Patient-Important Outcomes in Registered Diabetes Trials

Gunjan Y. Gandhi, MD, MSc  
M. Hassan Murad, MD, MPH  
Akira Fujiyoshi, MD, MPH  
Rebecca J. Mullan, MS  
David N. Flynn, BS  
Mohamed B. Elamin, MBBS  
Brian A. Swiglo, MD  
William L. Isley, MD†  
Gordon H. Guyatt, MD, MSc  
Victor M. Montori, MD, MSc

TRIALS MEASURING BIOCHEMICAL and surrogate markers may help researchers understand how and to what extent interventions could affect health. The value of these interventions remains unclear until trials test their effect on outcomes that are important to patients and clinicians. We have previously reported1 that only 21% of diabetes randomized clinical trials (RCTs) published in high-impact medical journals as of 2003 reported on patient-important outcomes, that is, death and quality of life (morbidity, pain, function)2;3; in contrast, 60% measured physiological or laboratory outcomes.1 Diabetologists have engaged in limited debate regarding the paucity of trials measuring patient-important outcomes.4;6

In the wake of policy debates concerning rosiglitazone, those involved have called for trials of antidiabetes agents that measure cardiovascular end points important to patients.7 Nevertheless, authors of a recent consensus statement on how to choose antidiabetic agents to treat patients with type 2 diabetes based their recommendations on trials that measure the effect of these agents on physiological or surrogate markers of poor validity.8 In both cases, the authors looked forward to ongoing trials’ assessing patient-important outcomes.

Are ongoing diabetes trials likely to be more informative to patients and clinicians? We used a novel approach of querying large public clinical trial registries to systematically determine the extent to which ongoing and future registered RCTs plan to measure patient-important outcomes in patients with diabetes.

Context Concerns about the safety and efficacy of diabetes interventions persist, in part because randomized clinical trials (RCTs) have not measured their effect on patient-important outcomes, ie, death and quality of life (morbidity, pain, function).

Objective To systematically determine the extent to which ongoing and future RCTs in diabetes will ascertain patient-important outcomes.

Data Sources On November 10, 2007, we searched primary RCT registries ClinicalTrials.gov (http://www.clinicaltrials.gov), International Standard Randomized Controlled Trial Number Register (http://isrctn.org), and Australian New Zealand Clinical Trials Registry (http://www.anzctr.org.au).

Study Selection We identified phase 2 through 4 RCTs enrolling patients with diabetes. Of 2019 RCTs, 1054 proved eligible. We randomly sampled 50% of the eligible RCTs (527 of 1054) and selected 436 registered since registration became mandatory (2004).

Data Extraction Pairs of reviewers working independently collected study characteristics and determined the outcomes measured and their type (physiological outcomes, surrogate outcomes thought to reflect an increased risk for patient-important outcomes, and patient-important outcomes).

Results Of the 436 registered RCTs included in this analysis, 24 (6%) had not started enrollment, 109 (25%) were actively enrolling, and 303 (69%) had completed enrollment. Primary outcomes were patient-important outcomes in only 78 of 436 RCTs (18%; 95% confidence interval [CI], 14%-22%), physiological and laboratory outcomes in 69 of 436 (16%; 95% CI, 13%-20%), and surrogate outcomes in 268 of 436 (61%; 95% CI, 57%-66%). Patient-important outcomes were reported as primary or secondary outcomes in 201 of 436 (46%; 95% CI, 41%-51%). In multivariate analysis, large trials (odds ratio [OR], 1.10; 95% CI, 1.02-1.19 for every additional 100 patients) and trials of longer duration (OR, 1.03; 95% CI, 1.01-1.06 for every additional 30 days) were more likely while parallel design RCTs (OR, 0.15; 95% CI, 0.09-0.44) and type 2 diabetes trials (OR, 0.23; 95% CI, 0.09-0.61) were less likely to assess patient-important outcomes as a primary outcome.

Conclusion In this sample of registered ongoing RCTs in diabetes, only 18% included patient-important outcomes as primary outcomes.

©2008 American Medical Association. All rights reserved.  

Author Affiliations: Division of Endocrinology, Diabetes, Metabolism and Nutrition (Drs Gandhi, Swiglo, Isley, and Montori), Knowledge and Encounter Research Unit (Drs Gandhi, Murad, Fujiyoshi, Swiglo, Elamin, Isley, and Montori, Ms Mullan, and Mr Flynn), and Division of Preventive, Occupational, and Aerospace Medicine (Dr Murad), College of Medicine, Mayo Clinic, Rochester, Minnesota, and Department of Clinical Epidemiology and Biostatistics, McMaster University, Hamilton, Ontario, Canada (Dr Guyatt).

†Died September 27, 2007.

