

Functional MRI Research Findings May Be Tantalizing, But They May Not Be True, Study Says

BY SUSAN FITZGERALD

ARTICLE IN BRIEF

A new analysis suggests that software problems may be responsible for as much as a 70 percent false-positive rate in functional MRI data, indicating that there is increased brain activity in a brain region under study when in fact there is not.

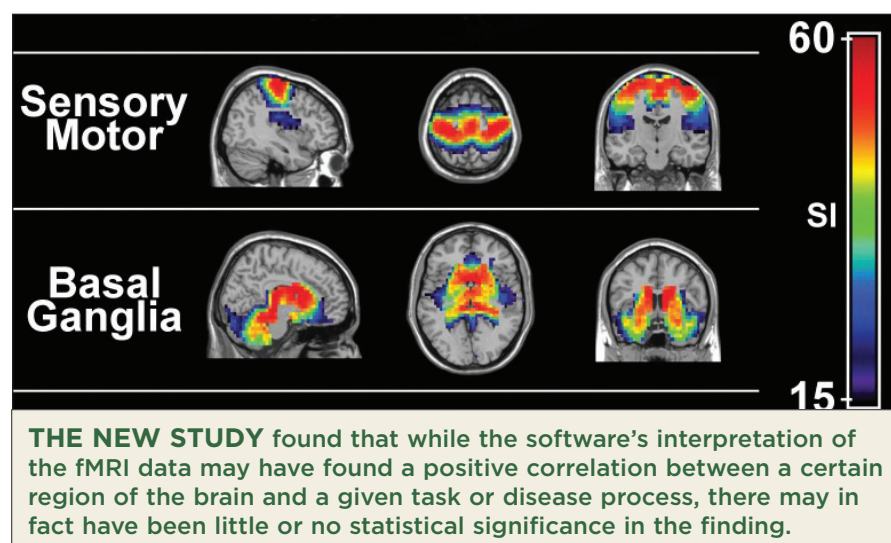
While it seems as though functional MRI (fMRI) can provide an answer to just about any question about brain function, a new analysis suggests that a portion of the findings published over the past 20 years may be questionable or outright wrong because of flaws in software programs used to interpret fMRI data.

The software problems can result in “false positive” rates of up to 70 percent, indicating that there is increased brain activity in a brain region under study when in fact there is not, according to an analysis published July 12 in the *Proceedings of the National Academy of Sciences* (PNAS).

The report initially concluded that as many as 40,000 fMRI studies may be dubious due to the software glitches, but the researchers later issued a correction that offered a less definitive estimate.

“These results question the validity of a number of fMRI studies and may have a large impact on the interpretation of weakly significant neuroimaging results,” said the researchers, headed by Anders Eklund, MD, PhD, of Linköping University in Sweden.

While the PNAS report does not repudiate the value of fMRI research, experts said it should serve as a cautionary note to researchers that they need to be meticulous in their research design,



statistical methods, and interpretation of data. They added that open sharing of fMRI data is needed so that outside researchers can better understand how published findings were arrived at and determine if results are reproducible.

“It would be misleading to say the majority of published (fMRI) studies are subject to this flaw, but it is fair to say a substantial fraction of studies could be problematic,” David Van Essen, PhD, Alumni endowed professor in the department of neuroscience at Washington University in St. Louis, told *Neurology Today*. Of particular concern might be findings that were only weakly statistically significant, he said.

Dr. Van Essen, who was not involved with the PNAS study, said he did not believe it would be worthwhile to go back and replicate past studies to see if they were valid or not, but rather take the lessons learned and conduct more rigorous research going forward.

“You need to get to know your data and the methods you use to analyze your data, and you need a critical understanding, as best you can, of the issues that lurk

that may have an impact on your careful interpretation of the data,” said Dr. Van Essen, who is a principal investigator for the Human Connectome Project, a major research effort to map brain connectivity. The project has made its methodologies and data available to other researchers.

STATISTICAL METHODS NOT VALIDATED

According to the PNAS report, there have been more than 40,000 published studies using fMRI to study brain function since the imaging tool came into use about 20 years ago. fMRI, which uses changes in blood flow to map brain activity, is used by neurologists and neuroscientists to study a variety of disorders, including epilepsy, Alzheimer’s disease, and brain injury.

fMRI has also become a popular tool in psychology and psychiatry research to explore how humans think, feel and form opinions. The public now is used to seeing colorful images of the brain “light up” while the person being scanned is considering such things as

sex, chocolate or gambling. One recent fMRI study showed the teenage brain devouring social media. The images give the impression, rightly or not, that the very spot in the brain where a specific thought or feeling arose has been pinpointed.

