Objectives

We sought to validate 4 angiographic measures as potential surrogates for clinical restenosis (target lesion revascularization [TLR]) after stent implantation.

Background

Given the low revascularization rates with drug-eluting stents (DES), an angiographic surrogate of TLR is desirable to reduce the sample size required to demonstrate efficacy in future trials of antirestenosis devices.

Methods

We evaluated 4 potential angiographic measures (late loss [LL] and percent diameter stenosis [%DS], both in-stent and in-segment) as a surrogate for TLR at 1 year. From 11 multicenter, prospective randomized stent trials, 9 comparing DES with bare-metal stents (BMS) and 2 comparing different DES, individual data on 5,381 patients with a single treated lesion and follow-up angiography at 6 to 9 months were analyzed.

Results

By 4 well-defined criteria of surrogacy, LL and %DS strongly predicted the risk of TLR, with in-segment %DS being the most highly predictive (p < 0.05). Differences in TLR risk were fully explained statistically by their differences in LL or %DS, although LL as a surrogate was dependent on vessel size whereas %DS was not. However, because of the curvilinearity of the logistic model, trials comparing 2 effective DES can have significant differences in mean LL and %DS but small expected differences in TLR risk, especially at the lower ranges of LL and %DS.

Conclusions

From in-stent and in-segment LL and %DS measures, logistic models can reliably estimate TLR rates for DES and BMS. These angiographic measures are thus suitable surrogate markers for clinical stent efficacy and can be used as primary end points in future DES trials to significantly reduce sample size. (J Am Coll Cardiol 2008; 51:23–32) © 2008 by the American College of Cardiology Foundation

Drug-eluting stents (DES) reduce angiographic restenosis in patients undergoing percutaneous coronary intervention (PCI). The most commonly measured clinical indicator of stent efficacy is target lesion revascularization (TLR), which is defined as recurrent ischemia due to angiographic restenosis within the stent or its margins necessitating repeat revascularization with either PCI or coronary artery bypass graft surgery. Of note, TLR rates are typically 30% to 60% lower than the corresponding binary restenosis rates, suggesting discordance between the ischemic thresholds of angiographic and clinical measures of restenosis (1,2).

See page 33
relevant 30% reduction in TLR with DES. Moreover, given the low frequency clinical event rates with DES, very large potential differences between stents (the “delta”) are often allowed in non-inferiority trials to make comparative DES studies practical, degrading confidence that the clinical performance of 2 devices are indeed similar.

To reduce the sample size required in superiority and non-inferiority trials, continuous angiographic indexes of long-term stent patency such as late lumen loss (LL) and follow-up percent diameter stenosis (%DS) have been proposed as surrogates of TLR for use in randomized studies (3–5). However, previous DES studies have suggested that the distribution of individual LL measures is asymmetric with a rightward skew and that the relation between LL and TLR is nonlinear (6,7). Whether angiographic measures are thus valid surrogates for TLR and, if so, what cutoff values correspond with clinical efficacy has not been established. These issues have important clinical and regulatory implications.

To address these issues, we systematically analyzed pooled patient-level data from 11 randomized, controlled DES trials. Specifically, we sought to determine the appropriateness of using the continuous angiographic end points of LL and follow-up %DS as surrogates of TLR after stent implantation; to characterize the relationship between clinical and angiographic measures of stent efficacy; and to assess the relative value in this regard of in-stent versus in-segment measures.

### Methods

#### Study population and protocols

Eleven randomized stent trials enrolling 8,726 patients were included in the analysis. Details of the study protocols have been reported previously (8–18). Protocol-specified angiographic follow-up was performed in a cohort of patients of varying size from each of these studies to further characterize vascular responses. In total, routine angiographic follow-up between 6 and 9 months and clinical follow-up at 1 year was performed in 5,381 patients with a single treated lesion, who comprise the study population (Table 1). All patients at baseline had symptoms or objective signs of myocardial ischemia due to coronary artery disease and at least one de novo stenosis in a native coronary artery treated with 1 or more study stents. Nine of the 11 randomized trials compared DES versus otherwise-identical BMS (8–16), and the remaining 2 trials compared 1 DES to another DES (Cypher vs. Taxus in REALITY; Prospective Randomized Multi-Center Head-to-Head Comparison of the Sirolimus-Eluting Stent [Cypher] and the Paclitaxel-Eluting Stent [Taxus] and Endeavor vs. Cypher in ENDEAVOR III; Randomized Comparison of Zotarolimus-Eluting and Sirolimus-Eluting Stents in Patients With Coronary Artery Disease) (17,18). In total, the pooled analysis comprised 6 arms treated with sirolimus-eluting stents (Cypher, Cordis Corp., Miami, Florida) (8–10,13,17,18), 5 arms using paclitaxel-eluting stents (4 with the polymer-based Taxus stent [Boston Scientific, Natick, Massachusetts]) (11,14,15,17), and 1 with a nonpolymeric paclitaxel-eluting stent (Achieve, Guidant, Indianapolis, Indiana) (12); 2 arms using zotarolimus-eluting stents (Endeavor, Medtronic, Minneapolis, Minnesota) (16,18), and 9 arms using bare metal stents (BX Velocity, Cordis Corp.; Express, Boston Scientific; Multilink Penta, Guidant, Indianapolis, Indiana; Driver, Medtronic) (8–16).