Corresponding Author: Victor M. Montori, MD, MSc, Knowledge and Encounter Research Unit, Mayo Clinic College of Medicine, 200 First St SW, Rochester, MN 55905 (montori.victor@mayo.edu).
Eligible registered RCTs exclusively studied patients with any form of diabetes receiving at least 2 interventions, 1 of which served as control. When trials were registered more than once, we retained the initial registration. We included registered RCTs regardless of their language, size, or their primary objectives. We did not exclude trials that were completed or closed for enrollment.

Search Strategy
On November 10, 2007, we searched for phase 2 to 4 RCTs enrolling patients with diabetes registered in 3 large public-access trial registries: ClinicalTrials.gov (http://www.clinicaltrials.gov), International Standard Randomized Controlled Trial Number Register (http://isrctn.org), and Australian New Zealand Clinical Trials Registry (http://www.anzctr.org.au). We chose to use these 3 registries because they meet the proposed WHO criteria for a primary register.

Assessment of Study Eligibility
Teams of 2 reviewers working independently screened all potentially eligible studies identified in the initial search. When a disagreement was noted, the first step in resolving it was by consensus (ie, the 2 reviewers discussed among themselves to reach a decision). If they did not agree, then we used arbitration (ie, a third reviewer who is an experienced diabetologist and methodologist made the decision). In this process, we often had to Google the trial to find if it had an official Web site that offered more details than what is in registries. The majority of the disagreements were resolved by consensus among the 3 reviewers. There were only 4 instances requiring arbitration. Eligible studies included RCTs enrolling patients with diabetes. They had to report outcomes of interest in the registries.

Because we planned to test up to 10 predictors of trials that sought to measure patient-important outcomes in multivariate analyses, we estimated that we needed at least 100 trials that measured patient-important outcomes and 100 trials that did not to be able to test these models with a reasonably low risk of overfitting. We extracted data from 30 eligible clinical trials as a pilot and noted that about 25% of these chose patient-important outcomes as primary outcome measures. This meant that we needed 400 trials. We chose to focus on trials registered after registration became mandatory in 2004.

Data Collection
Working in duplicate, we used a standardized form to extract data regarding study characteristics that included registration details, trial funding, trial phase, region and number of centers, planned sample size, enrollment status, whether trials were completed or were active, design, blinding, patient characteristics (age, diabetes type), and type of intervention (drug vs placebo, active intervention, usual care, or no intervention; nutrition; education, counseling or physiological interventions, or other interventions).

Outcome Classification
Two reviewers working independently and in a blinded fashion classified trial outcomes into 3 categories.

Patient-Important Outcomes: death and quality of life (major morbid events such as stroke, myocardial infarction, amputation, loss of vision, and end-stage renal disease; minor morbid events such as hypoglycemic events, delayed wound healing, infection, and visual disturbances; and pain and functional status).

Surrogate Outcomes: intermediate end points that may indicate disease progression and increased risk for patient-important outcomes (eg, glycated hemoglobin, cholesterol levels, worsening renal function, and need for retinal photocoagulation).

Physiological and Laboratory Outcomes: assessed response to physiological or laboratory maneuvers without direct tangible effects on patients (eg, insulin, C-peptide levels).

We used the $\kappa$ statistic to measure chance-adjusted interobserver agreement about the nature of outcomes.

Statistical Analyses
We used descriptive statistics to characterize the reporting of outcomes in registered diabetes trials. Univariate analyses explored associations between trial characteristics and likelihood of reporting patient-important outcomes. To test predictors, we constructed a multivariate model (using logistic regression) with patient-important outcomes (yes/no) as the dependent variable and the following planned independent predictors: study sponsor (for profit vs not for profit), trial design (parallel vs crossover), trial phase (3 or 4 vs 2), number of patients, type of diabetes (type 2 diabetes vs others). The rationale for choosing these predictors is (1) measuring patient-important outcomes, such as death or stroke, is less likely to occur in studies with crossover design; (2) phase 2 trials are probably less likely to report patient-important outcomes because they focus on dosing, safety, and feasibility of interventions; (3) smaller trials may be less
likely to report patient-important outcomes because they are not powered for these outcomes; (4) trials enrolling patients with type 2 diabetes, who are at higher risk due to comorbidities, may be more likely to assess patient-important outcomes; and (5) trials with shorter follow-up duration are expected to report less patient-important outcomes because many complications of diabetes require longer time to occur.

