



Effective Health Care Program

Testing of CYP2C19 Variants and Platelet Reactivity for Guiding Antiplatelet Treatment *Executive Summary*

Background

Burden of Disease and Clinical Setting

Approximately 82 million Americans currently suffer from some form of cardiovascular disease.¹ In the United States, coronary heart disease alone is the cause of 1 of every 6 deaths, and stroke, 1 of every 18 deaths.² There were approximately 7 million inpatient cardiovascular operations and procedures in the United States in 2007, of which 1 million were either percutaneous coronary intervention (PCI) or coronary artery bypass graft (CABG) surgeries.¹

Randomized controlled trials have established dual antiplatelet treatment with clopidogrel and aspirin as the current standard of care for medical and interventional management of acute coronary syndromes.³ Dual antiplatelet treatment is also recommended for patients undergoing PCI⁴ with placement of stents (either bare metal or drug eluting). Randomized controlled trials support the use of clopidogrel in patients who have experienced acute cerebrovascular events (e.g., stroke) and those with peripheral arterial disease.^{3,5-8} For patients with atrial fibrillation and contraindications to vitamin K antagonists, the ACTIVE A (Atrial Fibrillation Clopidogrel Trial with

Effective Health Care Program

The Effective Health Care Program was initiated in 2005 to provide valid evidence about the comparative effectiveness of different medical interventions. The object is to help consumers, health care providers, and others in making informed choices among treatment alternatives. Through its Comparative Effectiveness Reviews, the program supports systematic appraisals of existing scientific evidence regarding treatments for high-priority health conditions. It also promotes and generates new scientific evidence by identifying gaps in existing scientific evidence and supporting new research. The program puts special emphasis on translating findings into a variety of useful formats for different stakeholders, including consumers.

The full report and this summary are available at www.effectivehealthcare.ahrq.gov/reports/final.cfm.

Irbesartan for Prevention of Vascular Events) trial suggested that the combination of clopidogrel and aspirin is more effective than aspirin alone for preventing thromboembolic disease.⁹



Since the approval of clopidogrel by the U.S. Food and Drug Administration (FDA) for routine clinical use, the drug has become one of the most commonly prescribed agents in the United States. However, patient response to clopidogrel-based antiplatelet therapy is variable both between patients and across multiple measurements within a patient, with some patients showing no or minimal platelet response to clopidogrel administration (often termed clopidogrel “nonresponsiveness” or “resistance”). Alternatives to standard clopidogrel treatment include higher dose clopidogrel regimens and the use of other antiplatelet agents, such as prasugrel or ticagrelor.¹⁰⁻¹³ Given the availability of alternative antiplatelet strategies and concern about adverse clinical outcomes in clopidogrel nonresponders, research has focused on methods to identify patients who are unlikely to benefit from clopidogrel-based treatment. The question of identifying the optimal antiplatelet therapy may also carry cost implications because generic clopidogrel products are now available in the United States.^a

Clopidogrel Metabolism

To be biologically active, clopidogrel must be transformed to the active metabolite R-130964 by members of the CYP enzyme system, primarily the enzyme CYP2C19. R-130964 acts by binding irreversibly to the P2Y₁₂ receptor (the adenosine diphosphate [ADP] receptor) on the surface of platelets and inhibits platelet aggregation for the life cycle of the platelet.^{14,15}

However, the relationship between genotype and clinical outcomes is not straightforward. The fact that each individual carries two CYP2C19 alleles results in combinations of alleles of varying enzymatic activity. The combined effect of the two alleles on actual enzymatic activity levels depends on the “true” genetic model of CYP2C19 alleles (dominant, recessive, additive, or codominant). Unfortunately, the true underlying genetic model for CYP2C19 variants is not known with certainty.¹⁶ This is of particular concern, as the allele frequency of CYP2C19 variants is heterogeneous across populations of different ethnicities, resulting in different genotype prevalences. For example, data from the Third National Health and Nutrition Examination Survey (NHANES III) showed statistically significant heterogeneity in the prevalence of the *2 allele among non-Hispanic whites, Mexican-Americans, and non-Hispanic blacks. Non-Hispanic blacks had the highest prevalence of the *2 allele (18.3%) and of homozygotes for that allele (*2/*2; 3.8%).¹⁷ Studies have shown that the prevalence of the rare

allele is even higher in East Asian populations, with *2 allele frequencies as high as 30-40 percent.¹⁸⁻²⁰

