"If you think of standardization as the best that you know today, but which is to be improved tomorrow, you get somewhere."  
- Henry Ford, 1863–1947

We now have a wealth of data from randomized trials that sought to compare carotid artery stenting (CAS) and carotid endarterectomy (CEA) as strategies intended to reduce intermediate and long-term risk of stroke ipsilateral to a significant carotid artery stenosis. If the data from these trials were wholly comparable on the basis of trial construct, baseline operator expertise, requirements for equipment used, choice of primary endpoint, and rigor of ascertainment of all chosen endpoints, then we would have one of the most perfect and complete combined datasets the endovascular community has ever seen, and there would probably be no need for ongoing discussion.

This critique seeks to explore the similarities and differences in the design and execution of these trials and attempts to highlight common themes and lessons learned.

**TRIAL SIMILARITIES**

**Independent Review**

Regardless of some differences in baseline surgical risk and majority population in terms of symptom status, and in common with most randomized trials of carotid intervention, the trials listed on page 60 (Randomized Trials for Standard-Risk/High-Risk Population) were all designed as one-to-one randomized constructs with intended independent review of outcomes (ie, review by stroke physicians or vascular neurologists who were not intimately involved in the performance of the carotid interventions being compared). The principal investigators (PIs) of the listed trials were all neurologists, with the exception of CREST (Carotid Revascularization Endarterectomy Versus Stenting Trial), which had a vascular surgical co-PI at its inception. The self-audit of outcome data in this clinical arena is notoriously unreliable.1 Death is, without doubt, a hard endpoint and an unarguable outcome, evident to the nonspecialist. However, minor neurological deficit that may otherwise affect the primary powered endpoint of any trial comparing strategies in carotid intervention, and may very well affect the patients’ quality of life, may be easily overlooked by the operator.

**TRIAL DISCREPANCIES**

**Operator Experience**

A cursory glance at the stroke/death outcomes in Table 1 would suggest substantially better results for carotid artery stenting (CAS) within CREST than within the European trials. This is partly explained by more stringent entry requirements for operators within CREST than in EVA-3S (Endarterectomy Versus Angioplasty in Patients With Severe Symptomatic Carotid Stenosis), ICSS (International Carotid Stenting Study), and, in its latter half, SPACE, and in part due to inclusion of asymptomatic patients.

In the European trials, there is evidence to suggest that there were temporal changes in interventionist entry crite-
ria, most likely resulting from attempts to expedite recruitment. In SPACE, after a relaxation of entry requirements, the adverse event rate rose. The original sample size for a noninferiority margin of 2.5% was 1,900 patients. A preliminary predefined interim analysis at 3 years (after the inclusion of 369 patients in each arm) demonstrated an overall event rate of < 3.9%, leading to a sample size recalculation of 1,200 patients to show noninferiority. At the second preplanned interim analysis after the introduction of a "preliminary certificate" (allowing interventionists to perform CAS within the trial after performing 10 cases), the sample size required to prove noninferiority was recalculated again (2,500 patients), presumably because of downsizing of absolute risk reduction in the CAS arm. The implication is that there must have been a substantial increase in the endpoint rate in the CAS arm after relaxation of the entry criteria. Notably, there was a linear relationship between numbers of patients recruited into this trial and the event rate for CAS. This was not the case for carotid endarterectomy (CEA), suggesting within-trial learning for CAS but not for CEA. The authors of EVA-3S and ICSS stated that there was no demonstrable influence of operator experience on outcome. However, the trials did not allow an adequately powered comparison of experienced and inexperienced operators, and 85% of operators within EVA-3S had performed fewer than 50 cases, the absolute minimal baseline experience for CAS according to expert opinion. Assessment of an individual surgeon’s CEA performance is thought to require ≥ 200 cases to ensure sufficiently narrow 95% confidence intervals. Tellingly, the relationship between trial entry requirement for interventionists within EVA-3S, ICSS and SPACE, and CAS outcomes within these trials is remarkably linear.