Odds ratios (ORs) and their associated 95% confidence intervals (CI) characterize the strength of association between predictors and outcome. This is similar to setting a significance level at 5% and conducting 2-sided hypothesis testing. When this 95% CI does not include 1 then the respective P value would be <.05 and the association statistically significant. We entered data in duplicate and conducted analyses using StatsDirect 2.5.4 (StatsDirect Ltd, Cheshire, England). Considering that trials are usually powered to detect predetermined significant differences in their primary outcomes, we considered primary outcomes in our main analysis and conducted sensitivity analysis using both primary and secondary outcomes.

RESULTS

Search Results

The FIGURE describes the flow of articles. After rigorously screening 2019 potentially eligible RCTs in duplicate, 1054 proved eligible. We randomly sampled 50% of the eligible RCTs (527 of 1054) and selected those registered since registration became mandatory in 2004 (436 trials; trial numbers available online at www.jama.com).

Classification of Trial Outcomes

The majority of registered trials were phase 3 or 4, parallel design RCTs conducted in Europe or North America involving adults with type 2 diabetes, testing a drug and registered in ClinicalTrials.gov. Of these, 6% (24 of 436) had not started enrollment, 25% (109 of 436) were actively enrolling, and 69% (303 of 436) had finished enrollment. Sixty-four (21%) of the 303 trials that finished enrollment were designated as active and not completed. A MEDLINE search for publication status was conducted on a 10% random sample (46 trials) and revealed that 2 (4%) of the 46 included trials had been published.

Primary outcomes were patient-important outcomes in only 78 of 436 RCTs (18%; 95% CI, 14%-22%), physiological and laboratory outcomes in 69 of 436 (16%; 95% CI, 13%-20%), and