Furthermore, the CYP2C19 genotype is only one of many determinants of the effect of clopidogrel on platelet reactivity. For example, a genome-wide association study recently demonstrated that the *2 allele accounts for only 12 percent of the total observed variation in clopidogrel responsiveness in a selected white population.²¹ Several studies have demonstrated that environmental factors and patient characteristics, such as body mass index, diabetes, and smoking habits, can influence platelet reactivity.

Predicting Response and Guiding Antiplatelet Treatment

There are currently two main approaches to determine whether a patient will have a poor response to clopidogrel: (1) genetic testing to see whether the patient has a genotype that is associated with reduced ability to metabolize clopidogrel (a “poor-metabolizer” phenotype) and (2) direct testing of the patient’s blood while the patient is taking clopidogrel to see whether the platelets actually have become less prone to aggregate in response to specific agonists (phenotypic testing for platelet reactivity).

Genetic Tests for CYP2C19 Variants

Genetic testing for one or more genetic variants can be performed with various genotyping methods. For biallelic variants, these methods identify homozygotes for each variant and heterozygotes. Testing for CYP2C19 variants requires a sample of somatic genetic material, usually obtained from a blood sample or from buccal swabs. Because allelic variants at the CYP2C19 locus do not change over a person’s lifetime, testing done at any time point is representative of the person’s genotype.

Measurement of Platelet Reactivity

Phenotypic testing measures the reactivity of platelets while a patient is taking clopidogrel (on-clopidogrel platelet reactivity). Several assays for measuring platelet reactivity are available. These include rapid point-of-care platelet function assays (e.g., VerifyNow, Platelet Function Analyzer [PFA]-100, Plateletworks); measurements of mediators of reactivity (e.g., vasodilator-stimulated phosphoprotein [VASP] phosphorylation using flow cytometry); and functional assays (e.g., aggregometry using appropriate agonists). We refer to all these assays as “phenotypic tests” because they attempt to measure an intermediate clinical phenotype (platelet reactivity).²²

^a FDA release, available at www.fda.gov/NewsEvents/Newsroom/PressAnnouncements/ucm304489.htm; last accessed: October 16, 2012.

Scope

We performed a comparative effectiveness review regarding the utility of testing for CYP2C19 variants and platelet reactivity for guiding antiplatelet treatment. We evaluated the analytic validity, predictive utility, and comparative effectiveness of genetic and phenotypic tests as biomarker tests (and of relevant test-and-treat strategies) for guiding antiplatelet therapy in patient populations who are eligible for clopidogrel treatment.

Key Questions

On the basis of the original topic nomination and an extensive process of topic development and refinement, we formulated the following Key Questions to guide the review:

Key Question 1. In patient populations who are candidates for clopidogrel therapy, does genetic testing for CYP2C19 variants predict intermediate and clinical outcomes following treatment initiation?

- a. What is the analytic validity (technical test performance) of the various assays used for CYP2C19 genetic testing?
- b. What is the clinical validity (predictive accuracy) of genetic testing for predicting intermediate and clinical outcomes in patients who are receiving clopidogrel therapy?
- c. Do the following factors modify the association between genetic test results and clinical outcomes?
 - i. Comedications
 - ii. Patient-level factors (e.g., race or ethnicity, age, sex, disease severity, or comorbidities)
 - iii. Test-related factors (e.g., between-assay differences)
 - iv. System-level factors (e.g., settings where testing is performed)

Key Question 2. In patient populations receiving clopidogrel therapy, does phenotypic testing of platelet reactivity predict intermediate and clinical outcomes?