**Stroke/Death Ratios**

If the all-stroke to all-cause death ratio for CAS is com-

<table>
<thead>
<tr>
<th>Trial</th>
<th>Operator Requirements</th>
<th>CAS</th>
<th>CEA</th>
</tr>
</thead>
<tbody>
<tr>
<td>SAPPHIRE</td>
<td>Morbidity and mortality threshold, &lt; 6%. Median total volume, 64 (20–700).</td>
<td>Morbidity and mortality thresholds: &lt; 6% in symptomatic and &lt; 3% in asymptomatic patients. Median annual volume, 30 (15–100).</td>
<td></td>
</tr>
<tr>
<td>EVA-3S</td>
<td>Total experience, ≥ 12 CAS or ≥ 5 CAS + ≥ 30 non-CAS supra-aortic PTA/stents. Proctoring allowed. First ever CAS case could be performed within the trial with supervision.</td>
<td>≥ 25 CEA in the preceding year.</td>
<td></td>
</tr>
<tr>
<td>SPACEb</td>
<td>Total experience, ≥ 25 CAS. From 2002, &quot;certificate&quot; after 10 CAS cases. Proctoring allowed.</td>
<td>≥ 25 consecutive CEA with morbidity and mortality thresholds.</td>
<td></td>
</tr>
<tr>
<td>ICSSc</td>
<td>Thresholds set for morbidity and mortality (≤ 6%). Total experience, ≥ 50 PTA/stents (any territory); ≥ 10 CAS. CAS study day attendance (in an experienced unit) mandatory. Proctoring allowed. First ever CAS case could be performed within the trial with supervision.</td>
<td>Thresholds set for morbidity and mortality (≤ 6%). ≥ 50 CEA (annual rate, ≥ 10/y).</td>
<td></td>
</tr>
<tr>
<td>CREST lead-in</td>
<td>Total experience, ≥ 15 CAS with low complication rates and evaluation of technique. Candidates who were approved were trained in CREST study devices (animal models and participation in the Carotid Stent Operators Certification Programme). Those with ≥ 30 case experience performed 5–10 cases in the lead-in phase; those with &lt; 30 cases performed 10–20 in the lead-in.</td>
<td>The outcomes from the last 50 CEA procedures were reviewed by the surgical credentialing committee.</td>
<td></td>
</tr>
<tr>
<td>CREST randomized</td>
<td>Accreditation via the CREST lead-in phase.</td>
<td>Thresholds: &lt; 3% in asymptomatic patients and &lt; 6% in symptomatic patients &gt; 12 cases per year.</td>
<td></td>
</tr>
</tbody>
</table>

4Procedures performed in the presence of a physician instructor.
5SPACE data include 51 patients treated by interventionists with a “preliminary certificate.”
6ICSS data include 199 patients randomized at “probationary” centers (those who were being proctored and who had performed < 10 cases).
pared across the trials, the figure is reasonably consistent (see Randomized Trials for Standard-Risk/High-Risk Population chart). There are two notable outliers; in SAPPHIRE, the ratio is small (3) perhaps because this trial ensured a very high level of operator expertise (Table 1) and recruited a majority asymptomatic population. In EVA-3S, the ratio is 11.5, higher than in the other trials. Arguably, expertise in carotid stenting is reflected by the stroke/death ratio, and we can see from Table 1 that EVA-3S had the most lax trial entry requirements for CAS (but not for CEA).