### Table 1 Characteristics of Studies Included in Pooled Analysis

<table>
<thead>
<tr>
<th>Trial Name</th>
<th>Arm 1</th>
<th>Arm 2</th>
<th>Study Design</th>
<th>No. of Patients</th>
<th>Recruiting Centers</th>
<th>Diabetes Mellitus, %</th>
<th>Reference Vessel Diameter, mm</th>
<th>Lesion Length, mm</th>
</tr>
</thead>
<tbody>
<tr>
<td>TAXUS-IV</td>
<td>Taxus (SR) Express*</td>
<td>R, DB, MC</td>
<td>558</td>
<td>U.S.</td>
<td>24.4</td>
<td>2.78 ± 0.48</td>
<td>13.8 ± 6.6</td>
<td></td>
</tr>
<tr>
<td>TAXUS-V</td>
<td>Taxus (SR) Express*</td>
<td>R, DB, MC</td>
<td>989</td>
<td>Europe</td>
<td>18.9</td>
<td>2.80 ± 0.47</td>
<td>20.6 ± 7.5</td>
<td></td>
</tr>
<tr>
<td>TAXUS VI</td>
<td>Taxus (MR) Express*</td>
<td>R, DB, MC</td>
<td>417</td>
<td>Europe</td>
<td>26.6</td>
<td>2.61 ± 0.46</td>
<td>14.6 ± 5.9</td>
<td></td>
</tr>
<tr>
<td>SIRIUS</td>
<td>Cypher BX Velocity*</td>
<td>R, DB, MC</td>
<td>699</td>
<td>U.S.</td>
<td>23.3</td>
<td>2.75 ± 0.36</td>
<td>15.2 ± 6.2</td>
<td></td>
</tr>
<tr>
<td>E-SIRIUS</td>
<td>Cypher BX Velocity*</td>
<td>R, DB, MC</td>
<td>319</td>
<td>Europe</td>
<td>26.1</td>
<td>2.64 ± 0.33</td>
<td>13.2 ± 5.7</td>
<td></td>
</tr>
<tr>
<td>C-SIRIUS</td>
<td>Cypher BX Velocity*</td>
<td>R, DB, MC</td>
<td>88</td>
<td>Canada</td>
<td>18.9</td>
<td>2.65 ± 0.45</td>
<td>9.7 ± 3.3</td>
<td></td>
</tr>
<tr>
<td>RAVEL</td>
<td>Cypher BX Velocity*</td>
<td>R, DB, MC</td>
<td>218</td>
<td>Europe</td>
<td>24.9</td>
<td>2.96 ± 0.51</td>
<td>11.5 ± 4.6</td>
<td></td>
</tr>
<tr>
<td>DELIVER</td>
<td>Achieve ML Penta*</td>
<td>R, SB, MC</td>
<td>442</td>
<td>U.S.</td>
<td>26.1</td>
<td>2.76 ± 0.47</td>
<td>13.8 ± 5.3</td>
<td></td>
</tr>
<tr>
<td>REALITY II</td>
<td>Cypher Taxus (SR)</td>
<td>R, DB, MC</td>
<td>756</td>
<td>Europe + LA + Asia</td>
<td>20.2</td>
<td>2.78 ± 0.46</td>
<td>15.1 ± 6.7</td>
<td></td>
</tr>
<tr>
<td>ENDEAVOR II</td>
<td>Endeavor Driver*</td>
<td>R, DB, MC</td>
<td>521</td>
<td>Europe + AP + Aus + Canada</td>
<td>28.6</td>
<td>2.78 ± 0.46</td>
<td>15.1 ± 6.7</td>
<td></td>
</tr>
</tbody>
</table>

* Bare-metal stent. All other stents are drug-eluting stents. † Only patients with a single lesion treated and follow-up angiography are included in this analysis.

