Flow Diversion in the Treatment of Intracranial Aneurysm Trial

TO THE EDITOR: We read with interest the recent paper by Raymond et al.8 (Raymond J, Gentric JC, Dursaut TE, et al: Flow diversion in the treatment of aneurysms: a randomized care trial and registry. J Neurosurg [epub ahead of print November 4, 2016; DOI: 10.3171/2016.4.JNS152662]). The authors conclude that “Flow diversion was not as safe and effective as hypothe-sized. More randomized trials are needed to determine the role of flow diversion in the management of aneurysms.” We respectfully disagree with these conclusions.

Randomized controlled trials (RCTs) are held up as the pinnacle of clinical research. Many previous investigators as well as regulatory agencies have decried the dearth of RCTs evaluating flow diversion. The challenges in conducting RCTs in the setting of emerging technologies are well known to us, and we congratulate the authors on their ability to enroll patients in an RCT. However, not all RCTs are created equal, and a substantial degree of “pushback” has emerged by authors critical of the potential limitations of RCTs.3 Chief among these critiques is that the gauntlet of selection criteria in many RCTs results in a rarified population that is not representative of usual practice. In order to overcome this specific limitation, many investigators have proposed implementation of “pragmatic RCTs,” which have relatively broad inclusion criteria to encompass a wide array of patients. As a result, most current pragmatic RCTs are very large trials that typically utilize the electronic medical record to sweep up vast amounts of clinical data on large numbers of patients. Granted, the MR CLEAN trial was described by its investigators as “pragmatic,” based on the fact that a wide range of NIHSS scores and endovascular devices were included.1 However, in the MR CLEAN trial inclusion and exclusion criteria were still precisely defined, a factor that was likely responsible for its positive result.

Raymond and colleagues have recently introduced into the medical lexicon the idea of the “randomized care trial.”5 This term seems new, and both Google and PubMed searches fail to uncover a single prior publication using that specific term. The authors have previously claimed that “care” trials can cure nearly every ill of the RCT. They have written, “Care trials protect present pa-
tients from both unverifiable medicine and research performed for extraneous interests. They provide prudent care when evidence is lacking…. Care trials can identify which medical alternative should be standard therapy. In the meantime, they provide optimal care in the presence of uncertainty.”6 Unfortunately, there is little evidence to substantiate these claims.

The FIAT (Flow Diversion in Intracranial Aneurysm Treatment) trial is a small trial with unclear enrollment criteria that somehow captured a predominance of lesions that have been known, based on much, much larger previous trials, to be of high risk for flow diversion. Indeed, if the FIAT investigators had been in charge prior to regulatory approval of flow diversion we would never have learned that carotid artery aneurysms that were previously very difficult to treat can be easily cured with low complication rates, even with first-generation flow diversion technology. The investigators and regulators who designed and oversaw the PUFS (Pipeline for Uncoilable or Failed Aneurysms) trial understood the critical importance of clear enrollment criteria, unmet clinical needs, and rational design of trials.1

What were the enrollment criteria for FIAT? As written in the paper, “All patients harboring an aneurysm for which flow diversion was considered a promising treatment were eligible to participate.” Who decided that flow diversion was promising for these patients? What were their specialties? Who are these patients? What was considered “promising?” How many patients were excluded and why? And how did they end up with cohorts of patients known to be at high risk of treatment with an any device, with more than one-third of patients randomized to flow diversion harboring aneurysms distant from the internal carotid artery (ICA)? Further, the complications were concentrated in known “high-risk” aneurysms for flow diversion, including ruptured, distal anterior, and posterior circulation aneurysms. Indeed, rather than eliminating bias and providing data on usual care patients, FIAT reports on complications in very small cohorts of very rare aneurysms. How does this inform the community on the usual patient with small or medium-sized anterior circulation aneurysms? It doesn’t.

Just as not all RCTs are created equal, not all aneurysms are created equal. The natural history, treatment efficacy, and complication rate vary widely among aneurysms of different size, territory, and presentation. We fully agree
with the FIAT authors’ suggestion that we remain in the early stages of investigation regarding flow diversion. Accordingly, it seems premature to shift from explanatory trials—focused on clearly selected, relatively homogeneous groups of patients—to all-encompassing pragmatic trials. Trials such as PUFS, ASPIRE, and PREMIER have enrolled well-defined groups of patients that likely allow “generalization” to other, similar aneurysms. Almost counterintuitively, the results of the “pragmatic” FIAT trial, instead of being “generalizable,” are impossible to apply to any patient group, given unclear enrollment criteria and substantial heterogeneity.