<p>| Table 1. Description of 436 Included Randomized Trials and Their Primary Outcomes |
|---------------------------------|------------------------------|----------------|----------------|----------------|</p>
<table>
<thead>
<tr>
<th>Trial Characteristics</th>
<th>Patient Important</th>
<th>Surrogate</th>
<th>Laboratory</th>
<th>Other Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trials</td>
<td>78 (18)</td>
<td>268 (61)</td>
<td>69 (16)</td>
<td>21 (5)</td>
</tr>
<tr>
<td>Registry</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ClinicalTrials.gov</td>
<td>55 (71)</td>
<td>224 (84)</td>
<td>58 (84)</td>
<td>16 (76)</td>
</tr>
<tr>
<td>IsRCTN</td>
<td>15 (19)</td>
<td>37 (14)</td>
<td>7 (10)</td>
<td>4 (19)</td>
</tr>
<tr>
<td>Australian New Zealand</td>
<td>8 (10)</td>
<td>7 (2)</td>
<td>4 (6)</td>
<td>1 (5)</td>
</tr>
<tr>
<td>Year</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2004</td>
<td>3 (4)</td>
<td>15 (6)</td>
<td>0</td>
<td>1 (5)</td>
</tr>
<tr>
<td>2005</td>
<td>22 (28)</td>
<td>92 (34)</td>
<td>23 (33)</td>
<td>6 (29)</td>
</tr>
<tr>
<td>2006</td>
<td>26 (33)</td>
<td>113 (42)</td>
<td>32 (47)</td>
<td>7 (33)</td>
</tr>
<tr>
<td>2007</td>
<td>27 (35)</td>
<td>48 (18)</td>
<td>14 (20)</td>
<td>7 (33)</td>
</tr>
<tr>
<td>Phase</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>12 (15)</td>
<td>50 (19)</td>
<td>22 (32)</td>
<td>5 (24)</td>
</tr>
<tr>
<td>3</td>
<td>34 (44)</td>
<td>130 (49)</td>
<td>24 (35)</td>
<td>6 (28)</td>
</tr>
<tr>
<td>4</td>
<td>14 (18)</td>
<td>53 (20)</td>
<td>15 (22)</td>
<td>5 (24)</td>
</tr>
<tr>
<td>Not reported</td>
<td>18 (23)</td>
<td>35 (12)</td>
<td>8 (12)</td>
<td>5 (24)</td>
</tr>
<tr>
<td>Region of work</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>North America</td>
<td>31</td>
<td>134</td>
<td>31</td>
<td>10</td>
</tr>
<tr>
<td>Europe</td>
<td>29</td>
<td>112</td>
<td>29</td>
<td>9</td>
</tr>
<tr>
<td>Australia</td>
<td>9</td>
<td>33</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Asia</td>
<td>7</td>
<td>60</td>
<td>9</td>
<td>4</td>
</tr>
<tr>
<td>Africa</td>
<td>7</td>
<td>15</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>South America</td>
<td>0</td>
<td>31</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Not reported</td>
<td>4</td>
<td>24</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>No. of centers</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Single</td>
<td>19 (25)</td>
<td>57 (21)</td>
<td>25 (36)</td>
<td>3 (14)</td>
</tr>
<tr>
<td>Few (&lt;10)</td>
<td>13 (17)</td>
<td>31 (12)</td>
<td>14 (20)</td>
<td>2 (9)</td>
</tr>
<tr>
<td>Many (&gt;10)</td>
<td>22 (28)</td>
<td>92 (34)</td>
<td>8 (12)</td>
<td>7 (33)</td>
</tr>
<tr>
<td>Not reported</td>
<td>24 (30)</td>
<td>88 (33)</td>
<td>22 (32)</td>
<td>9 (44)</td>
</tr>
<tr>
<td>Diabetes type</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>10 (13)</td>
<td>35 (13)</td>
<td>16 (23)</td>
<td>4 (19)</td>
</tr>
<tr>
<td>2</td>
<td>27 (37)</td>
<td>203 (76)</td>
<td>43 (62)</td>
<td>12 (57)</td>
</tr>
<tr>
<td>1 or 2</td>
<td>22 (30)</td>
<td>19 (7)</td>
<td>6 (9)</td>
<td>3 (14)</td>
</tr>
<tr>
<td>Other (gestational, etc)</td>
<td>0</td>
<td>4 (1)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Type not specified</td>
<td>19 (20)</td>
<td>7 (3)</td>
<td>4 (6)</td>
<td>2 (10)</td>
</tr>
<tr>
<td>Age of study participants, y</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥18</td>
<td>47 (60)</td>
<td>168 (63)</td>
<td>40 (58)</td>
<td>12 (57)</td>
</tr>
<tr>
<td>&lt;18</td>
<td>25 (32)</td>
<td>83 (31)</td>
<td>24 (35)</td>
<td>8 (38)</td>
</tr>
<tr>
<td>Not reported</td>
<td>6 (8)</td>
<td>17 (6)</td>
<td>5 (7)</td>
<td>1 (5)</td>
</tr>
<tr>
<td>Study population size</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤50</td>
<td>12 (15)</td>
<td>32 (12)</td>
<td>32 (46)</td>
<td>2 (9)</td>
</tr>
<tr>
<td>51-200</td>
<td>23 (30)</td>
<td>67 (25)</td>
<td>26 (38)</td>
<td>6 (25)</td>
</tr>
<tr>
<td>&gt;200</td>
<td>40 (51)</td>
<td>159 (59)</td>
<td>10 (15)</td>
<td>10 (48)</td>
</tr>
<tr>
<td>Not reported</td>
<td>3 (4)</td>
<td>10 (4)</td>
<td>1 (1)</td>
<td>3 (14)</td>
</tr>
<tr>
<td>Trial design</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parallel</td>
<td>65 (83)</td>
<td>245 (91)</td>
<td>53 (77)</td>
<td>19 (90)</td>
</tr>
<tr>
<td>Crossover</td>
<td>12 (15)</td>
<td>17 (6)</td>
<td>14 (20)</td>
<td>2 (10)</td>
</tr>
<tr>
<td>Not reported</td>
<td>1 (2)</td>
<td>6 (3)</td>
<td>2 (3)</td>
<td>0</td>
</tr>
</tbody>
</table>
surrogate outcomes in 268 of 436 (61%; 95% CI, 57%-66%). Included trials and the classification of primary outcomes are described in Table 1. Patient-important outcomes were reported as primary or secondary outcomes in 201 of 436 (46%; 95% CI, 41%-51%). Included trials and the classification of primary and secondary outcomes are described in Table 2. Because of poor reporting, we could not classify the outcomes in 21 RCTs. Of the 201 RCTs reporting patient-important outcomes, 27 assessed mortality, 38 major morbidity, 87 minor morbidity, 33 quality of life, 25 pain, and 15 functional status, as primary or secondary end points. Chance-adjusted interreviewer agreement for outcome determination ranged from $k = 0.64$ to 0.88 between teams of reviewers.

### Predictors of Reporting Patient-Important Outcomes

The results of univariate and multivariable logistic regression analyses are summarized in Table 3 depicting 2 models; the first is based on primary outcomes and the second is based on both primary and secondary outcomes.

### Primary Outcomes

In multivariate analysis, independent predictors of patient-important outcomes were larger trial size, longer trial duration, and crossover-design RCTs. Trials of patients with type 2 diabetes were significantly less likely to report patient-important outcomes. Study sponsor and trial phase were not associated with consideration of patient-important outcomes. The exclusion of phase 2 trials from analyses changed the proportion of trials reporting patient-important outcomes as a primary outcome from 18% to 19%.

### Primary and Secondary Outcome

In multivariate analysis, the independent predictors of patient-important outcomes were trial length and trial phase (phase 3 or 4 vs phase 2). Trials of patients with type 2 diabetes were significantly less likely to report patient-important outcomes. Trial size, study sponsor, and trial design (parallel vs crossover) were not associated with consideration of patient-important outcomes. The exclusion of phase 2 trials from analysis changed the proportion of trials reporting patient-important outcomes as their primary or secondary outcome from 46% to 52%.