AP = Asia Pacific; Aus = Australia; DB = double-blind; DM = diabetes mellitus; ENDEAVOR = Randomized Comparison of Zotarolimus-Eluting and Sirolimus-Eluting Stents in Patients With Coronary Artery Disease; LA = Latin America; MC = multicenter; MR = moderate release; R = randomized; RAVEL = A Randomized Comparison of a Sirolimus-Eluting Stent With a Standard Stent for Coronary Revascularization; REALITY = Prospective Randomized Multi-Center Head-to-Head Comparison of the Sirolimus-Eluting Stent (Cypher) and the Paclitaxel-Eluting Stent (Taxus); SB = single-blind; SIRIUS = Sirolimus-Eluting Stent in Coronary Lesions; SR = slow release; TAXUS = Paclitaxel-Eluting Stents in the Treatment of Longer Lesions. Focus on Patients With Diabetes; UB = unblinded.
A wide range of de novo lesions in native coronary arteries were evaluated in these trials. Earlier pivotal studies enrolled patients with relatively simple lesions (RAVEL [A Randomized Comparison of a Sirolimus-Eluting Stent With a Standard Stent for Coronary Revascularization], SIRIUS [Sirolimus-Eluting Stent in Coronary Lesions], TAXUS-IV [Paclitaxel-Eluting Stents in the Treatment of Longer Lesions. Focus on Patients With Diabetes], DELIVER, ENDEAVOR II and III) (8,9,11,12,16,18). Later studies enrolled patients with more complex lesions potentially having greater risk of restenosis, including E-SIRIUS, C-SIRIUS, TAXUS-V, and REALITY, which targeted lesions in small vessel (<3.0 mm in diameter) (10,13,14,17), and TAXUS-V, TAXUS-VI, and REALITY, which enrolled patients with long lesions (>30 mm in length) (14,15,17). Inclusion of bifurcations and ostial lesions was also allowed in the REALITY trial (17). All other studies excluded complex stenoses such as ostial, bifurcation, excessive calcification, total occlusions, and thrombus-containing lesions. In TAXUS-V, TAXUS-VI, all the SIRIUS trials, and REALITY trial, the use of multiple stents was allowed (9,10,13–15,17). In most studies, conventional stent implantation with predilation was mandated (8,9,11,12,14–16,18). Direct stenting was allowed only in E-SIRIUS, C-SIRIUS, and REALITY (10,13,17). Only the REALITY trial allowed the enrolment of patients with multiple lesions (17). From the REALITY database, only patients undergoing stent implantation in a single lesion were included in the present analysis.

Target lesion revascularization was defined as the need for repeat revascularization at the site of stent implantation, including the 5 mm proximal and distal persistent zones, with either associated ischemia or a severe follow-up %DS (>70% by quantitative angiography). Independent clinical event committees for each trial adjudicated all TLR events.

**Angiographic analysis methodology.** All angiograms were analyzed by independent core laboratories: the Brigham and Women’s Hospital, Boston, Massachusetts (9–11,13–16,18), Cardiovascular Research Foundation, New York, New York (12), and Cardialysis, Rotterdam, the Netherlands (8,17). Similar methodology was used in all 3 laboratories. Quantitative coronary angiography was performed with the CMS Medis system (Leiden, the Netherlands) in all studies but 2, in which the CASS system (PIE Medical, Maastricht, the Netherlands) was used (8,17). Measurements of minimal lumen diameter (MLD) and reference vessel diameter (RVD) were performed at baseline, after final intervention and at follow-up, and used to calculate the %DS = (1 – MLD/RVD) × 100. In 9 trials, the RVD was obtained from averaging 5-mm segments proximal and distal to the target lesion location, whereas the interpolated RVD at the lesion site was used to calculate %DS in 2 trials (8,17). We calculated LL as the change in MLD from the final post-PCI angiogram to follow-up, and it was calculated both in-stent and in the entire analysis segment, also including the 5-mm proximal and distal stent margins (also called in-segment).

**Statistical methods.** All statistical analyses were conducted using SAS version 9.1 (SAS Institute, Cary, North Carolina). The statistical criteria for evaluating whether LL and follow-up %DS are useful surrogate end points for TLR entailed 4 main steps, as follows.

**DETERMINATION OF WHETHER THE POTENTIAL SURROGATE EXHIBITS STRONG CONSISTENT EVIDENCE OF TREATMENT DIFFERENCES WITHIN EACH TRIAL.** For this purpose, the z-score was used, defined as the observed treatment difference divided by its standard error. For LL and %DS, this was the t statistic obtained from a 2-sample t test modified to permit unequal variances in the 2 treatment groups. For comparative purposes, the z-score also was obtained for the treatment difference in percentage with TLR. For such a difference in percentages, z is the square-root of the chi-square statistic. The larger the value of z, the stronger the evidence of a treatment difference (e.g., z scores of 1.96, 3.29, and 6.11 are associated with p values of 0.05, 0.001, and 0.000000001, respectively).