David F. Kallmes, MD
Waleed Brinjikji, MD
Alejandro A. Rabinstein, MD
Mayo Clinic, Rochester, MN

While a received view and the current regulation separate clinical interventions and comparison of patient outcomes, one within the domain of care and the other within the domain of research, care trials assess the outcome of interventions as they are practiced in the best medical interest of participants. Care trials are pragmatic as opposed to explanatory trials, design concepts introduced in 1967. Kallmes and colleagues appear to think that it is too early for pragmatic trials because of a belief that, in evaluating innovative treatments, there is a proper chronology of methods based on the selection of patients: “it seems premature to shift from explanatory trials—focused on clearly selected, relatively homogeneous groups of patients—to all-encompassing pragmatic trials.”

To discuss this issue, we have ordered various research methods (preclinical, clinical, and epidemiological) according to the selection of research subjects in Fig. 1. Dr. Kallmes participated in PUFS, which was a case series lacking a comparator group that was used for device approval (Fig. 1B). If, as he and his colleagues assert, “the investigators and regulators who designed and oversaw the PUFS trial understood the critical importance of … rational design of trials,” we believe that they would have done more than rename their case series, the weakest method, as a clinical trial (Fig. 1C). Ethical care using innovative devices remains research, and the protection of participating patients requires randomized allocation, as we will see shortly.

The authors’ approach is to start with the preclinical method, since the idea that you can only learn from the careful selection of identical research subjects (of the same purebred species) comes from the animal laboratory (Fig. 1A). According to this line of thinking, the more selected and homogeneous your research subjects, the more generalizable the results of your experiment. In clinical medicine, the truth is exactly the opposite: the more select the patient group, the less generalizable the results. This common mistake confuses the problem of who should be studied in order for the results to be generalizable with the problem of how to re-identify the same patients in the future (to make clinical categories of “natural kinds”).

The authors complain of the “limitations of RCTs.” However, here we are confronted with our own human limitations. The reasoning that “RCTs are difficult—therefore let’s use an alternative” is common, but not the best solution. The complaint is typically followed by a good idea, like PROS (pragmatic registry-based observational studies), which is only a new name to re-introduce the same old case series and observational studies.

The authors’ second methodological choice, then, is epidemiological, as we know from past publications. However, even there, canons of statistical inference, such as random sampling to validly infer that results from the sample apply to the population, are rarely if ever respected (Fig. 1B). The end result of starting with a small case series of selected patients, without a comparison group, followed by observational studies of treatments performed in all other patients, is that innovations are practiced just as if they were standard care and proper RCTs comparing the innovation with standard care are never performed.

References

Disclosures
Dr. Kallmes reports a consultant relationship with Medtronic and receipt of support from Medtronic, MicroVention, Sequent Medical, NeuroSigma, and Codman Neurovascular.

Response
We thank the authors for their interest in our study. They acknowledge “the dearth of RCTs evaluating flow diversion” and the “challenges in conducting RCTs in the setting of emerging technologies,” and it is precisely to address these challenges that the care trial methodology was conceived. In fact, FIAT is the prototype care trial.
Let’s compare a case series like PUFS with a randomized trial (FIAT). What did we learn from PUFS? Flow diversion in 108 or so carefully selected “uncoilable” proximal carotid aneurysms led to aneurysm occlusion in many patients (64%-81%); unfortunately, the treatment could also be harmful or fatal in other patients (2.6%-11.7%), with 44 serious adverse events, 3 major strokes, and 3 deaths reported. We could attack PUFS with the same criticisms leveled at FIAT: PUFS concerned rare aneurysms; the authors mixed extra- and intradural lesions, which we know from much, much larger experience differ in terms of presentation, natural history, response to various treatments, complication rates, etc. They did not distinguish those patients with a good circle of Willis, who

FIG. 1. Schematic representation of various ways to study innovative treatments and patient characteristics in preclinical or clinical studies.
could benefit from parent vessel occlusion, from those without a good circle of Willis or those with aneurysms that truly could not be treated with coil embolization, the true fusiform dilatations, from those that could have been treated with stents. This reasoning has no end. We explored the idea of designing multiple explanatory trials in which some colleagues said they would have participated, but close to a dozen trials would have been necessary just for flow diversion. This avenue was hopeless.