“Despite the popularity of fMRI as a tool for studying brain function, the statistical methods used have rarely been validated using real data,” the researchers reported in PNAS, noting that the methodologies rely on a “variety of assumptions.”

To test the reliability of three widely-used fMRI software programs (SPM, FSL and AFNI), the researchers used resting-state fMRI data obtained from brain scans of 499 healthy controls. They used the data to perform 2,880,000 group analyses to see how the group analyses differed based on the software package and software settings.

Given that all the scans were done while the subject was in a resting state, there should not have been any notable differences in brain activity found in the comparisons of the groups. “In theory, we should find 5 percent false positives (for a significance threshold of 5 percent), but instead we found that the most common software packages... can result in false-positive rates of up to 70 percent,” the researchers reported.

The problem stemmed from how the software interpreted thousands of bits of information on small units called voxels that are gathered during an fMRI scan and then put together to create a “picture” of brain activity. Simply put, software was not adequately controlling for differences in brain activity among voxels or clusters of voxels that could have occurred simply by chance. While the software’s interpretation of the fMRI data may have found

Continued on page 10

Human Connectome Project

Continued from page 7

It might also help us with other issues, such as early childhood education and the importance of nutrition. How do we take this information to provide insight into designing proper educational tools stimulating cognitive developments that have long-lasting effects? How do we provide insights into nutritional effects on

early brain development to lower income countries?

WHAT ARE SOME OF THE CHALLENGES IN DOING THIS STUDY?

The scan is very sensitive to movement, so when you scan children who are sick, they are typically sedated. But these are healthy children, so there is no sedation. And they’re children, so they

don’t follow instructions, and this is the challenge.

We have developed methods based on our experience on imaging typically developing children in the past ten years. When we’re working with newborns, the parents bring them in during nap time, so that allows us to do the imaging. But around two or three years old, it’s very, very difficult because they are too young to train, but too old to be napping. Therefore, a mock scanner, simulating a real

MR scanner, is used to train the subjects so they feel comfortable in a MR scanner and can hold still during the time of the imaging. We also ask parents to bring the child’s favorite DVDs so that they can watch movies while being scanned. •

LINK UP FOR MORE INFORMATION:

- The Baby Connectome Project.

Epilepsy Surgery Pays for Itself Within Nine Years, According to French Study

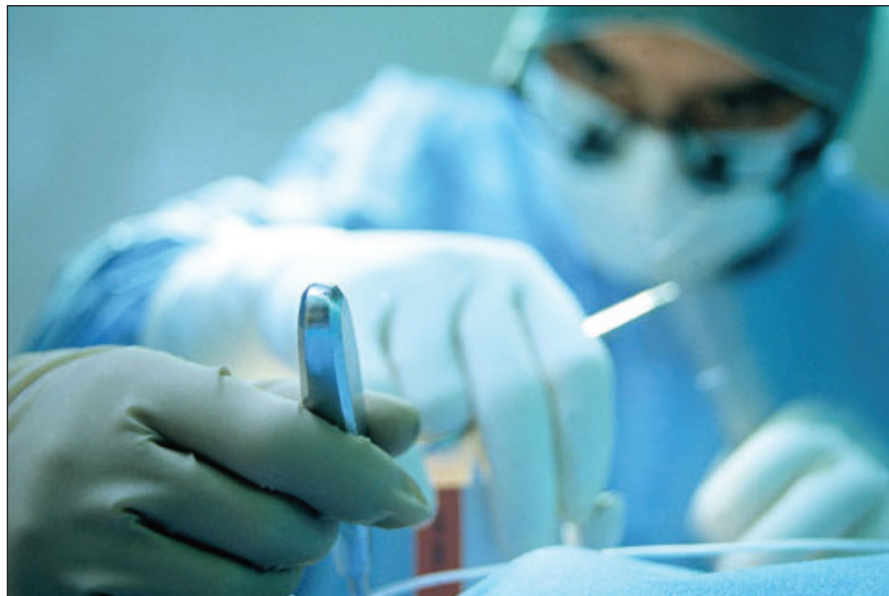
BY GINA SHAW

ARTICLE IN BRIEF

A prospective observational study of adult patients with partial intractable epilepsy in France found that direct costs of care became significantly lower the third year after epilepsy surgery compared with medical therapy alone.

Epilepsy surgery is cost-effective compared with continued medical care for people with drug-resistant partial epilepsy, paying for itself in the near term (within about nine years), according to a new study published in the September 5 online edition of *Epilepsia*.

The prospective observational study involved a cohort of patients recruited from 15 epilepsy units in France between 2001 and 2013. An original group of 289 underwent pre-surgical evaluation; of those, seven were wrongly included and 10 lost to follow-up. A total of 119 patients underwent surgery,



THE STUDY AUTHORS attributed the decline in costs after epilepsy surgery to sharp decreases in the costs of antiepileptic drugs (AEDs) and hospitalization in the surgical group from the second year, compared with the continuous increase in the costs of AEDs in the medical group.