- a. What is the analytic validity (technical test performance) of the various assays used in phenotypic testing of platelet reactivity?
- b. What is the clinical validity (predictive accuracy) of phenotypic testing for predicting intermediate and clinical outcomes in patients who are receiving clopidogrel therapy?
- c. Do the following factors modify the association between phenotypic test results and clinical outcomes?

- i. Comedications
- ii. Patient-level factors (e.g., race or ethnicity, age, sex, disease severity, or comorbidities)
- iii. Test-related factors (e.g., between-assay differences)
- iv. System-level factors (e.g., settings where testing is performed)

Key Question 3. What is the comparative effectiveness of alternative test-and-treat strategies (including a no-testing strategy) for therapeutic decisionmaking regarding antiplatelet therapy among patients who are candidates for clopidogrel-based treatment?

- a. What is the comparative effectiveness of the following testing strategies on therapeutic decisionmaking, platelet reactivity during followup, and clinical outcomes in patients who are candidates for antiplatelet treatment?
 - i. Genetic testing for CYP2C19
 - ii. Genetic testing for CYP2C19 followed by phenotypic testing for platelet reactivity
 - iii. Phenotypic testing for platelet reactivity
 - iv. No testing
- b. How do modifying factors (e.g., race or ethnicity, age, sex, comorbidities, diet, or the time between conducting the test and obtaining results) affect the association of alternative phenotypic or genetic test-and-treat strategies and patient outcomes? Alternative test-guided treatments can include nonclopidogrel antiplatelet agents or high-dose clopidogrel regimens.

Key Question 4. What are the potential adverse effects or harms from genetic or phenotypic testing per se or from test-directed treatments?

Analytic Framework

We developed an analytic framework (Figure A) that maps the Key Questions within the context of populations, interventions, comparators, and outcomes of interest, as well as the chain of logic that evidence must support to link the interventions to health outcomes. Analytic and clinical validity were straightforward to represent in the analytic framework (Key Questions 1a, 1b, 2a, and 2b). Regarding treatment decisionmaking (Key Question 3a), we conceptualized the analytic framework as a decision problem, wherein patients' disease can be managed with one of the following approaches (depicted from top to bottom in the flow diagram):

- Undergo genetic testing and then base the treatment decision on the test results.

