Success, Failure, and Transparency in Biomarker-Based Drug Development
A Case Study of Cholesteryl Ester Transfer Protein Inhibitors

Spencer Phillips Hey, PhD; Jessica M. Franklin, PhD; Jerry Avorn, MD; Aaron S. Kesselheim, MD, JD

Background—Although biomarkers are used as surrogate measures for drug targeting and approval and are generally based on plausible biological hypotheses, some are found to not correlate well with clinical outcomes. Over-reliance on inadequately validated biomarkers in drug development can lead to harm to trial subjects and patients and to research waste. To shed greater light on the process and ethics of biomarker-based drug development, we conducted a systematic portfolio analysis of cholesterol ester transfer protein inhibitors, a drug class designed to improve lipid profiles and prevent cardiovascular events. Despite years of development, no cholesterol ester transfer protein inhibitor has yet been approved for clinical use.

Methods and Results—We searched PubMed and Clinicaltrials.gov for clinical studies of 5 known cholesterol ester transfer protein inhibitors: anacetrapib, dalcetrapib, evacetrapib, TA-8995, and torcetrapib. Published reports and registration records were extracted for patient demographic characteristics and study authors’ recommendations of clinical usage or further testing. We used Accumulating Evidence and Research Organization graphing to depict the portfolio of research activities and a Poisson model to examine trends. We identified 100 studies for analysis that involved 96,944 human subjects. The data from only 41,201 (42%) of the human subjects had been presented in a published report. For the 3 discontinued cholesterol ester transfer protein inhibitors, we found a pattern of consistently positive results on lipid-modification end points followed by negative results using clinical end points.

Conclusions—Inefficiencies and harms can arise if a biomarker hypothesis continues to drive trials despite successive failures. Regulators, research funding bodies, and public policy makers may need to play a greater role in evaluating and coordinating biomarker-driven research programs. (Circ Cardiovasc Qual Outcomes. 2017;10:e003121. DOI: 10.1161/CIRCOUTCOMES.116.003121.)

Key Words: biomarkers ■ cholesterol ester transfer proteins ■ ethics, research ■ lipids ■ lipoproteins
WHAT IS KNOWN

• Despite a compelling biomarker-based rationale, cholesteryl ester transfer protein inhibitors have not demonstrated any meaningful clinical benefits.
• A comprehensive review of development programs for cholesteryl ester transfer protein inhibitors has not been performed.

WHAT THE STUDY ADDS

• Less than half of the patient–subject data from cholesteryl ester transfer protein inhibitor trials conducted through 2016 have been published or made available to the wider scientific and medical communities.
• The majority of published data reports positive findings on biomarkers; however, the negative clinical results have important implications for future research in this area.
• The entire portfolio of cholesteryl ester transfer protein testing, which consistently shows that lipid-modifying biomarkers do not predict clinical cardiovascular benefit, should be made known to patients and investigators involved in trials of drugs in this class.

The testing of the CETP class of drugs makes it a useful case for examining the scientific, social, and ethical implications of biomarker-driven drug development. There are examples of successful drug classes that have emerged after initial failure: for example, dabigatran (Pradaxa), an oral anti-coagulant that acts directly by inhibiting thrombin, received FDA approval only after several previous drugs in the same class failed in development,10 and several additional novel oral anti-coagulants have followed. Similarly, the thiazolidinediones, such as troglitazone (Rezulin) and rosiglitazone (Avandia) to manage diabetes mellitus, were found to cause severe hepatotoxicity and myocardial infarction, respectively, although the thiazolidinedione, pioglitazone, remains in use. But the CETP trajectory raises the question of how the ethical obligations of research stakeholders might evolve as new evidence about biomarkers accumulates. These are especially timely questions in light of recent legislation calling for greater use of biomarkers in drug evaluation and the increasing concern about the efficiency of clinical research activities.11–13

To better understand the process of biomarker-based drug development, we examined the CETP inhibitor class in depth, systematically mapping the state of accumulating evidence to provide insight into the scientific and ethical basis for continued development efforts in this area.14

Methods

Study Design and Sources Used

We performed a keyword search of PubMed (filtered by clinical trials) and Clinicaltrials.gov in October 2015, using CETP inhibitor names and their known variants linked by a Boolean “or” operator. The list of keywords included anacetrapib (MK-0859), dalcetrapib (JTT-705), evacetrapib (LY2484595), TA-8995, and torcetrapib (CP-529,414). These database searches were then supplemented with a manual review of references in the trial reports.