88 were operable but did not have surgery, and 65 were not operable.

Direct medical costs for both groups peaked in the first year of

inclusion, with presurgical evaluation costing a mean of 9,073 euros in the surgical group and 6,089 euros in the control group. Mean costs for the surgical group during the inclusion year, with surgery included, were 21,517 euros.

Compared with the control group, mean direct medical costs were significantly lower in the surgical group from the second year ($p=0.02$), with a highly significant difference from the third year ($p < 0.001$). “This was due mainly due to the sharp decrease in costs of AEDs [antiepileptic drugs] and hospitalization in the surgical group from the second year, whereas costs of AEDs showed a continuous increase in the medical group,” wrote the authors, led by Marie-Christine Picot, head of the epidemiology unit at the University Hospital of Montpellier in France.

The researchers calculated the value of the discounted incremental cost-effectiveness ratio (ICER) — the extra
Continued on page 11

fMRI Research

Continued from page 8

a positive correlation between a certain region of the brain and a given task or disease process, there may in fact have been little or no statistical significance in the finding.

fMRI TECHNOLOGY STILL EVOLVING

Dr. Van Essen said fMRI is a “remarkable tool” tool that is advancing the understanding of brain functions. But he said many of the findings reported so far are akin to pinpointing a mountain on a map rather than providing a detailed description of the terrain.

“Knowing the locations of ‘peaks’ is only the starting point for getting a deeper understanding of the neurobiology,” he said.

Dr. Van Essen said “methodologies are improving rapidly” and some of the software issues have already been addressed. Still, “I would say in general that investigators using neuroimaging and readers of those studies need to be mindful of a host of issues that are becoming evident in the field and are still generally underappreciated.”

Prashanthi Vemuri, PhD, assistant professor of radiology at Mayo Clinic in Rochester, MN, who does imaging research on Alzheimer’s disease and mild cognitive impairment, said the reported concerns are important and “will make the field stronger.”

She said a challenge researchers face in interpreting data from brain scans is to sort out the “noise” from the “signal.” The so-called noise may be due to physiological fluctuations, scanner noise, measurement errors, or even some unconsidered or unknown factor, she told *Neurology Today*, which may lead to a false-positive finding that a certain region of the brain is showing activity during a given task or a lack of activity due to neurodegenerative disease.

Dr. Vemuri said she did not anticipate that major findings on AD would be overturned because of concerns raised in the new report, though she said that fMRI findings related to AD may be more reliable than those related to mild cognitive impairment because the so-called signal is stronger due to the advanced state of disease.

“I think more rigorous reporting of methods and doing more sensitivity analyses to make sure the findings stand” will increase confidence, she said.

The report also serves as a reminder that fMRI is still largely considered a research tool, not a definitive diagnostic one. While fMRI has the potential to diagnose neurodegenerative disease early or predict whether a patient will recover from traumatic brain injury, the technology is evolving.

“There is currently no clinical indication for the use of fMRI in the evaluation of patients in impaired states of consciousness — coma, persistent vegetative state, or minimally conscious state — and it should not be used to influence clinical decision-making,” said David Thomas Jones, MD, assistant professor of neurology and radiology at the Mayo Clinic in Rochester, MN.

Lorina Naci, PhD, the L’Oreal for Women in Science Research Excellence Fellow at the Brain and Mind Institute at the University of Western Ontario, said the *PNAS* report has brought into focus methodological problems that have been increasingly recognized in the fMRI field. Specifically, it calls attention to erroneous assumptions about how brain activity in one voxel relates to activity in nearby ones.

“If we find one voxel that shows activity, what is the likelihood that the voxel next door will be active simply by chance?”

Dr. Naci asked. The tendency has been to underestimate the chance factor, she said, and estimate a higher-than-reality likelihood that neighboring voxels, or brain activity clusters, are due to whatever task or disease process is being studied.

Dr. Naci, who does fMRI research involving brain-injured patients in non-communicative states, said the explosion in fMRI studies on headline-grabbing topics may have helped create the impression that the technology has an unlimited ability to reveal the brain’s most intricate workings and secrets.

As with any field of science, she said there needs to be some healthy skepticism about fMRI so that the ultimate goal of devising “better treatments for patients” can be achieved. •

LINK UP FOR MORE INFORMATION:

- Eklund A, Nichols TE, Nuttson H. Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates. *Proc Natl Acad Sci USA* 2016;113(28):7900-7905.
- Iyengar S. Commentary: Case for fMRI data repositories. *Proc Natl Acad Sci USA* 2016; 113(28):7699-7700.