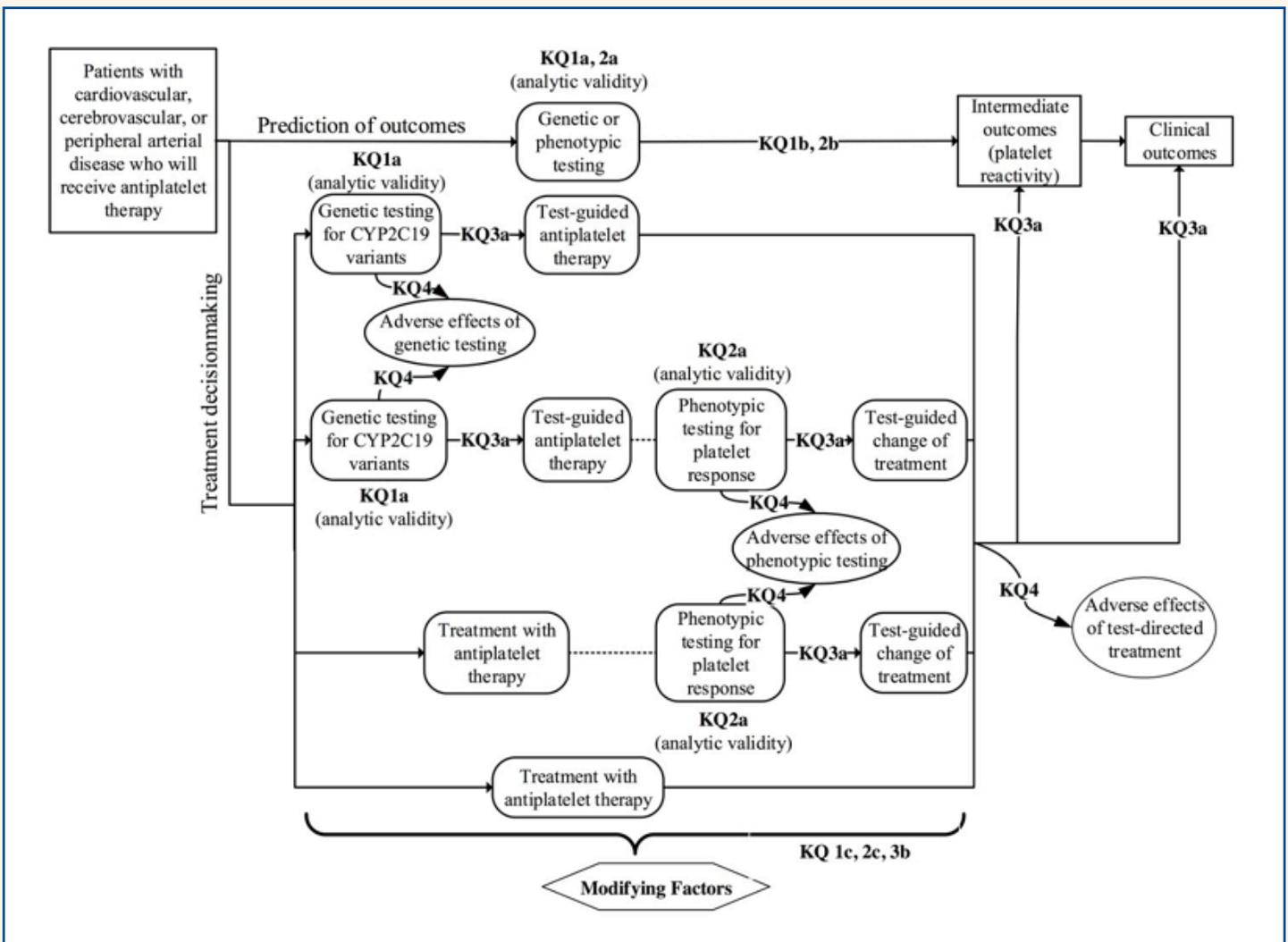
- Undergo genetic testing and then base the treatment decision on the test results. After receiving therapy for an adequate period of time, undergo phenotypic testing for platelet reactivity and use the results to decide whether the treatment strategy should be modified.
- Receive standard treatment directly and, after an appropriate amount of time, undergo phenotypic testing for platelet reactivity and use the results to decide whether the treatment strategy should be modified. Use of phenotypic testing (but not genetic testing) as a monitoring test can be considered a variation of this strategy in which the test is repeatedly performed.

- Receive antiplatelet therapy without undergoing any testing (the current standard of care).

The above strategies were identified as the most prevalent in published studies by preliminary searches conducted in preparation of this review. Additional variations of these strategies were uncovered by the full evidence review.

Modifiers of the effects of testing on outcomes, in terms of both predictive ability and decisionmaking, were reviewed in Key Questions 1c, 2c, and 3b. Tests and test-directed treatments may be associated with harms, investigated in Key Question 4.

Figure A. Analytic framework



Note: KQ = Key Question.

Methods

Literature Search, Study Selection, and Data Extraction

We conducted literature searches for studies in MEDLINE®, the Cochrane Central Trials Registry, and the Cochrane Database of Systematic Reviews (from inception through July 27, 2012) without any language restriction. Our search included terms for the populations, tests, and drugs of interest. (See Appendix A of the full report for complete search strings, which were extensively validated against previous reviews on the tests of interest.) We also performed searches of the Human Genome Epidemiology Network (HuGENet) database and National Institutes of Health Genetic Association Database, using the same cutoff date (July 27, 2012). Finally, we performed a targeted search of the FDA Web site (last search performed on April 25, 2012).