Inclusion and Exclusion Criteria

From our initial search results, duplicates were removed and registry records with incoherent properties (eg, unpublished trials that were registered as phase 3, but only planned to enroll a few dozen patients–subjects) were set aside as registration errors. In vitro studies, in vivo non-human studies, letters, reviews, nonprimary analyses, and studies lacking adequate methodological reporting were excluded from our analysis.

Data Extraction

If a published report was available for a trial, one of us (S.P.H.) initially extracted the trial phase, primary end point, study completion date, publication date, sample size, and authors’ qualitative recommendation. Author recommendations were double-coded for 25% of our sample, according to the following scheme: studies were coded as positive if they reported achieving their primary end point with statistical significance (if powered to do so) and if the authors explicitly endorsed further testing or the clinical usage of the experimental drug. Studies were coded negative if they did not achieve their primary end point with statistical significance or the authors explicitly recommended against further testing of the experimental drug. Studies were coded mixed if the authors did not make a strong recommendation either for or against further testing or emphasized troubling safety signals. Terminated studies were also classified as negative.

If no published report was available for a study, then we sought to obtain as much of the relevant data as possible from the Clinicaltrials.gov record, using the stated “primary completion date” or “completion date” in place of the date of publication. If no study completion date could be found in the public records, we e-mailed the corresponding authors to request the information.

Data Analysis

To represent the research portfolio, we used Accumulating Evidence and Research Organization graphing.14 This method depicts studies as individual nodes on a graph, arranged by time of publication (or study completion) along the x-axis, grouping studies first by the type of CETP inhibitor and then by the primary outcome.

To evaluate the trends in number of early-phase trials and number of patients per year before and after 2007—the year that the decisive, negative trial for torcetrapib was published15—we fit a segmented Poisson regression model with terms for trend before and after 2007. Based on initial analyses and visual inspection of data, we did not include a term for an immediate change in 2007. We also used a Kaplan–Meier analysis to examine time from study completion to publication of results.

Results

Our search identified 108 studies, from which we identified 100 for analysis (Figure 1) that involved 96,944 human subjects.15–49 The demographic properties of this sample, broken down by CETP agent, are described in the Table. Eli Lilly’s testing program for evacetrapib was the shortest, comprising 4 years. Roche’s program for dalcepatrrib was the longest, lasting 13 years although Merck’s and Amgen’s programs for anacetrapib and TA-8995, respectively, are still ongoing.

At the time of our analysis, 62% of CETP inhibitor studies have been published, and the data from only 41,201 (42%) of the human subjects has been presented in a published report. No results have yet been deposited on Clinicaltrials.gov although federal regulations for these trials only require
results to be made available on Clinicaltrials.gov if a medication has been licensed by the FDA. Half of the studies (50/100; 50%) in our sample were pharmacokinetic or pharmacodynamic investigations, 5% primarily examined dose and safety issues, 34% used lipid-level modification (Δ lipid) end points, 6% used a nonlipid biomarker (such as carotid intima media thickness or flow-mediated dilation), and 5% used a clinical cardiovascular disease end point (including one ongoing trial of anacetrapib).

**Publication Outcomes**

Most (53/61; 87%) of the published trials in our sample reached positive conclusions in terms of their stated goals. All (23/23; 100%) of the published lipid-modification trials reported positive results in modification of cholesterol levels (although there are 4 unpublished, terminated studies). However, no published trial using a cardiovascular disease end point was positive.15,30 By contrast, 3 negative trials of torcetrapib using nonlipid biomarker end points have been published.38–40 These studies found that torcetrapib increased systolic blood pressure but did not affect the rate of change in the maximum intima media thickness or atheroma volume.39

Only 3 published studies fit our criteria for mixed conclusions; all of them involving dalcetrapib.32,34 One report described significant interactions between dalcetrapib and the lipase inhibitor, orlistat, in healthy volunteers; the other described the results of 2 trials demonstrating alterations in the pharmacokinetics of dalcetrapib for subjects with hepatic or renal impairment.34 These mixed results were all published subsequent to the decisive, negative phase 3 trial of dalcetrapib.30

For our time-to-publication analysis, no completion date information could be found for 18 studies. This left 82 studies to evaluate the time from completion of the study until publication of study results. As shown in Figure 2, 21% (95% confidence interval, 11%–39%) of studies remain unpublished 6 years after completion.