### Limitations and Strengths

The inferences from this review are limited by inherent limitations of trial registries. Some trials did not report all the variables assessed in this review. Sample size of trials, if reported, did not contain information about how it was determined and which of the outcomes drove the sample-size calculations. Furthermore, registries did not indicate whether a trial was a pilot study, in which case surrogate end points may be justified. There was significant variability in the reporting of details of trial design as well as outcomes within and between the 3 trial registries. In addition, the registries have limited advanced search capabilities compared with MEDLINE or other electronic databases of published literature. Another key area that was lacking is the ability to allow Web site links in critical fields. Researcher-controlled Web sites may be changed at any time, therefore potentially affecting the integrity of the record allowing for information deficits or unacknowledged changes after the study starts enrolling. These factors may have had an effect on the completeness of assessment of trial description and on the reproducibility of our outcomes classification.

There are few trials reporting physiological and laboratory outcomes represented in our findings compared with our previous report based on published diabetes RCTs: while 119/199

---

Table 1. Description of 436 Included Randomized Trials and Their Primary Outcomes (cont)

| Trial Outcomes, No. (%) | Patient Important | Surrogate | Laboratory | Other Outcomes$^a$
|------------------------|-------------------|-----------|------------|----------------
| **Type of intervention** |                   |           |            |                |
| Drug vs Placebo        | 29 (37)           | 103 (38)  | 40 (58)    | 6 (29)         |
| Active drug            | 23 (29)           | 98 (37)   | 19 (28)    | 6 (29)         |
| Usual care             | 2 (3)             | 3 (1)     | 3 (4)      | 0              |
| Other                  | 1 (1)             | 5 (2)     | 1 (1)      | 0              |
| **Education or lifestyle changes** | 10 (13) | 35 (13) | 0 | 7 (33) |
| Nutrition              | 2 (3)             | 15 (6)    | 4 (6)      | 0              |
| Other interventions$^c$| 11 (14)           | 9 (3)     | 2 (3)      | 2 (10)         |
| **Funding**            |                   |           |            |                |
| Not-for-profit sources | 34 (44)           | 68 (25)   | 26 (38)    | 8 (38)         |
| For-profit sources     | 34 (44)           | 179 (67)  | 31 (45)    | 12 (57)        |
| **Mixed sources**      | 10 (12)           | 17 (6)    | 10 (14)    | 1 (5)          |
| Not reported           | 0                 | 4 (2)     | 2 (3)      | 0              |
| **Planned trial duration, median (range), d** | 232 (7-2555) | 183 (3-1825) | 136 (1-548) | 180 (3-1461)

*Abbreviation: ISRCTN, International Standard Randomized Controlled Trial Number.

$^a$Examples of other outcomes: cost, utilization of health services, and anthropomorphic changes.

$^c$Examples of other interventions: surgical procedures, acupuncture, and screening tests.
(60%) of published RCTs in our previous report assessed physiological and laboratory measures, we noted that only 57 of 436 (13%) registered trials in the present study measured physiological and laboratory outcomes. This suggests that RCTs’ assessing such outcomes are not getting registered, that registration bias leads to overregistration of trials measuring patient-important outcomes, or that publication bias differentially affects trials that measure patient-important outcomes. This concern is reinforced because many major journals have endorsed a policy of mandatory RCT registration prior to publication,10 the policy statement indicated the registries we used as meeting all their requirements, and trialists tend to submit trials measuring patient-important outcomes to higher-impact journals (ie, 62% of published diabetes RCTs appeared in the top general medical journals). Thus, our results may represent an overestimate of the extent to which future RCTs plan to measure patient-important outcomes and thus provide an excessively optimistic forecast.

The validity of the findings is strengthened by our reasonably reproducible judgments about the type of outcomes and our parsimonious analyses. We used trial registries as a novel source of data for studying and forecasting clinical trial methodology.