We considered both comparative and noncomparative studies for Key Questions pertaining to prognostic ability but focused on comparative studies of alternative test-and-treat strategies. We did not include non-English-language studies. We excluded narrative reviews, editorials, letters to the editor, and other papers not presenting primary research data. We also excluded studies reporting exclusively on healthy individuals. Studies conducted in all relevant care settings were included. We contacted the authors of the primary studies to verify cases of suspected overlap.

A single investigator extracted data from each study; quantitative results were verified by a second reviewer. Disagreements were resolved by consensus involving a third investigator. We extracted information on the following items: patient selection criteria, population characteristics, sample size, study design, analytic details, and outcomes.

Risk-of-Bias Assessment of Individual Studies

For assessing the risk of bias, we followed recently updated guidance from the Agency for Healthcare Research and Quality “Methods Guide for Effectiveness and Comparative Effectiveness Reviews” (Methods Guide), available on the Effective Health Care Web site.^{23,24} Two independent reviewers evaluated the risk of bias for each study, and disagreements were resolved by consensus including a third reviewer.

For studies of analytic validity (Key Questions 1a and 2a), we compiled a list of 11 items for assessing quality and completeness of reporting based on a recent AHRQ Methods Report.²⁵

For studies of predictive ability (Key Questions 1b, 1c, 2b, and 2c), we based our assessment on the recently proposed Quality Assessment of Diagnostic Accuracy Studies (QUADAS)-2 instrument,²⁶ a new version of the validated QUADAS list of quality items²⁷⁻²⁹ for systematic reviews of medical tests. We used the number of items scored as having been adequately addressed (i.e., indicative of low risk of bias) to classify studies into three categories (A, B, or C) indicating low, moderate, and high risk of bias, respectively.

Finally, for studies providing information on test-and-treatment strategies (Key Questions 3 and 4) we used a combination of items from the QUADAS-2 tool and the Cochrane risk-of-bias tool.³⁰

Data Synthesis

We summarized our findings according to the order of the Key Questions. Within each Key Question, results were organized on the basis of the populations assessed and clinical indications for clopidogrel use, index tests used, and outcomes assessed.

Meta-Analysis

We performed random-effects inverse-variance meta-analysis when at least three studies were available on sufficiently similar populations using the same test and assessing the same outcomes.³¹ Between-study heterogeneity was assessed on the basis of the Q statistic³² (considered statistically significant when its p-value [P_Q] was less than 0.1). Between-study inconsistency was assessed using the I^2 index.³³ Prior to the review, we decided not to combine studies of different phenotypic tests for platelet reactivity because they are based on different principles of measurement. Similarly, we decided not to combine trials providing information about effect modification due to heterogeneous populations enrolled in trials comparing different pairs of interventions (i.e., the magnitude and direction of effect modification by the tests of interest were likely to vary among different treatment comparisons).

In the absence of consensus on the correct genetic model, we assumed a dominant model for all minor alleles because this is the model used in previous CYP2C19 meta-analyses and because it allowed fullest use of the data. We performed sensitivity analyses assuming a recessive or additive model.

We used hazard or incidence rate ratios in our meta-analyses whenever available or extractable from the reviewed studies. When such statistics were not reported and could not be calculated, we used risk (proportion) ratios because they approximate the relative incidence rate.

For case-control studies, we used odds ratios because they are valid statistics for these designs and approximate the incidence rate ratio or risk ratio, depending on sampling methods.³⁴⁻³⁶ For parsimony, we refer to all these statistics as relative risks (RRs).