**EXAMINATION OF THE STRENGTH OF THE RELATIONSHIP BETWEEN THE POTENTIAL SURROGATE (LL OR %DS) AND THE CLINICAL OUTCOME (TLR).** For the patients in each clinical trial, the c-statistic was used to measure the strength of association between each quantitative outcome (e.g., in-stent LL) and the binary outcome TLR. The c-statistic is defined as the area under the receiver operator characteristic curve but can be more clearly understood as follows: for any 2 randomly selected patients, one with and one without TLR, c is the probability that the former has the greater value of the quantitative surrogate. A value of c = 1 means perfect discrimination, so an effective surrogate has a value of c close to 1.

The relationship of TLR to each of the 4 potential surrogates LL and %DS, both in-stent and in-segment is modeled using logistic regression applied to all 5,381 patients in the 11 trials. For a given LL the log odds of TLR depends on the RVD, and this necessitates a bivariate logistic regression in which the log odds of TLR are linearly related to LL and to RVD grouped in 3 intervals, <2.5 mm, 2.5 to 3 mm, and ≥3 mm. These logistic models (1 for each of the 4 potential surrogates) are then used to predict each individual patient’s probability of TLR. For each treatment group in each trial, the predicted number with TLR equals the sum of these individual probabilities. Actual and predicted percentages of patients with TLR for all 22 treatment groups in the 11 trials are then compared to examine the extent to which the potential surrogate can reliably predict the TLR rate in these trials, and hence can be relied on to predict the true expected TLR rate in any future trials in which TLR might not be actually assessed.
Determinations of Whether the Potential Surrogate in Each Clinical Trial Statistically Explains the Observed Treatment Difference in the Clinical Outcome. The Prentice criterion of surrogacy (19) entails fitting 2 logistic regression models to all the data within a trial comparing DES and BMS: 1) log odds TLR = α1 − β1 T, where T = 1 if DES, 0 if BMS; and 2) log odds TLR = α2 − β2 T − γ × potential surrogate (e.g., in stent LL). Then, the estimated percentage of treatment effect explained by the surrogate = 100 × (β1 − β2)/β1 %. With a perfect surrogate, the true amount explained should be 100%. The Prentice criterion of surrogacy was explored in the 5 largest trials comparing DES and BMS, as this criterion is best explored in large trials in which treatment differences in the clinical outcome TLR are very pronounced.

Determination of Whether the Extent of the Size of the Treatment Effect on TLR Links Closely to the Size of the Treatment Effect on the Potential Surrogate (e.g., In-Segment LL) Across All the Trial Studies. The Hughes criterion (20) of surrogacy was used to demonstrate across trials the extent to which the magnitude of treatment effect on TLR is closely linked to the magnitude of mean treatment difference in the quantitative potential surrogates LL and %DS. This is best examined graphically by plotting on a single scatter graph the %TLR on the vertical axis and the mean of the potential surrogate on the horizontal axis for each treatment group in each of the 11 trials.

Results

The included studies and characteristics of patients with a single treated lesion and angiographic follow-up are shown in Table 1. Table 2 provides the LL, follow-up %DS, and 1-year TLR rates for each of the 11 trials.

Treatment difference in %DS, LL, and TLR in each trial. For each of the 11 trials, the strength of evidence for a treatment difference in LL and %DS (both in-stent and -segment) expressed by the z-score (the observed difference in means divided by its standard error) is summarized in Table 3. As shown in this table, the z-scores for LL and follow-up %DS are similar (and markedely greater than the z-scores for TLR), signifying that LL and %DS discriminate equally well between treatment groups and much more significantly so than does TLR. For both LL and %DS, the z-scores were greater for the in-stent compared with the in-segment measure.

The relationship of LL and follow-up %DS to the clinical outcome TLR. The c-statistic expresses the strength of association of TLR to each of the 4 potential surrogate end points (Table 4) and was examined for each study comparing DES and BMS, as this criterion is best explored in large trials in which treatment differences in the clinical outcome TLR are very pronounced.