More importantly, how can the results of series like PUFS impact our practice? We are told that “if the FIAT investigators had been in charge prior to regulatory approval of flow diversion we would never have learned that carotid artery aneurysms that were previously very difficult to treat can be easily cured with low complication rates, even with first-generation flow diversion technology.” Not so. We can “learn” the exact same facts in FIAT, although the FIAT investigators were more careful than to conclude that patients are “easily cured with low complication rates.” FIAT indeed provides the same numbers as PUFS for the same patients, the proximal carotid artery aneurysms judged untreatable by other means, shown in Table 4 of our article (registry): aneurysm occlusion in 77%, morbidity-mortality in 4.5%. But more than raw numbers are needed before adopting a new treatment in practice.

FIAT can be analyzed according to as many predefined anatomic (or other) criteria as one wishes, but there is no need to hide the outcome of flow diversion for non-PUFS patients, for treatment “indications” typically evolve as soon as a device has been approved, as the authors demonstrate when they shift the subject matter from “[uncoilable] carotid artery aneurysms that were previously very difficult to treat,” introduced at the beginning of the letter, to “the usual patient with small or medium-sized anterior circulation aneurysms” at the end, or as they themselves witnessed when they published their observational studies on the use of flow diversion.9

Our provisional conclusions on flow diversion are more prudent because FIAT provides more than an observational registry. The question of interest to clinical care is: Does flow diversion improve patient outcomes? To that question, the registry portion of FIAT cannot provide an answer, any more than PUFS or other observational studies, because there are no clinical outcomes from standard treatment options with which to compare flow diversion. When an innovation becomes available we are quick to judge patients as “untreatable by other means,” and as we discussed at length in our article, this notion is just as unstable, subjective, and arbitrary as the notion of “uncoilable” in PUFS.

What sets FIAT apart is that it is a real trial. Obviously, all patients have other options, because just recently, flow diversion did not exist, and patients were still managed one way or another. In contrast to registries of selected patients, where selection serves to show that therapy can work in the most favorable cases, without a valid comparison, randomized allocation provides a means to measure whether it works better than what was previously available. Unfortunately, the only valid comparisons so far have yet to show that flow diversion improves outcomes for any verifiable kind of patients.

Questions raised by the authors about FIAT are answered in Table 1.

The results of the FIAT trial so far are perfectly generalizable. What are those results? That there is still a lack of evidence that flow diversion improves the outcome of any kind of patients, including those proximal carotid artery aneurysm patients that were used to introduce the Pipeline device. This is not to say that it will be impossible to show that flow diversion is beneficial; only that we need more randomized evidence before we can rationally adopt this treatment as better. Trials remain necessary if we want to offer innovative approaches, rather than standard treatment options, until they are shown to improve patient outcomes. For which kind of patients? How could anybody know before doing the trial? Professor Kallmes and colleagues can be assured that, as soon as better outcomes are shown for any group of patients, homogeneous or not, such as proximal carotid artery PUFS aneurysms or cavernous aneurysms, randomized allocation will be interrupted, and the innovation will be prescribed as the

### Table 1. Responses to the concerns raised about FIAT

<table>
<thead>
<tr>
<th>Question or Criticism raised about FIAT</th>
<th>Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>What were the enrollment criteria for FIAT?</td>
<td>“All patients harboring an aneurysm for which flow diversion was considered a promising treatment were eligible to participate.”</td>
</tr>
<tr>
<td>Who decided that flow diversion was promising for these patients?</td>
<td>Doctors who are currently and will be making these decisions in the future</td>
</tr>
<tr>
<td>What were their specialties?</td>
<td>Neurosurgeons and neuroradiologists with endovascular expertise</td>
</tr>
<tr>
<td>Who are these patients?</td>
<td>Everyday care patients</td>
</tr>
<tr>
<td>What was considered “promising”?</td>
<td>The “unmet clinical needs” mentioned in the letter</td>
</tr>
<tr>
<td>How many patients were excluded and why?</td>
<td>Only patients for whom flow diversion was not a consideration</td>
</tr>
<tr>
<td>How did they end up with cohorts of patients known to be at high risk of treatment with any device, with more than 1/3 of patients randomized to flow diversion harboring aneurysms distant from the ICA?</td>
<td>The trial is at the service of patients. As is obvious from the wording of the question, these poor patients are at high risk with any treatment; the clinician felt that flow diversion might offer them the best chance of a good outcome</td>
</tr>
<tr>
<td>Complications were concentrated in known “high-risk” aneurysms for flow diversion, including ruptured, distal anterior, and posterior circulation aneurysms.</td>
<td>Complications are always “concentrated” in “known high-risk” patients, whatever “high risk” means. This is why absolute numbers should not impact clinical decisions—rather, only relative risks from randomized trials</td>
</tr>
</tbody>
</table>
new standard treatment. In the meantime, how should we manage the next patient? If we are considering flow diversion, we should recommend trial participation.