Inclusion of Asymptomatic Patients
EVA-3S, SPACE, and ICSS recruited symptomatic patients exclusively. Of a total of 2,502 patients randomized within CREST, 1,181 were asymptomatic. This may have diluted the procedural risks across this trial for both CAS and CEA, because asymptomatic patients can expect lower procedural hazards than symptomatic patients during carotid intervention.6-8

Myocardial Infarction
Myocardial infarction (MI) was included as an outcome event in EVA-3S, ICSS, and CREST, although it was not evaluated in SPACE. In ICSS, the primary endpoint was the difference in fatal or disabling stroke in any territory at 3 years, and this has yet to be reported (an interim analysis reflecting 120-day safety data has been published).9 In EVA-3S, MI was considered a secondary endpoint. In CREST, MI formed part of a composite primary endpoint. There is a substantial disparity among the trials with respect to the magnitude of difference in MI rates between CAS and CEA limbs (Table 2). This raises the question of ascertainment bias. Scrutiny of this bias requires both a thorough evaluation of the definitions of MI employed within the trials but also an exploration of the adjudication process around independent review of this endpoint.

EVA-3S. MI was defined by at least two of the following criteria: typical chest pain lasting 20 minutes or more; serum levels of creatinine kinase MB or troponin at least twice the upper limit of normal range; and new Q waves on at least two adjacent derivations or predominant R waves in V1 (R wave ≥ 1 mm > S wave in V1).

### Table 2. MI as an Outcome Event

<table>
<thead>
<tr>
<th></th>
<th>CEA (%)</th>
<th>CAS (%)</th>
<th>RR/HR (95% CIs)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>EVA-3S</td>
<td>2/259 (0.8)</td>
<td>1/261 (0.4)</td>
<td>0.5 (0.04–5.4)</td>
<td>0.62</td>
</tr>
<tr>
<td>ICSS</td>
<td>4/857 (0.5)</td>
<td>3/853 (0.4)</td>
<td>Not recordedc</td>
<td>NS</td>
</tr>
<tr>
<td>CREST</td>
<td>28/1240 (2.3)</td>
<td>14/1262 (1.1)</td>
<td>0.5 (0.26–0.94)</td>
<td>0.03</td>
</tr>
</tbody>
</table>

a30-day outcomes.
b120-day outcomes.
cNot recorded as an individual component endpoint.
Abbreviations: CI, confidence interval; HR, hazard ratio; RR, relative risk.

### Table 3. Cardiovascular Risk Variables in ICSS

<table>
<thead>
<tr>
<th>Risk Variable</th>
<th>CAS (%)</th>
<th>CEA (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cardiac failure</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>Angina</td>
<td>10</td>
<td>9</td>
</tr>
<tr>
<td>Previous MI</td>
<td>18</td>
<td>18</td>
</tr>
<tr>
<td>Previous CABG</td>
<td>13</td>
<td>14</td>
</tr>
<tr>
<td>Atrial fibrillation</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>Other embolic source</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

Abbreviations: CABG, coronary artery bypass grafting; MI, myocardial infarction.
ICSS. MI was defined as the presence of two of the following three criteria: specific cardiac enzymes more than twice the upper limit of normal; history of chest discomfort for at least 30 minutes; or the development of specific abnormalities (eg, Q waves) on a standard 12-lead electrocardiograph.

SAPPHIRE. The definition of MI in this trial was “a creatinine kinase < 2 times the upper limit of normal with a positive MB fraction.”

MI was included in a composite primary endpoint because the investigators understood that patients with carotid artery disease are also likely to be at substantial risk for MI.

CREST. MI was defined by a creatinine kinase MB or troponin level that was twice the upper limit of the normal range or higher according to the center’s laboratory, in addition to either chest pain or symptoms consistent with ischemia or electrocardiogram (EKG) evidence of ischemia, including new ST-segment depression or elevation of more than 1 mm in two or more contiguous leads according to the core laboratory.