Table 2

<table>
<thead>
<tr>
<th>Trial Name</th>
<th>In-Stent</th>
<th>In-Segment</th>
</tr>
</thead>
<tbody>
<tr>
<td>TAXUS-IV</td>
<td>0.39 ± 0.50</td>
<td>0.92 ± 0.58</td>
</tr>
<tr>
<td>TAXUS-V</td>
<td>0.49 ± 0.61</td>
<td>0.90 ± 0.62</td>
</tr>
<tr>
<td>TAXUS-VI</td>
<td>0.39 ± 0.56</td>
<td>0.99 ± 0.59</td>
</tr>
<tr>
<td>SIRIUS</td>
<td>0.17 ± 0.45</td>
<td>1.00 ± 0.70</td>
</tr>
<tr>
<td>E-SIRIUS</td>
<td>0.21 ± 0.41</td>
<td>1.06 ± 0.61</td>
</tr>
<tr>
<td>C-SIRIUS</td>
<td>0.08 ± 0.33</td>
<td>1.02 ± 0.69</td>
</tr>
<tr>
<td>RAVEL</td>
<td>-0.01 ± 0.33</td>
<td>0.80 ± 0.53</td>
</tr>
<tr>
<td>DELIVER</td>
<td>0.81 ± 0.60</td>
<td>0.98 ± 0.57</td>
</tr>
<tr>
<td>REALITY</td>
<td>0.09 ± 0.38</td>
<td>0.30 ± 0.42</td>
</tr>
<tr>
<td>ENDEAVOR II</td>
<td>0.61 ± 0.46</td>
<td>1.03 ± 0.59</td>
</tr>
<tr>
<td>ENDEAVOR III</td>
<td>0.60 ± 0.48</td>
<td>0.15 ± 0.35</td>
</tr>
</tbody>
</table>

Table 3

<table>
<thead>
<tr>
<th>Trial</th>
<th>LL</th>
<th>%DS</th>
<th>TLR</th>
</tr>
</thead>
<tbody>
<tr>
<td>TAXUS-IV</td>
<td>11.6</td>
<td>12.5</td>
<td>8.9</td>
</tr>
<tr>
<td>TAXUS-V</td>
<td>10.5</td>
<td>10.0</td>
<td>7.3</td>
</tr>
<tr>
<td>TAXUS-VI</td>
<td>10.6</td>
<td>10.5</td>
<td>7.2</td>
</tr>
<tr>
<td>SIRIUS</td>
<td>18.7</td>
<td>18.5</td>
<td>13.1</td>
</tr>
<tr>
<td>E-SIRIUS</td>
<td>14.5</td>
<td>14.0</td>
<td>11.1</td>
</tr>
<tr>
<td>C-SIRIUS</td>
<td>8.2</td>
<td>8.1</td>
<td>5.8</td>
</tr>
<tr>
<td>RAVEL</td>
<td>13.7</td>
<td>11.9</td>
<td>9.8</td>
</tr>
<tr>
<td>DELIVER</td>
<td>3.0</td>
<td>2.7</td>
<td>2.3</td>
</tr>
<tr>
<td>REALITY</td>
<td>7.4</td>
<td>3.3</td>
<td>4.5</td>
</tr>
<tr>
<td>ENDEAVOR II</td>
<td>9.2</td>
<td>8.5</td>
<td>7.7</td>
</tr>
<tr>
<td>ENDEAVOR III</td>
<td>8.3</td>
<td>6.5</td>
<td>4.4</td>
</tr>
</tbody>
</table>

Values are expressed as % where indicated or as mean ± SD.

Abbreviations as in Tables 1 and 2.
groups combined. The highest c-statistic in most trials was for the association of the in-segment %DS to TLR, with average $c = 0.95$ across all 11 trials. The c-statistics relating TLR to the other 3 measures (in-stent LL, in-stent %DS, and in-segment LL) were slightly smaller, but similar to one another, with average $c = 0.90$ across all 11 trials.

Figure 1, which combines data from all 11 trials, shows the observed proportions with TLR for 12 intervals of LL and 10 intervals of follow-up %DS. For the latter, it is evident that for any given LL the risk of TLR depends on the RVD, whereas the relationship between follow-up %DS and TLR does not depend on vessel size. A sharp monotonic increase in the risk of TLR was present once follow-up %DS reaches 50% up to a TLR rate of 90% for patients with %DS between 70% and 90%. For the small number of patients with %DS $\geq 90\%$, the observed TLR rate decreased to 63%.

Logistic regression was used to model the relation between the 4 angiographic surrogates and TLR, adjusting for vessel size (RVD). The consequent smooth logistic curves are shown in Figure 1. The actual regression equations relating in-stent LL and in-segment %DS to TLR are given in the Online Appendix.

Table 5 shows the predicted number of patients with TLR (based on the aforementioned logistic model of in-stent LL and in-segment %DS) for each treatment group in the 11 trials compared with the actual observed numbers. Close agreement between predicted and actual numbers was present in all 44 instances, indicating that the models have excellent goodness of fit. This is further illustrated in the upper graph of Figure 2, which plots the actual TLR versus in-stent LL predicted TLR rates for all 22 treatment groups (13 DES and 9 BMS). All points are within one standard error of the 45° diagonal line of equality. Similar results were obtained when the predicted percent with TLR was based on a logistic model relating risk of TLR to in-segment follow-up %DS (Fig. 2, lower graph).