**Development Trajectory and Trends**

Figure 3 depicts the Accumulating Evidence and Research Organization graph for this portfolio of 100 investigations. In the 3 failed developmental trajectories (torcetrapib, dalcetrapib, and evacetrapib), there is a pattern of positive lipid-modification trials appearing over a period of ≈4 years that

---

**Table. Demographic Properties of Sample Studies**

<table>
<thead>
<tr>
<th>Manufacturer</th>
<th>Dalcetrapib</th>
<th>Torcetrapib</th>
<th>Anacetrapib</th>
<th>Evacetrapib</th>
<th>TA-8995</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total N studies</td>
<td>28</td>
<td>17</td>
<td>21*</td>
<td>26</td>
<td>8</td>
<td>100</td>
</tr>
<tr>
<td>% N published</td>
<td>82%–85%</td>
<td>71%–76%</td>
<td>90%–92%</td>
<td>19%–67%</td>
<td>38%–69%</td>
<td>62%</td>
</tr>
<tr>
<td>% N with results on ct.gov</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Total Patient enrollment</td>
<td>21,530</td>
<td>25,841</td>
<td>34,736</td>
<td>14,088</td>
<td>749</td>
<td>96,944</td>
</tr>
<tr>
<td>% Patient in published reports</td>
<td>85%</td>
<td>71%</td>
<td>10%</td>
<td>5%</td>
<td>70%</td>
<td>42%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>End points (N positive/N published)</th>
<th>PK/PD</th>
<th>Dose/safety</th>
<th>Δ lipid</th>
<th>Other marker</th>
<th>CVD</th>
</tr>
</thead>
<tbody>
<tr>
<td>PK/PD</td>
<td>12 (7/10)</td>
<td>1 (1/1)</td>
<td>12 (12/12)</td>
<td>20 (2/2)</td>
<td>5 (2/2)</td>
</tr>
<tr>
<td>Dose/safety</td>
<td>2 (2/2)</td>
<td>0</td>
<td>1 (1/1)</td>
<td>1 (1/1)</td>
<td>1 (0/0)</td>
</tr>
<tr>
<td>Δ lipid</td>
<td>9 (8/8)</td>
<td>12 (7/7)</td>
<td>7 (5/5)</td>
<td>4 (2/2)</td>
<td>2 (1/1)</td>
</tr>
<tr>
<td>Other marker</td>
<td>3 (2/2)</td>
<td>3 (0/3)</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>CVD</td>
<td>2 (0/1)</td>
<td>1 (0/1)</td>
<td>1 (NA)</td>
<td>1 (0/0)</td>
<td>0</td>
</tr>
</tbody>
</table>

Proportion of positive results refers only to published studies. Δ lipid indicates lipid-level modification end points; CVD, cardiovascular disease; and PK/PD, pharmacokinetic/pharmacodynamic.

*Because 1 trial of anacetrapib is ongoing, a denominator of 20 and 99 is used to calculate the percentage of published trials.
concludes with ≥1 large, negative phase 3 trials. Three of the 5 trajectories (anacetrapib, evacetrapib, and TA-8995) also do not include any trials with nonlipid biomarkers as primary end points. After the Investigation of Lipid Level Management to Understand its Impact in Atherosclerotic Events phase 3 trial of torcetrapib, there is an obvious spike in pharmacokinetic/pharmacodynamic research activities in the other 4 trajectories. There is also seemingly greater caution about conducting phase 3 trials and a shifting of research burdens to earlier phases.

Results of our trend analysis are presented in Figure 4. We found that there was no difference in trend in the number of trials before and after 2007. However, the overall rate of trials per year was significantly higher after 2007 (1.29 versus 8.75; \( P=0 \)). We also found that before 2007, there was an average annual increase of 36% in the number of patient–subjects involved in trials each year. The annual increase after 2007 was smaller (16%) and was different from the previous trend (\( P<0.001 \)).

For our time-to-publication analysis, no completion date information could be found for 18 studies. This left 82 studies to evaluate the time from completion of the study until publication of study results. As shown in Figure 4, 21% (95% confidence interval, 11%–39%) of studies remain unpublished 6 years after completion.

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Year</th>
<th>02</th>
<th>03</th>
<th>04</th>
<th>05</th>
<th>06</th>
<th>07</th>
<th>08</th>
<th>09</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
</tr>
</thead>
<tbody>
<tr>
<td>FK/PD</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Figure 2.** Kaplan–Meier plot of time to publication.
Comprehensively mapping the research portfolio of CETP inhibitors, an experimental drug class driven by lipid-modification biomarkers, we observed a consistent pattern across each of the 3 abandoned drug trajectories of uniform reporting of positive study results on lipid-modification outcomes, followed by negative results on clinical outcomes. This pattern simultaneously confirms the mechanistic hypothesis that CETP inhibitors raise HDL and lower LDL while disproving the clinical hypothesis that this biological mechanism is a sufficient indicator of clinical cardiovascular benefit.