The CREST criteria appropriately included non–Q wave MI, although the EVA-3S and ICSS protocols accepted only transmural MI as a relevant consideration. The Carotid Stenting Trials’ Collaboration (incorporating the PI of ICSS, SPACE, and EVA-3S) appears to view non–Q wave MI as irrelevant yet puts great faith in the findings of the ICSS diffusion-weighted imaging study in which much was made of an increase in new hyperintensities on diffusion-weighted imaging magnetic resonance imaging of the brain in the CAS limb compared with the CEA limb. The fate and clinical relevance of these lesions is not known with any degree of confidence. A recent study in which a fairly comprehensive cognitive function test battery was employed to compare outcomes in CAS and CEA patients versus a control population suggests that new brain lesions after carotid revascularization are not associated with cognitive decline.

The Carotid Stenting Trials’ Collaboration considered that the different rates of MI in the trials could be explained by the “substantially higher baseline prevalence of cardiovascular disease in CREST.” The devil is in the detail. Baseline cardiac comorbidity in CREST was described as “previous cardiovascular disease,” which is an
all-encompassing definition. Baseline rates for CAS and CEA were 42.4% and 45%, respectively. In EVA-3S, baseline cardiovascular risk was specified only as “previous MI,” thus explaining the low incident rates (10.7% and 13.1% for CAS and CEA, respectively). In ICSS, a number of cardiovascular risk variables were considered (Table 3).

Patients may clearly have had more than one risk factor, rendering a summation of percentage risk overly simplistic. Even allowing for the fact that patients with atrial fibrillation were excluded from CREST, the overall cardiovascular risk profile for ICSS looks similar to CREST.

Relevance of Periprocedural MI

A procedure-related MI (whether Q wave or non–Q wave) is a very relevant consideration for any patient. A non–Q wave MI will confer an increase in the risk of death by a factor of 6 and an increase in the 6-month risk of further MI by a factor of 27.14

Ascertainment and Adjudication

The CREST protocol mandated postprocedure screening for MI with EKG and enzyme assay. In EVA-3S and ICSS, it would appear that only patients with cardiac symptoms postprocedure were evaluated (by EKG and cardiac enzyme assay). A specific MI adjudicating committee oversaw outcome events in CREST; this was not the case for EVA-3S and ICSS.

Mandatory CAS equipment. Tightly proscribed trials ensure standardized practice within them but compromise generalizability. The regulatory and reimbursement environment in the United States dictates mandatory use of approved carotid stents and embolic protection devices (EPDs), and this often results in use of a particular manufacturer’s stent and EPD in combination. Although these constraints ensure majority operator experience of particular stent/EPD combinations within trials such as CREST, thereby mitigating against learning curve issues with particular devices, there are two important counter arguments.

First, lesion-specific stenting is becoming a firmly entrenched belief in the European Union. This means that operators select the stent system and EPD according to factors such as the patient’s anatomy, lesion characteristics, symptoms, and cardiovascular status. Of

(Continued on page 95)
course, this presupposes a level of baseline experience that supports this degree of complex decision-making. Many interventionists performing CAS, certainly within EVA-3S and ICSS, simply did not have the baseline experience to be able to make these sorts of value judgements about CAS equipment. Second, restricted use of stents/EPDs may limit the validity of inferences regarding generic CAS outcomes in the real world.

WEAKNESSES AS EXECUTED
Both EVA-3S and SPACE were stopped early (the former on the grounds of safety, the latter for reasons of insufficient funding and futility). Generic problems of truncated trials include credibility issues, imprecision (wide confidence intervals), and bias (suspended trials may stop on a “random high”). When noninferiority trials (such as SPACE) are terminated early, the hazards of treatment in the “experimental” limb may be systematically overestimated, especially when the number of outcome events is small. The SPACE investigators concluded “that noninferiority of CAS was not proven.” The upper limits of the confidence intervals of both the intention-to-treat (ITT) and the per-protocol (PP) analyses exceeded the preset delta of 2.5% (the margin of difference between CAS and CEA outcomes thought to be clinically relevant); however, because the lower limits of both analyses crossed zero, the superiority of CEA cannot be assumed. Note the value of both ITT and PP analyses in noninferiority trials: In superiority trials, the ITT analysis tends to be conservative, underestimating differences in the treatment effect and protecting against type 1 errors. In noninferiority trials, underestimates of the treatment effect can bias toward noninferiority, inflating the false positive effect, hence the additive value of the PP analysis. SPACE failed to prove the noninferiority of CAS due to lack of statistical power and insufficient numbers. There were no significant differences in the primary endpoint of ipsilateral stroke/death between CAS and CEA; the absolute difference was four outcome events in 600 patients per group.