**Relationship between treatment differences in TLR and LL or %DS.** Applying the Prentice criterion of surrogacy to the 5 largest trials comparing DES and BMS, we present in Table 6 for each trial the percentage of treatment effect on TLR that is explained by each of the four potential surrogate outcomes. In-segment LL and follow-up %DS each explain most of the treatment effect in TLR, such that in each trial any residual treatment effects on TLR, after accounting for the influence of LL and %DS, fall well short of statistical significance. In-stent LL and follow-up %DS were both estimated to explain all of the treatment effect in most of the trials. This unusual phenomenon may be because in-stent measures reflect “pure efficacy” without taking account of edge effects, and hence tend to slightly exaggerate the overall superiority of DES compared with BMS.

**Relating the size of treatment effect on TLR to the size of treatment effects on LL and %DS.** Applying the Hughes criterion of surrogacy, in Figure 3 we show the reductions in mean LL and mean follow-up %DS (both in-stent and in-segment) that were achieved for each trial of a DES compared with a BMS were clearly related to a corresponding reduction in risk of TLR. Of the 4 plotted associations in Figure 3, the tightest link to TLR reduction is seen for in-segment LL. In contrast, in the REALITY trial, which compared 2 DES with relatively low LL, the rates of TLR were similar with the 2 devices despite a highly significant difference in in-stent and in-segment mean LL and follow-up %DS.

**Discussion**

Before the availability of DES, the major limitation of PCI was restenosis, which necessitated repeat intervention in 20% to 30% of cases after BMS implantation (21). In contrast, currently approved DES are notable for much lower rates of clinical and angiographic restenosis in the mostly noncomplex lesion types enrolled in the pivotal randomized trials and, as a result, DES have been widely adopted for the majority of patients undergoing PCI. This new standard of care has made it difficult to realistically evaluate next-generation DES and novel antirestenosis therapies because the frequencies of the customary measures of clinical effectiveness (such as TLR) are so low that very large sample sizes are required for superiority and noninferiority testing. Although this problem may be overcome by restricting enrolment to high-risk lesions and patients, recruitment rates would markedly slow, and the results would not be applicable to a broader cross section of patients. Thus, the need exists for alternative measures that can reliably be used as surrogates for TLR, allowing smaller trial sizes to accelerate evaluation of potentially more efficacious and cost-effective DES alternatives.

Ellis et al. (6) previously demonstrated a strong association between the continuous angiographic measure of
LL and individual risk of TLR. However, this analysis was limited by modest sample size from a single study (TAXUS-IV). The present report describing the detailed results of a pooled patient level analysis from 11 randomized contemporary DES trials involving 5,381 patients with systematic angiographic follow-up is the most extensive investigation to date evaluating the potential utility of angiographic surrogates of TLR. This analysis demonstrates that by 4 different commonly accepted criteria of surrogacy, both LL and follow-up %DS, whether measured in-stent or in-segment, are valid surrogates for TLR. The risk of TLR in individual patients is strongly related to LL and %DS. Also, the treatment differences in TLR rates within each trial were entirely explained by treatment differences in LL and %DS. This strong of a linkage has rarely been observed in other fields; for example, blood pressure differences between antihypertensive regimens do not fully explain their differences in risk of stroke (22). Finally, the size of the treatment differences in %TLR linked closely to the respective size of the mean differences for these surrogate markers, especially for in-segment LL (20).