The use of registries, while promising, has several challenges. The search engines are not user friendly and do not serve the needs of heterogeneous users. As others have reported, we have also found duplicate trial registration.11 Finally, we found inconsistency in the extent of use of the registries’ fields and in the detail reported (in some cases linking the registration to study or sponsor Web sites in which the information promised was not available). Notwithstanding these concerns, clinical trial registries represent a potentially comprehensive resource for forecasting evidence in a given field thanks to journal editors’ policy initiatives requiring prospective trial registration,12 and the World Health Organization’s efforts to standardize and integrate trial registries.13

<p>| Table 2. Description of 436 Included Randomized Trials and Their Primary and Secondary Outcomes |
|-------------------------------------------------|---------------------------------|-----------------|-----------------|-----------------|-----------------|</p>
<table>
<thead>
<tr>
<th><strong>Trial Characteristics</strong></th>
<th><strong>Total No.</strong></th>
<th><strong>Patient Important</strong></th>
<th><strong>Surrogate</strong></th>
<th><strong>Laboratory</strong></th>
<th><strong>Other Outcomes</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Registry</td>
<td>407 (93)</td>
<td>161 (37)</td>
<td>57 (13)</td>
<td>17 (4)</td>
<td></td>
</tr>
<tr>
<td>ClinicalTrials.gov</td>
<td>146 (36)</td>
<td>139 (86)</td>
<td>50 (87)</td>
<td>17 (100)</td>
<td></td>
</tr>
<tr>
<td>ISRCTN</td>
<td>39 (19)</td>
<td>18 (11)</td>
<td>6 (11)</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Australian New Zealand</td>
<td>15 (6)</td>
<td>4 (3)</td>
<td>1 (2)</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Year</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2004</td>
<td>6 (3)</td>
<td>12 (8)</td>
<td>1 (2)</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>2005</td>
<td>58 (29)</td>
<td>59 (37)</td>
<td>21 (37)</td>
<td>6 (35)</td>
<td></td>
</tr>
<tr>
<td>2006</td>
<td>80 (40)</td>
<td>66 (39)</td>
<td>30 (52)</td>
<td>3 (18)</td>
<td></td>
</tr>
<tr>
<td>2007</td>
<td>57 (28)</td>
<td>25 (16)</td>
<td>5 (9)</td>
<td>8 (47)</td>
<td></td>
</tr>
<tr>
<td>Region of work</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>North America</td>
<td>72 (17)</td>
<td>71 (17)</td>
<td>27 (20)</td>
<td>13 (14)</td>
<td></td>
</tr>
<tr>
<td>Europe</td>
<td>89 (21)</td>
<td>63 (15)</td>
<td>22 (16)</td>
<td>5 (6)</td>
<td></td>
</tr>
<tr>
<td>Australia</td>
<td>25 (10)</td>
<td>20 (8)</td>
<td>2 (1)</td>
<td>1 (1)</td>
<td></td>
</tr>
<tr>
<td>Asia</td>
<td>38 (10)</td>
<td>31 (10)</td>
<td>10 (16)</td>
<td>4 (5)</td>
<td></td>
</tr>
<tr>
<td>Africa</td>
<td>8 (2)</td>
<td>8 (2)</td>
<td>2 (3)</td>
<td>0 (0)</td>
<td></td>
</tr>
<tr>
<td>South America</td>
<td>14 (4)</td>
<td>21 (6)</td>
<td>5 (8)</td>
<td>2 (2)</td>
<td></td>
</tr>
<tr>
<td>No. of centers</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Single</td>
<td>43 (21)</td>
<td>33 (21)</td>
<td>21 (37)</td>
<td>7 (41)</td>
<td></td>
</tr>
<tr>
<td>Few (&lt;10)</td>
<td>31 (15)</td>
<td>17 (11)</td>
<td>10 (18)</td>
<td>2 (12)</td>
<td></td>
</tr>
<tr>
<td>Many (≥10)</td>
<td>64 (33)</td>
<td>52 (32)</td>
<td>7 (12)</td>
<td>6 (35)</td>
<td></td>
</tr>
<tr>
<td>Not reported</td>
<td>63 (31)</td>
<td>59 (36)</td>
<td>19 (33)</td>
<td>2 (12)</td>
<td></td>
</tr>
<tr>
<td>Diabetes type</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>41 (20)</td>
<td>14 (9)</td>
<td>9 (16)</td>
<td>2 (12)</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>102 (51)</td>
<td>132 (81)</td>
<td>39 (67)</td>
<td>11 (64)</td>
<td></td>
</tr>
<tr>
<td>1 or 2</td>
<td>32 (16)</td>
<td>11 (7)</td>
<td>6 (11)</td>
<td>4 (24)</td>
<td></td>
</tr>
<tr>
<td>Other (gestational, etc)</td>
<td>2 (1)</td>
<td>1 (1)</td>
<td>1 (2)</td>
<td>0 (0)</td>
<td></td>
</tr>
<tr>
<td>Type not specified</td>
<td>24 (12)</td>
<td>3 (2)</td>
<td>2 (4)</td>
<td>0 (0)</td>
<td></td>
</tr>
<tr>
<td>Age of study participants, y</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥18</td>
<td>163 (81)</td>
<td>148 (82)</td>
<td>50 (88)</td>
<td>17 (100)</td>
<td></td>
</tr>
<tr>
<td>&lt;18</td>
<td>20 (10)</td>
<td>5 (3)</td>
<td>3 (5)</td>
<td>0 (0)</td>
<td></td>
</tr>
<tr>
<td>Not reported</td>
<td>18 (9)</td>
<td>8 (5)</td>
<td>4 (7)</td>
<td>0 (0)</td>
<td></td>
</tr>
<tr>
<td>Study population size</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤50</td>
<td>28 (14)</td>
<td>22 (14)</td>
<td>26 (45)</td>
<td>2 (12)</td>
<td></td>
</tr>
<tr>
<td>51-200</td>
<td>50 (25)</td>
<td>51 (32)</td>
<td>16 (28)</td>
<td>5 (29)</td>
<td></td>
</tr>
<tr>
<td>&gt;200</td>
<td>117 (58)</td>
<td>82 (50)</td>
<td>13 (23)</td>
<td>7 (41)</td>
<td></td>
</tr>
<tr>
<td>Not reported</td>
<td>6 (3)</td>
<td>6 (4)</td>
<td>2 (4)</td>
<td>3 (18)</td>
<td></td>
</tr>
<tr>
<td>Trial design</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parallel</td>
<td>174 (86)</td>
<td>147 (91)</td>
<td>42 (73)</td>
<td>16 (94)</td>
<td></td>
</tr>
<tr>
<td>Crossover</td>
<td>20 (10)</td>
<td>10 (6)</td>
<td>13 (23)</td>
<td>1 (6)</td>
<td></td>
</tr>
<tr>
<td>Not reported</td>
<td>7 (4)</td>
<td>4 (3)</td>
<td>2 (4)</td>
<td>0 (0)</td>
<td></td>
</tr>
</tbody>
</table>