LESSONS LEARNED
Apart from emerging trends with respect to lower procedural hazards following CAS than CEA in women (in the European trials), patients presenting with ocular symptoms, and patients with contralateral occlusion (explored at the roundtable discussion), the issue of age at treatment is an intriguing and consistent finding.

Octogenarians
The matter of the safety record for CAS in octogenari-
ans is the subject of ongoing debate. Higher rates of procedural hazards were noted in the CREST lead-in phase; thereafter, octogenarians were excluded from the lead-in phase, although they remained eligible for the randomized phase. A recent meta-analysis of pooled data from EVA-3S, SPACE, and ICSS demonstrated that in patients < 70 years of age, the procedural risks of stroke/death were 5.8% for CAS and 5.7% for CEA (relative risk, 1.00; confidence interval, 0.68–1.47). The excess risk of CAS appeared to be in patients 70 years or older, where the risk of CAS versus CEA was 12% versus 5.9%, (relative risk, 2.04; confidence interval, 1.48–2.82). A comparable pattern was seen in CREST; prespecified analyses demonstrated an interaction between age and treatment effect, with a crossover at approximately 70 years. It is important to appreciate that the influence of age at treatment on outcome for CAS is not a stochastic phenomenon, and outcomes can be plotted against age as a continuous variable, providing a predictive model that can be used to help direct decision making.

There have been recent reports, particularly from experienced operators working with proximal protection devices, suggesting very acceptable outcomes in this population. It is not clear whether baseline experience of the operator, type of EPD, or a combination of both was ultimately responsible for improved results. A point worth considering is the influence of the EPD on microembolic burden, which may vary from one type of device (proximal vs distal filter, etc.) to another. The elderly, with reduced cerebrovascular reserve, may fare better with devices that not only entrap macroemboli but also reduce the procedural microembolic burden.

CONCLUSION

Unfortunately, there seem to be irreconcilable differences between the current randomized trials, and it remains an inconvenient truth that the differences are perhaps most marked between the American and European trials. The intertrial variability for the European trials appears superficial, and the trials are similar enough to have allowed a (preplanned) meta-analysis of their datasets, which has recently been published. There are some common themes, to include lower procedural hazards in patients with ocular rather than cortical symptoms, in females (certainly in the European trials), and in those with contralateral occlusion. These factors will require further exploration as they were not a priori endpoints (and the trials were not powered to definitively answer questions about these subsets).

The impact of age on outcome is unmistakable but there is an increasing body of evidence to suggest that octogenarians can be treated safely by CAS in experienced centers employing modified CAS technique. ■

Sumaira Macdonald, MBChB (Comm.), FRCP, FRCR, PhD, of the Steering and Technical Management Committees of the Asymptomatic Carotid Surgery Trial (ACST-2), is a Consultant Vascular Radiologist and Honorary Clinical Senior Lecturer at Freeman Hospital in Newcastle upon Tyne, England. She was a collaborator and proctor (physician instructor) in the International Carotid Stenting Study (ICSS). She has disclosed that she receives grant/research funding from Abbott Vascular, Cordis Corporation, Bard Peripheral Vascular, ev3 Inc., Medtronic Invatec, Pyramed, and W. L. Gore & Associates, and that she is a paid consultant to Abbott Vascular, ev3 Inc., Medtronic Invatec, and W. L. Gore & Associates. Dr. Macdonald may be reached at sumaira.macdonald@nuth.nhs.uk.