Figure 1 Impact of Angiographic Measures on TLR

Proportion of patients with target lesion revascularization (TLR) as a function of in-stent and in-segment late loss (LL) and follow-up percent diameter stenosis (%DS), stratified by reference vessel diameter (RVD) using individual patient data from all 11 trials. **Left column** = observed proportions with TLR for 12 intervals of LL and 10 intervals of follow-up %DS. **Right column** = corresponding smoothed logistic curves.
LL as a surrogate of clinical stent efficacy. The relationship between individual patient LL and risk of TLR was found to be well represented by a logistic curve, confirming the observation introduced by Ellis and co-workers (6). Although LL would thus appear to be a useful surrogate marker for TLR in DES versus BMS trials, the nonlinearity of this relationship makes the use of LL as a surrogate in comparative DES versus DES studies more problematic. In REALITY, despite the highly significant difference in both in-stent and in-segment late loss between Cypher and Taxus, the 1-year rates of TLR were similar: 4.4% versus 5.3%, respectively. Conversely, the large difference in LL between the Cypher and Endeavor stents in the ENDEAVOR III trial was associated with a larger relative difference in TLR between the two stents (3.6% with Cypher and 6.9% with Endeavor). This disparity between REALITY and ENDEAVOR III may be explained by the fact that the greater LL of the Endeavor stent places it on the rising slope of the LL-TLR relation, whereas the mean LL of Cypher and Taxus place the majority of patients on the flatter part of the curve where similar absolute differences in LL translate into smaller differences in TLR.
Follow-up %DS as a surrogate of clinical stent efficacy. Follow-up %DS was equally effective as LL in predicting TLR. The risk of TLR was small and nearly flat at less than 50% DS, increased sharply between %DS of 50% to 80%, and then reached a plateau at greater than 80% DS. Of note, the TLR rate decreased to approximately 63% for the small number of occluded (or near-occluded) vessels, which likely is explained by physician decision to treat approximately one-third of such patients conservatively if clinically stable or with a low likelihood of sustained vessel patency with repeat revascularization.

The present study demonstrates that %DS and LL are roughly equally effective surrogates. However, although it has been argued that LL best captures the physiological mechanism whereby stents reduce TLR (3–5), several advantages favor the use of %DS in practice. First, and perhaps most importantly, the impact of LL on the likelihood of TLR varies with vessel size, whereas the %DS–TLR relationship is vessel size independent. Second, as the difference in 2 measures of MLD obtained from 2 different angiograms at different points in time, LL is inherently subject to more measurement error than follow-up %DS. In contrast, %DS is based on evaluation at a single time point. Finally, %DS is conceptually more intuitive and easier to apply in clinical practice than LL.

In-stent versus in-segment measures as angiographic surrogates. In-stent LL and follow-up %DS measurements both assess the magnitude of absolute and relative neointimal hyperplasia within the stent, providing an accurate assessment of the antiproliferative effect of DES. However, in-stent measures reflect only the pure biologic potency of an antirestenotic device. In-segment measurements additionally account for the magnitude of lumen renarrowing that occurs at the margins of the stent, which may reflect stent/balloon mismatch, drug diffusion effects, and so on. Because isolated stenoses at stent edges represent an increasingly greater proportion of TLR events with DES than BMS (9,11), in-segment measures might be a wise choice as a clinical event surrogate.

Use of quantitative angiographic measures to reduce clinical trial sample size requirements. Both %DS and LL had significantly greater ability to discriminate between treatments as compared with relying on the binary outcome TLR. For instance, in each of the 5 largest trials comparing DES and BMS, the $z$-score for TLR was slightly more than half of the $z$-score for in-segment LL, with a mean ratio of 0.54. As a result, markedly fewer patients would be required in a clinical trial to demonstrate efficacy using any of the quantitative angiographic measures as primary end point rather than binary TLR. Specifically, the number of patients required to detect a statistically significant treatment difference is inversely proportional to the square of the expected $z$-score. Hence, future trials comparing stents would require approximately 71% fewer patients using in-segment LL rather than TLR as the efficacy measure (1 – $0.54^2 \times 100 = 71\%$). As discussed previously, however, for comparisons of 2 DES, both with relatively low LL, the detection of
significant differences in LL or follow-up %DS does not imply that sizeable differences in TLR exist.

It should be noted that the logistic models in the present report describe the relationship between individual angiographic measures (LL and %DS) and the clinical efficacy variable TLR. The logistic models cannot be used to predict the expected TLR rate of a specific stent cohort by simply using the observed group mean LL or %DS, which ignores the width and skewness of the distribution. The predictive logistic models in Figure 1 and the Online Appendix may be directly applied, however, to any individual patient’s angiographic data to provide an estimated probability of TLR. To determine the predicted %TLR for an entire study group, the probability of TLR for each patient is then averaged across the entire study cohort to determine the expected group TLR rate. Such an approach is a more direct use of the observed trial data rather than relying on the mean LL or %DS, and avoids the need for more complex power transformations to remove skewness (3–5).

