at a time, and it takes a long time. And she said, you don't have to."

The team's DOE design tweaked 5 variables over 19 trials, focusing on the reproducibility of spectra of the opportunistic pathogen *Pseudomonas aeruginosa*. The total time expenditure was just several hours of instrument time—far less than the several weeks he anticipated spending with the one-parameter-at-a-time approach.

In the end, the change was worth it, and according to Sandrin, the final data revealed some unexpected findings. For one thing, though reproducibility tends to rise with increasing laser shots at the MALDI target, there is a limit. "About 500 shots is where we started to reach that plateau of number of shots versus reproducibility," he says. More surprisingly, spectral reproducibility tended to increase as peak resolution decreased—an observation that Sandrin calls "wildly counterintuitive."



Results from DOE optimization experiments for CRISPR/Cas9-mediated genome editing. The desired outcome is loss of red fluorescent protein expression. Credit: B. Steyer.

"What that suggested to us right away was that we were probably overemphasizing the importance of resolution." By tweaking their acquisition software accordingly, Sandrin's team boosted the reproducibility of automated data collection from 90% to 97%. And not only for *Pseudomonas*. "When we apply those same optimized settings to *Klebsiella* and *Serratia*, you can see the same tightening up of the data," he says.

Emerging applications

DOE is particularly popular in the worlds of engineering, pharmaceuticals, cosmetics, and the like, where it is used especially to optimize manufacturing processes. But as Sandrin's experience shows, the method seems to be slowly working its way into the life sciences.

Benjamin Steyer, an MD-PhD student at the University of Wisconsin, Madison, used DOE to optimize CRISPR/Cas9-mediated genome editing in new cell types using nonviral liposomal transfection, with the ultimate goal of translating the technique to the clinic. "Having the highest gene knock-out or gene correction [rate] in a specific organ, like the eye or liver, may mean the difference between a successful therapy and an unsuccessful one," he explains. CRISPR/Cas9 may seem an odd choice for DOE—countless researchers have successfully applied it in their labs—yet as a technique, it is "much more parametrically complex" than, for example, RNA interference, Steyer says, as Cas9 and sgRNA "can be delivered in different forms and at different molar ratios." Among other findings, Steyer's DOE analysis demonstrated that the key factor in optimizing delivery in his hands was not the nucleic acid concentration at all but instead the amount and type of transfection reagent itself.

Synthetic biologist Douglas Densmore of Boston University uses DOE to aid in the design and optimization of synthetic genetic circuits and biosynthetic pathways—an approach he and his team, working with Eric Young, Christopher Voigt, and D. Benjamin Gordon at MIT, codified in a new software tool called Double Dutch (www.clothocad.org/doubleDutch/).

Synthetic biology is fertile ground for DOE, says Nicholas Roehner, a postdoc in the Densmore lab and lead author on the Double Dutch project. One of the goals of synthetic biology is to develop tools and component libraries so that a researcher can rationally design complex genetic circuits that will behave as anticipated. In practice, that process is complicated by the fact that some components behave differently depending on their context. As a result, there's usually trial and error in-

volved—and as the circuits get ever more complicated, sometimes including up to 20 genes and associated control elements, those strategies can quickly become unmanageable.

DOE, Roehner explains, "can give you that regression model that lets you make some rough predictions" about what combinations to try instead. But the output of a DOE is typically a table of numbers, such as -1, 0, and +1, which correspond to the parameter levels the user wants to test. Double Dutch is a web-based tool that converts that table into instructions for building the commensurate set of genetic circuits. It tries to balance the need to use the fewest number of circuit elements with other requirements, for instance, the fact that the same "strong" promoter cannot be used over and over again in a single circuit because of the possibility of homologous recombination in vivo. "You want to use as few parts as possible but at the same time not repeat parts within the different pathways," Roehner explains.

A programmer and "computationalist" with a background in bioengineering, Roehner was comfortable with statistics when he entered the Densmore lab, but he had never used DOE. "It sounded a little to me like cheating," he admits. When he joined the lab, he envisioned building rigorous mathematical models based on precise measurements of all of the thermodynamic and kinetic variables that might possibly drive a particular biological system. Then reality hit.

David Muddiman, the Jacob and Betty Belin professor of chemistry at North Carolina State University, has applied DOE to mass spectrometry methods. Credit: North Carolina State University.

"These days," he says, "I really think all the different approaches of going about designing biological systems—whether you're doing DOE, whether you're doing more complicated modeling of genetic circuits, or you're doing, say, a directed evolution approach—I think are all very complementary to each other. They all have their individual strengths and weaknesses, and each one is better at solving some problems than the other."

Parameters, parameters

For Kumar, however, DOE already had a proven track record. Prior to joining the PhD program in Michigan, Kumar had worked at a drug company in India on pharmaceutical process development and product scale-up. "I had to use DOE extensively as part of my job there," she says.

By the time she arrived in Ann Arbor, she had "forgotten all about" DOE though—until she ran headlong into PMEDSAH. "It occurred to me that DOE has not been considered at all as a property-prediction tool in biomaterials," she says. "And I thought, this could be an interesting material to demonstrate the uses of DOE."

Rather than running the idea by Lahann, Kumar forged ahead, devoting a summer to the project. She focused on three variables—overall catalyst quantity, the ratio of one catalyst to another, and reaction time. In total, she tested 45 variable combinations, searching for impact on polymer layer thickness and wettability. The analysis suggested that, contrary to conventional wisdom, catalyst concentration had a big impact on the resulting polymer architecture.

She presented the work in a group meeting as a *fait accompl*i. "I didn't even know she was working on that," Lahann says, "she just presented basically almost a kind of a rudimentary story of how the paper would look like."

Lahann thinks the work represents a novel application of DOE. The approach is commonly used in formulation optimization, small molecule synthesis, and batch chemistry, he says. But, "to my knowledge, this is one of the first examples where somebody really uses it to optimize … polymers on surfaces."

As an MD-PhD student at the Univeristy of Wisconsin, Benjamin Steyer used a DOE approach to boost his CRISPR/Cas9 experiments. Credit: B. Steyer.

Kumar proved her model worked by correctly predicting conditions that could create polymer designs the team had previously discovered via trial and error. Then Lahann asked her to put the model to the test by predicting conditions to make a polymer layer that the group otherwise could not make—a thin but highly hydrophobic layer with a different internal architecture. Again, the model delivered.

"We can now deliberately access different regimes, different architectures in those polymer coatings at will. And by doing that, we can set very different properties of those coatings predictively. And that's really amazing," Lahann says.

Those findings energized the lab. One member has now applied DOE to optimize microparticle capsules for cells, while another is using it to optimize the growth of cancer cells in a bioreactor.

Anyone can do DOE, Kumar says—the statistics aren't difficult, and there are tutorials and software to help. All that's required, she concedes, is time—and a good working knowledge of your experimental system. "You have to be a real expert in the process you're studying."

Written by Jeffrey Perkel, Ph.D.

BioTechniques 61:9-14 (July 2016)
doi: 10.2144/000114430