(continued)
and found that 9% measured patient-important outcomes. Montori et al reported that 42 of 199 (21%) RCTs published in 2003 in top biomedical journals reported patient-important outcomes as their primary outcomes, whereas surrogate outcomes were reported in 38 of 199 (19%) and the rest of the trials measured various laboratory and physiological outcomes. Although it is difficult to directly compare, the use of surrogate end points seems to have increased from 19% of trials in 2003 to 61% of trials noted in this investigation.

**Implications**

One potential reason for the apparent increase in trials measuring surrogate end points is the preference of researchers and funding agencies to obtain results faster, with fewer patients and at lower costs. A major downside of such trials is that the results cannot be used with confidence in patient care because they do not provide information about benefits that patients would consider important, given the paucity of validation of surrogate end points in diabetes and elsewhere.

### Table 2. Description of 436 Included Randomized Trials and Their Primary and Secondary Outcomes (cont)

<table>
<thead>
<tr>
<th>Trial Characteristics</th>
<th>Primary Outcomes</th>
<th>Surrogate Outcomes</th>
<th>Laboratory Outcomes</th>
<th>Other Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type of intervention</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Drug vs Placebo</td>
<td>61 (30)</td>
<td>82 (51)</td>
<td>28 (49)</td>
<td>5 (29)</td>
</tr>
<tr>
<td>Active drug</td>
<td>77 (38)</td>
<td>49 (30)</td>
<td>15 (26)</td>
<td>7 (41)</td>
</tr>
<tr>
<td>Usual care</td>
<td>3 (1)</td>
<td>4 (2)</td>
<td>1 (2)</td>
<td>0</td>
</tr>
<tr>
<td>Other</td>
<td>2 (1)</td>
<td>3 (2)</td>
<td>1 (2)</td>
<td>1 (6)</td>
</tr>
<tr>
<td>Education or lifestyle changes</td>
<td>32 (16)</td>
<td>14 (9)</td>
<td>4 (7)</td>
<td>2 (12)</td>
</tr>
<tr>
<td>Nutrition</td>
<td>8 (4)</td>
<td>7 (4)</td>
<td>5 (9)</td>
<td>1 (6)</td>
</tr>
<tr>
<td>Other interventions</td>
<td>18 (9)</td>
<td>2 (1)</td>
<td>3 (5)</td>
<td>1 (6)</td>
</tr>
<tr>
<td>Funding</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not-for-profit sources</td>
<td>73 (36)</td>
<td>36 (22)</td>
<td>24 (42)</td>
<td>4 (24)</td>
</tr>
<tr>
<td>For-profit sources</td>
<td>103 (51)</td>
<td>116 (73)</td>
<td>25 (44)</td>
<td>12 (70)</td>
</tr>
<tr>
<td>Mixed sources</td>
<td>22 (11)</td>
<td>7 (4)</td>
<td>8 (14)</td>
<td>1 (6)</td>
</tr>
<tr>
<td>Not reported</td>
<td>3 (2)</td>
<td>2 (1)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Planned trial duration, median (range, d)</td>
<td>252 (3-2555)</td>
<td>168 (3-1825)</td>
<td>49 (3-1825)</td>
<td>90 (1-1825)</td>
</tr>
</tbody>
</table>

Abbreviation: ISRCTN, International Standard Randomized Controlled Trial Number.