Tim E. Darsaut, MD
University of Alberta Hospital, Mackenzie Health Sciences Centre, Edmonton, AB, Canada

Robert Fahed, MD
Centre Hospitalier de l’Université de Montréal, Notre-Dame Hospital, Montreal, QC, Canada

Jean-Christophe Gentic, PhD
Centre Hospitalier de l’Université de Montréal, Notre-Dame Hospital, Montreal, QC, Canada

Jean Raymond, MD
Centre Hospitalier de l’Université de Montréal, Notre-Dame Hospital, Montreal, QC, Canada

TO THE EDITOR: We would like to thank Foreman et al.1 for their study (Foreman PM, Griessenauer CJ, Kicielinski KP, et al: Reliability assessment of the Biffi Scale for blunt traumatic cerebrovascular injury as detected on computer tomography angiography. J Neurosurg 127:32–35, July 2017), which is a valuable effort to look at the reliability of CT angiography (CTA) for grading of blunt traumatic cerebrovascular injury (TCVI). CTA is increasingly being used for TCVI screening, despite its reported widely variable sensitivity and specificity.2 Digital subtraction angiography (DSA) continues to be considered the “gold standard,” although its use after CTA is heterogeneous, i.e., some groups using it in CTA report negative results (due to reported low sensitivity of CTA), while others advocating its use in CTA report positive results (due to high false-positive results with CTA).3 Reliability of grading on CTA would help determine the appropriate use of subsequent DSA.

In this study, the authors included the study patients from a data set in which all CTA-positive patients underwent DSA. There is inherent selection bias as there was no control group and the graders were aware that all the study cases were DSA-confirmed cases. Could the authors comment as to whether it is still their practice to perform DSA in low-grade injuries noted on CTA?

Most of the low sensitivity of CTA is ascribed to low-grade injuries.3 The authors state that they used 6-mm axial slices. Is there a reason for using such thick slices? One would assume that thin slices would be preferred for subtle changes in luminal caliber (Grade I injuries). The authors found the CTA grades matched the DSA grade in 75% of cases; could they comment on whether the mismatched grades were more frequent with the lower-grade injuries?

The challenge with noninvasive imaging is to identify and grade the lower-grade injuries (Grade I and II), especially at the skull base. Would the authors agree that if it is reasonable to combine the lower grades, there would be no reason to perform DSA unless an endovascular intervention is contemplated? The dynamic nature of the injury and progression of grade of injury is equally likely to be appreciated on repeat CTA. We would also appreciate if the authors could elaborate on their experience with repeat imaging and its utility.

Accurate identification and grading is particularly challenging in lower grades in the presence of atherosclerotic disease and artifacts, such as at the skull base and with

References

INCLUDE WHEN CITING
Published online April 28, 2017; DOI: 10.3171/2016.12.JNS163176.
©AANS, 2017

Biffi Scale for blunt traumatic cerebrovascular injury


INCLUDE WHEN CITING
Published online April 28, 2017; DOI: 10.3171/2016.12.JNS163176.
©AANS, 2017

Biffi Scale for blunt traumatic cerebrovascular injury

TO THE EDITOR: We would like to thank Foreman et al.1 for their study (Foreman PM, Griessenauer CJ, Kicielinski KP, et al: Reliability assessment of the Biffi Scale for blunt traumatic cerebrovascular injury as detected on computer tomography angiography. J Neurosurg 127:32–35, July 2017), which is a valuable effort to look at the reliability of CT angiography (CTA) for grading of blunt traumatic cerebrovascular injury (TCVI). CTA is increasingly being used for TCVI screening, despite its reported widely variable sensitivity and specificity.2 Digital subtraction angiography (DSA) continues to be considered the “gold standard,” although its use after CTA is heterogeneous, i.e., some groups using it in CTA report negative results (due to reported low sensitivity of CTA), while others advocating its use in CTA report positive results (due to high false-positive results with CTA).3 Reliability of grading on CTA would help determine the appropriate use of subsequent DSA.