Other Analyses

To assess the impact of study-level characteristics on estimates of the effect size, we used univariable random-effects meta-regression.³³ Subgroup and meta-regression analyses were performed for factors reported at the group level. Predefined subgroups of interest were those defined by race or ethnicity, sex, specific assay used, and clinical setting of test use (e.g., short-term administration of clopidogrel during treatment of acute cardiac events or PCI vs. chronic clopidogrel use). We also explored temporal trends in the reported effect sizes using meta-regression with year of publication as the covariate. We used Egger's regression-based test³⁷ to assess the presence of small-study effects.^{37,38}

Strength of the Body of Evidence

We graded the strength of the body of evidence for the Key Questions following the Methods Guide and recently updated recommendations for the Evidence-based Practice Center (EPC) Program.^{23,39} Briefly, the strength of evidence was graded as high, moderate, low, or insufficient on the basis of four dimensions: risk of bias (described above), consistency, directness, and precision. We assessed the consistency of the data as either “no inconsistency” or “inconsistency present” (or “not applicable” if only one study). We also assessed the sparseness of the evidence. We considered evidence to be sparse if it was from only one study with a small sample size. Strength ratings were assigned on the basis of our level of confidence that the evidence reflected the true effect for the major comparisons of interest.

Assessing Applicability

We assessed applicability of the study findings on the basis of the individual study eligibility criteria and baseline characteristics of the included populations, following recommendations in the Methods Guide and recently updated recommendations for the EPC Program.^{23,39} We did not assess the applicability of studies regarding the analytic validity of the tests of interest (Key Questions 1a and 2a) because technical test performance does not directly inform medical decisions, although it is a prerequisite for the clinical use of tests.⁴⁰

Results

The literature search yielded 10,475 citations (10,374 from electronic databases, 77 from scientific information packets, and 24 from hand-searching). Of these, 1,419 articles were reviewed in full text. After full-text review, 326 were judged to have met the inclusion criteria for one or more Key Questions. We summarized the findings of the report according to the order of the Key Questions. Within each Key Question, results are organized for each appropriate subgroup on the basis of the populations assessed and clinical indications for clopidogrel use, index tests assessed, and outcomes assessed.

Key Question 1a: Analytic Validity of Tests for Genotyping CYP2C19 Variants

Eligible Studies

We identified 11 studies reporting information on the analytic validity of genotyping methods for detecting CYP2C19 variants. We also reviewed four FDA 510(k) summaries on genetic testing.

Summary of Findings

Primary studies generally indicated excellent test-retest reliability and interassay agreement. FDA 510(k) summaries did not report analyses on samples from populations and genes of interest to our review. However, the documents provided further evidence that genotyping methods have high test-retest reliability and indicated that rates of interassay agreement were high.

Primary studies reported limited information on the methods used to assess analytic validity. This probably reflects the fact that the primary focus of all included publications was not the tests' analytic validity but rather their clinical utility. Generally, studies provided adequate information on the genotyping methods used. However, they provided little information on the use of positive or negative control samples, the handling of uninterpretable results, and the test detection limits. Four studies reported information on the reproducibility of genotyping across different genotyping methods, but no study assessed reproducibility across operators. No study was conducted as part of an interlaboratory standardization project.

Key Question 1b: Predictive Value of Genetic Testing for CYP2C19 Variants

Eligible Studies

The 106 studies addressing Key Question 1b were described in 98 publications, 8 of which described 2 studies each. The vast majority of studies (100, or 94%) were of patients with ischemic heart disease. Three studies

enrolled patients with different forms of vascular disease (coronary, cerebrovascular, or peripheral arterial); one enrolled patients with cerebrovascular disease; one enrolled a mixed population of patients with manifest atherothrombotic disease along with asymptomatic patients at high risk for atherothrombotic disease; and one enrolled patients with atrial fibrillation who were not candidates for vitamin K antagonist therapy.