The existence of the nonpositive pharmacokinetic/pharmacodynamic trials of dalcetrapib also suggests that it is not simply the case that only drugs with uniformly positive early-phase trials will be advanced to the later phases of testing. On the contrary, negative or inconclusive phase 1 and 2 results can play a valuable role in helping to define the boundaries of clinical usage for a novel biomarker or drug—as is found in the mixed dalcetrapib studies, which explored safety signals in patients taking multiple medications or having comorbidities. Therefore, publication bias and biases in study design may provide a more plausible explanation for the rarity of negative early-phase results across this research portfolio.

The rate of nonpublication across this portfolio is also high and amounts to tens of thousands of human subjects. Although a time delay between trial completion and publication can explain the low proportion of published participant data for some of these drugs, this justification is certainly not available for torcetrapib, which was abandoned 10 years ago, and highly questionable for dalcetrapib, abandoned 4 years ago. Failing to publish or make publicly available the data from \(\approx 30\%\) of participants (in the case of torcetrapib) should be ethically unacceptable, particularly given that these data could be of great value for protecting participants in trials of other CETP inhibitors. Fortunately, recent changes to the US federal regulations will require that trial outcomes such as these are deposited at ClinicalTrials.gov within 1 year of the completion date.\(^7^1\)

The lack of nonlipid biomarker trials in the anacetrapib, evacetrapib, and TA-8995 trajectories is also problematic. Particularly after the failure of torcetrapib, several preclinical studies and commentaries were published describing the limitations of using HDL as a biomarker of cardiovascular benefit.\(^8,^7^2\) That would have been an opportunity for reconsideration of the future of CETP research programs under the assumption that raising HDL (and lowering LDL) by this mechanism is a sufficient end point alone.

Nevertheless, our results may be encouraging from a regulatory perspective because it confirms that the FDA insistence on clinical outcome data, rather than on biomarker data, for this class of drugs makes regulatory sense. However, ensuring an efficient biomarker research portfolio and adequately protecting human subjects may require the FDA to do more. For example, the FDA oversees all Investigational New Drug applications, which permit human trials to move forward. The results (and even the very existence) of failed trials are often not made public by FDA, and their details are considered the private property of the companies that conducted the research. Yet, subsequent Investigational New Drug applications are supposed to include reference to other Investigational New Drugs if pertinent.\(^7^3\) The FDA, therefore, has a privileged perspective on the entire research portfolio for a given drug class and could serve as a valuable relay for such information—alerting institutional review boards (IRBs), research participants, funding bodies, or other stakeholders to the array of ongoing activities that could inform their judgments to approve, participate in, or fund research.

IRBs, in their responsibility to ensure that a study minimizes risks and maximizes its contribution to generalizable knowledge, may also need to exercise greater scrutiny when evaluating a biomarker-driven trial. Valid judgments of social value, for example, require that an IRB systematically evaluate alternative designs and research procedures. Knowledge of existing data for a particular drug class could help to structure and facilitate this systematic analysis. For example, when judging the benefit that is likely to emerge from a phase 2 study designed to measure the HDL-raising capacity of a new agent, it would be valuable for an IRB to know whether there are other ongoing trials that are already evaluating this same end point or how many previous trials have ended in failure and resulted in patient harm.

Potential research participants should also be informed about the totality of research activities surrounding a given trial during the informed consent process. Narrative reviews and summary statistics about the known risks and benefits of the trial procedures can convey some of this information. However, there is considerable evidence that potential participants frequently fail to fully understand the state of evidence and how it justifies the activities and procedures to which they are consenting.\(^7^4\) Providing existing risk–benefit data, an Accumulating Evidence and Research Organization graph, study trends, or time-to-publication information for potential research participants as they contemplate enrolling in a trial would help participants’ understanding, enhancing their capacity to make an informed decision. This information
could also serve as a useful point of conversation between the researcher and potential participants.

This approach to mapping a research portfolio also accords well with the recent recognition from the US National Heart, Lung, and Blood Institute on the need to focus its funding on questions of high social value. The National Heart, Lung, and Blood Institute has announced that it will no longer be funding many smaller studies and will instead favor larger studies that address questions of greater clinical importance. Even though the CETP research portfolio was privately funded, understanding more about the patterns and dynamics of research portfolios can help to facilitate this analysis, complementing quantitative meta-analytic approaches—like that recently deployed to examine the association between LDL and reduction in cardiovascular risk across all LDL-modifying drug classes—by elucidating systematic differences between highly productive and disappointing domains of research.