In this study, the authors included the study patients from a data set in which all CTA-positive patients underwent DSA. There is inherent selection bias as there was no control group and the graders were aware that all the study cases were DSA-confirmed cases. Could the authors comment as to whether it is still their practice to perform DSA in low-grade injuries noted on CTA?

Most of the low sensitivity of CTA is ascribed to low-grade injuries.3 The authors state that they used 6-mm axial slices. Is there a reason for using such thick slices? One would assume that thin slices would be preferred for subtle changes in luminal caliber (Grade I injuries). The authors found the CTA grades matched the DSA grade in 75% of cases; could they comment on whether the mismatched grades were more frequent with the lower-grade injuries?

The challenge with noninvasive imaging is to identify and grade the lower-grade injuries (Grade I and II), especially at the skull base. Would the authors agree that if it is reasonable to combine the lower grades, there would be no reason to perform DSA unless an endovascular intervention is contemplated? The dynamic nature of the injury and progression of grade of injury is equally likely to be appreciated on repeat CTA. We would also appreciate if the authors could elaborate on their experience with repeat imaging and its utility.

Accurate identification and grading is particularly challenging in lower grades in the presence of atherosclerotic disease and artifacts, such as at the skull base and with
fracture fragments. The authors selected 40 patients, and it would be instructive if the authors could specify whether they excluded such patients. Only 6 of 40 patients in the study were Grade I, representing the distribution at the authors’ institution. However, larger published studies report a much higher percentage of Grade I injuries. The conclusion that CTA grading is reliable may be an overstatement based on the current study if 25% of cases were discordant with DSA.

Xiao Wu, BS
Ajay Malhotra, MD, MMM
Yale School of Medicine, New Haven, CT

References
2. Malhotra A, Wu X, Kalra VB, Schindler J, Matouk CC, For-

Disclosures
The authors report no conflict of interest.

Response
We appreciate the thoughtful critique by Mr. Wu and Dr. Malhotra regarding our study evaluating the reliability of the Biffl Scale for grading of TCVI. The primary objective of the paper was to assess the reliability of CTA in this setting, specifically the consistency of interpretation between different raters (intrarater reliability) and by the same rater (intrarater reliability). Reliability should not be confused with validity or accuracy and does not reflect a comparison with DSA. While sensitivity and specificity of CTA as compared to the DSA “gold standard” are critical for determining the usefulness of CTA, this was not the objective of the current paper.

A control group with negative screening CT angiograms was not included in our study. We acknowledge that this was indeed a regrettable limitation. The inclusion of a negative CTA control group would have shed light on the ability of CTA to distinguish common forms of TCVI from patients without TCVI. The relative distribution of injuries evaluated in the current study resulted from selection of patients with CTA- and DSA-proven TCVI. While the database included all injury grades, the requirement for both tests potentially selected for a greater number of high-grade injuries.

Our study found that CTA cannot consistently differentiate between Biffl Grade I, II, and III lesions. Although early reports using DSA indicated that stroke risk correlated with type of TCVI (e.g., that the risk of stroke is greater with Type II injuries compared with Type I lesions), we prefer to manage all patients with asymptomatic Grade I–IV lesions identified on CTA the same way, using antiplatelet agents and observation. Repeat CTA is rarely performed during the acute hospitalization and tends to be reserved for follow-up visits at 3 to 6 months afterward. DSA is reserved for cases in which a lesion has become symptomatic and endovascular treatment is anticipated.

In contrast to the situation with Grade I–III lesions, CTA can reliably identify Grade IV lesions (complete occlusion of the artery). Grade IV carotid artery lesions are both the least common and the most dangerous TCVI subtype. In terms of stroke risk, Grade IV vertebral artery injuries appear to occupy an intermediate position between Grade I–III lesions and carotid Grade IV lesions. Therefore, CTA can reliably identify TCVI lesions that carry the highest risk of stroke.

The slice thickness of the CTA images reported in the Methods section was erroneous. Axial slice thickness ranged from 0.7 mm to 2.5 mm, with sagittal and coronal slices ranging from 0.4 mm to 2.5 mm and 0.3 mm to 2.5 mm, respectively. We appreciate the opportunity to correct this mistake.

Paul M. Foreman, MD
Mark R. Harrigan, MD
University of Alabama at Birmingham, Birmingham, AL
Christoph J. Griessenauer, MD
Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA

References