Limitations of angiographic surrogates of TLR. Follow-up angiography in large studies is always <100% complete, most commonly because asymptomatic patients refuse re-study. However, angiographic follow-up rates in the pre-specified angiographic cohorts were >80% in most of the studies in the present pooled analysis. We would thus recommend that future trials using angiographic measurements as surrogates require similarly high rates of angiographic follow-up to achieve reliable conclusions. A potential second limitation is that 3,345 of the 8,726 patients were not enrolled in the angiographic follow-up cohort of the randomized trials. However, because consecutive patients from each trial were enrolled in each angiographic substudy, it is likely that the study cohort of the present analysis is representative. Third, although blinded clinical event committees adjudicated all cases of TLR with discretion to include only events with documented ischemia, some TLR events were driven by angiographic follow-up alone in asymptomatic patients with a severe %DS (>70% by core lab assessment, which typically corresponds to a >85% operator assessed visual stenosis), with the rationale that these patients would soon likely become symptomatic and require revascularization. The exact frequency of which this “oculostenotic reflex” contributed to reported TLR rates in all the studies is unknown, but was ~10% in the TAXUS-II, -IV, -V, and -VI trials (G. Stone, unpublished data, March 1, 2007) and is thus unlikely to have materially impacted the present analysis. Nonetheless, the angiographic measures discussed herein may be considered surrogates for either ischemia-driven TLR, or in a small proportion of patients, a severe recurrent angiographic stenosis (the prevention of which is also desirable). Moreover, routine angiographic follow-up increases TLR event rates in patients receiving both BMS and DES, though to a similar relative degree (1). These biases may somewhat accentuate the magnitude of the TLR relationship with LL and %DS. TLR itself is also an imperfect measure of recurrent ischemia since the decision whether or not to perform revascularization can be affected by patient and physician preferences. In addition, LL and %DS are surrogates only of TLR after stent implantation and not of the need for future revascularization elsewhere in the coronary tree from pre-existing or progressive atherosclerosis in nonstented segments.

Finally, an important caveat when considering the use of angiographic efficacy surrogates to reduce clinical sample size in new device trials is that the ability to detect differences in relatively low frequency safety events between 2 antirestenosis therapies (for which there are no acceptable angiographic surrogate measures) will be diminished. For example, previous large randomized trials demonstrated that the achievement of greater luminal dimensions with directional atherectomy compared with balloon angioplasty was associated with an increase in periprocedural myocardial infarction (23); such a relationship might have been missed in a smaller trial powered for angiographic end points only. Additional strategies are thus needed to ensure new device safety, such as large-scale, simple randomized trials without planned follow-up angiography, or observational patient registries. In this regard, emerging devices could first be evaluated in modest-sized phase 2 studies with routine angiographic follow-up to demonstrate antirestenosis efficacy as a surrogate of low clinical TLR. Once a device has passed this first “screening” test, a larger-scale pivotal study (without follow-up angiography) could then be performed with primary safety end points. Conversely, if the angiographic surrogate end points for clinical efficacy are not met in the initial study, resources can then be diverted toward the development of more beneficial devices to improve patient outcomes.

Conclusions

Applying well-defined rigorous criteria to an extensive database of randomized trials, we have demonstrated that within certain constraints, LL and follow-up %DS are suitable surrogate markers for TLR in trials evaluating DES and BMS. This finding has important practical and regulatory implications for future trials investigating the efficacy of new DES and antirestenosis devices.

Author Disclosures

Dr. Pocock has received research grants from Boston Scientific, has received honoraria from Boston Scientific and Conor Medical, and is a consultant/on the advisory board for Boston Scientific. Dr. Lansky has received research grants from Cordis, Boston Scientific, Medtronic, Abbott, and Guidant. Dr. Mehran is on the Speakers’ Bureau for The Medicines Company and has received honoraria from The Medicines Company. Dr. Popma has received research grants from Cordis, Medtronic, and Boston Scientific; is on the Speakers’ Bureau for Cordis, Boston Scientific, and Medtronic; and is a consultant/on the advisory board for
Cordis, Boston Scientific, and Medtronic. Dr. Dangas has received research grants from Cordis. Dr. Moses has received honoraria and is a consultant to and on the advisory board for Cordis. Dr. Kandzari has received research grants from Cordis, Boston Scientific, and Medtronic; is on the Speakers’ Bureau for Medtronic; is a consultant on the advisory board for Cordis, Boston Scientific, and Medtronic; and is a full-time employee of Cordis Corp. Dr. Ellis has received research grants from Cordis and is a consultant to and on the advisory board for Boston Scientific, Medtronic, and Cordis Johnson & Johnson. Dr. Stone is a consultant to and has received research grants and lecture fees from Boston Scientific Corp., Abbott Vascular Devices, Reva Medical, Xtent, and St. Jude Medical; is on the board of directors for and owns equity in Devax Corp.; and is a past consultant to and has received research grants from Medtronic Corp. and Guidant Corp.

Reprint requests and correspondence: Dr. Gregg W. Stone, Columbia University Medical Center, The Cardiovascular Research Foundation, 111 East 59th Street, 11th Floor, New York, New York 10022. E-mail: gs2184@columbia.edu.

REFERENCES


APPENDIX

For the logistic regression equations relating in-stent LL and in-segment %DS to TLR, please see the online version of this article.