This study has some limitations. First, our literature search may not have captured all relevant CETP research activities. However, because we are not offering quantitative estimates for drug effect sizes, we do not see this as a serious threat to our conclusions. On the contrary, the existence of additional studies would only further support our concerns about inefficiency and potentially amplify the apparent problems with nonpublication or unavailability of data.

Second, our analysis is largely based on publicly accessible information derived from registration records and published trial reports. The resulting reconstruction of the accumulating state of evidence over time may not adequately reflect the true state of knowledge because it evolved in real time for the 5 manufacturers, who also had access to their own proprietary data. Placement of nodes according to the date of publication (when available), rather than the study completion date, may also not correspond to the precise time when evidence became available to decision makers.

Third, our decision to stratify the Accumulating Evidence and Research Organization graph by study end point could obscure variations in the hypotheses among investigations. As a result, the 20 pharmacokinetic/pharmacodynamic investigations of evacetrapib, for example, do not definitively show nonproductive activity. Some or all of investigations could have explored scientifically relevant questions, such as drug–drug interactions in patients taking the CETP along with a statin.

Conclusions
This analysis provides insight into the risks and challenges for biomarker-driven drug development. When used as surrogate measures, biomarkers can be helpful in the diagnosis and management of disease, but they must be rigorously evaluated and validated. Systematic reviews and meta-analyses of biomarker research portfolios can help researchers clarify the steps in the drug development pathway for a promising class of drugs.

There might be scientific explanations accounting for each case of failure that would justify continuing testing on other CETP inhibitors, because as noted, it is sometimes the Nth drug in a class that first demonstrates clinical benefit. Our current system of research oversight allows individual companies and investigators to make their own determinations about whether the theory underlying a given biomarker remains plausible enough to continue supporting development. However, minimizing risks to volunteers and safeguarding the integrity of the research enterprise requires sufficient transparency so that successive research efforts and regulatory mechanisms can adequately balance the knowledge of previous failed research efforts with the prospect that subsequent efforts are designed to offer a favorable risk/benefit ratio. This is hampered by the unavailability of companies’ data on the clinical details of failed drug trials.

The possibility that the investigators, IRBs, and participants involved with the activities related to CETP inhibitors were not aware of all the other ongoing research activities when they initiated, approved, or consented to, these studies reinforces the importance of making such information available in a timely fashion. Such an outcome may require rethinking of current data protection rules and enhancing the transparency of ongoing research programs. Efforts in this direction are already underway.

It is possible that at some point, a CETP inhibitor will be found to be clinically useful. In the meantime, our review has generated some pressing questions: What evidence is sufficient to decisively disconfirm a biomarker hypothesis? What groups, if any, should be empowered to act on such accumulating evidence? When should trial information be made accessible to investigators, IRBs, and patients? What patterns of research activity distinguish an efficient from inefficient research portfolio, and who should make such choices? Considering answers to these questions may help investigators and policymakers arrive at consensus over the structure of the research enterprise and allocation of limited human and material resources in the context of increased investment in biomarker-based drug development.

Sources of Funding
This study was completed at the Program on Regulation, Therapeutics, and Law Biomarker Research Consortium, funded by the Laura and John Arnold Foundation. Dr Kesselheim’s work is also funded by the Engelberg Foundation and the Harvard Program in Therapeutic Science. Dr Kesselheim has also received unrelated research grants from the FDA.

Disclosures
None.

References
2. The President’s Council of Advisors on Science and Technology. Report to the President on propelling innovation in drug discovery, development, and evaluation. Executive Office of the President. 2012.
by guest on January 17, 2018 http://circoutcomes.ahajournals.org/ Downloaded from


Success, Failure, and Transparency in Biomarker-Based Drug Development: A Case Study of Cholesteryl Ester Transfer Protein Inhibitors
Spencer Phillips Hey, Jessica M. Franklin, Jerry Avorn and Aaron S. Kesselheim

_Circ Cardiovasc Qual Outcomes_. 2017;10:
doi: 10.1161/CIRCOUTCOMES.116.003121
_Circulation: Cardiovascular Quality and Outcomes_ is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2017 American Heart Association, Inc. All rights reserved.
Print ISSN: 1941-7705. Online ISSN: 1941-7713

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://circoutcomes.ahajournals.org/content/10/6/e003121

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in _Circulation: Cardiovascular Quality and Outcomes_ can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at:
http://www.lww.com/reprints

Subscriptions: Information about subscribing to _Circulation: Cardiovascular Quality and Outcomes_ is online at:
http://circoutcomes.ahajournals.org//subscriptions/