### Table 3. Univariate and Multivariate Analyses of Trial Characteristics and the Likelihood of Reporting Patient-Important Outcomes

<table>
<thead>
<tr>
<th>Predictor Variables</th>
<th>Primary Outcomes</th>
<th>Any Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Univariate</td>
<td>Multivariate</td>
</tr>
<tr>
<td>Trial duration</td>
<td>1.04 (1.01-1.00)</td>
<td>1.03 (1.01-1.00)</td>
</tr>
<tr>
<td>No. of patients</td>
<td>1.05 (1.02-1.08)</td>
<td>1.10 (1.02-1.19)</td>
</tr>
<tr>
<td>Study sponsor</td>
<td>0.64 (0.31-1.29)</td>
<td>0.58 (0.21-1.56)</td>
</tr>
<tr>
<td>Trial design</td>
<td>0.64 (0.32-1.33)</td>
<td>0.15 (0.05-0.44)</td>
</tr>
<tr>
<td>Trial phase</td>
<td>1.32 (0.69-2.51)</td>
<td>1.19 (0.46-3.10)</td>
</tr>
<tr>
<td>Type of DM</td>
<td>0.52 (0.25-1.08)</td>
<td>0.23 (0.09-0.61)</td>
</tr>
</tbody>
</table>

Abbreviation: DM, diabetes mellitus.
ning to measure patient-important outcomes as secondary end points need to overcome the temptation to select report outcomes with statistically significant results, as well as to report these findings transparently and carefully. Journals also may need to publish the less-than-interesting results to enable meta-analyses of these results to produce precise enough estimates that can guide practice. This approach is not without controversy. For example, experts have expressed concerns about the validity of pooling cardiovascular events across trials of rosiglitazone, since trials reported few events (with some reporting no events), and these trials did not have coronary events as primary end points. The limited number of events across inconsistent trials makes pooled estimates imprecise despite pooling. Whether the field could be advanced more effectively with fewer and larger trials measuring patient-important outcomes deserves greater debate.

The extent to which our findings apply to other chronic conditions is uncertain, but we have noted similar reliance on surrogates in conditions other than diabetes. Examples include the Effect of Ezetimibe Plus Simvastatin Versus Simvastatin Alone on Atherosclerosis in the Carotid Artery (ENHANCE) trial which assessed the effect of ezetimibe and simvastatin on plaque progression, not on cardiovascular events, and trials of testosterone in men assessed the effect on bone markers and lipid fractions, not on fractures or cardiovascular events.

We believe the time has come for a broad consensus on a standard set of important outcomes for patients with diabetes trials, similar to the Outcome Measures in Rheumatology (OMERACT) initiative. The OMERACT approach allows for the uniform measurement of outcomes in RCTs with emphasis on outcomes that experts—and ultimately patients—thought would better capture the experience of rheumatological conditions.

CONCLUSIONS

Only 18% of ongoing RCTs in diabetes will measure outcomes of importance to patients as primary end points.

Author Contributions: Dr. Montori had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. Dr. Murad and Gandhi contributed equally to this article.

Study concept and design: Gandhi, Murad, Isley, Montori.

Acquisition of data: Gandhi, Murad, Fujiyoshi, Mullan, Flynn, Elamin, Swiglo.

Analysis and interpretation of data: Gandhi, Murad, Fujiyoshi, Flynn, Elamin, Swiglo, Isley, Guyatt, Montori.

Drafting of the manuscript: Gandhi, Murad, Fujiyoshi, Flynn, Isley, Guyatt, Montori.

Critical revision of the manuscript for important intellectual content: Gandhi, Murad, Fujiyoshi, Mullan, Flynn, Elamin, Swiglo, Isley, Guyatt, Montori.

Statistical analysis: Gandhi, Murad, Flynn, Montori.

Administrative, technical, or material support: Gandhi, Murad, Flynn, Elamin, Montori.

Study supervision: Montori.

Financial Disclosures: None reported.

Funding/Support: There were no funding sources for this study.

Additional Information: For a list of the trial registry identifiers used in this study, see the article online at http://www.jama.com.

REFERENCES


5. Richter B, Berger M. Randomized controlled trials remain fundamental to clinical decision making in type II diabetes mellitus: a comment to the debate on randomized controlled trials. BMJ. 2007;334(7597):786.