The 106 studies had intermediate to large sample sizes: median number of enrolled individuals = 277, 25th percentile = 98, 75th percentile = 802, minimum = 30, maximum = 5,148. They were conducted recently (median year of start of enrollment, 2006, with 75% beginning enrollment after 2004), reflecting the relatively recent widespread availability of genetic testing for CYP2C19 variants. The majority of enrolled patients were men and the median age was 64 years. Across studies, the median proportions of patients with dyslipidemia and hypertension were both over 60 percent. The median proportions of patients with diabetes mellitus and patients who smoked were both greater than 25 percent. Overall, 94 percent of studies had a longitudinal (cohort) design; 11 of these were genetic substudies consisting of prospectively followed clopidogrel-treated groups from randomized trials.^b

Overall, studies had moderate risk of bias: 12 studies were rated as quality A, 88 studies were rated as quality B, and 6 were rated as quality C. We caution that this aggregate risk-of-bias rating can be misleading, especially in the presence of poor reporting, because it assigns the same weight to all items.

Summary of Findings

The two most common genotyping methods were TaqMan genotyping (44 studies; 42%) and polymerase chain reaction with restriction fragment length polymorphism analysis (PCR-RFLP; 13 studies; 12%). In the majority of cases, analyses were conducted on genetic material isolated from blood (92 studies; 87%). Among the 56 studies that reported the genotyping success rate, the median was 100 percent (minimum = 74%; maximum = 100%). Violations of Hardy-Weinberg equilibrium (on the basis of an exact goodness-of-fit test) were not more common than would be expected by chance.

Below, findings are presented for studies providing information on the ability of genetic testing for CYP2C19 variants to predict clinical outcomes (57 studies) or platelet reactivity (74 studies) during followup.

Clinical Outcomes

Several clinical outcomes of interest were reported: all-cause mortality, cardiac mortality, acute coronary syndromes, stent thrombosis, stroke, major adverse cardiovascular events (MACE), bleeding events, and need for revascularization. Under a dominant genetic model, loss-of function CYP2C19 alleles were statistically significantly associated with stent thrombosis (RR=1.52; 95% confidence interval [CI], 1.17 to 1.97); cardiovascular mortality (RR=1.98; 95% CI, 1.13 to 3.46); and MACE (RR=1.20; 95% CI, 1.04 to 1.39). Under a dominant genetic model, gain-of-function alleles (CYP2C19*17) were statistically significantly associated with reduced risk of MACE (RR=0.82; 95% CI, 0.74 to 0.92). Studies on the predictive value of CYP2C19 variants were judged to have moderate risk of bias. There was some indication of systematic differences between larger and smaller studies (loss-of-function alleles: $p < 0.001$, $p = 0.002$, and $p = 0.049$ for stent thrombosis, cardiovascular mortality, and MACE, respectively; gain-of-function alleles: $p = 0.046$ for bleeding events). There was also substantial risk of selective outcome reporting for outcomes other than MACE.

Sensitivity analyses using alternative genetic models (recessive and additive for the variant alleles) were based on a minority of studies and possible only for the association of loss-of-function alleles with MACE and stent thrombosis. Generally, these analyses were congruent with analyses using a dominant model because they also indicated significant association between loss-of-function alleles and adverse clinical outcomes. Effect sizes using both the recessive and additive models were larger than those under the dominant model.

Intermediate Outcome: Platelet Reactivity

The intermediate outcome of platelet reactivity was reported either as a continuous variable (in 61 studies) or according to a threshold of reactivity (e.g., high vs. low; in 39 studies). The most common assays for assessing reactivity were light-transmission aggregometry (LTA), the VerifyNow P2Y12 assay, and the VASP assay. For platelet reactivity as a continuous outcome, the mean or median reactivity was generally higher among clopidogrel-treated patients with one or two loss-of-function alleles than those with no loss-of-function alleles. For platelet reactivity as a categorical outcome, studies generally showed that platelet reactivity above the threshold used (or in higher quantiles compared with lower quantiles of reactivity) was more common in clopidogrel-treated patients with one or two

^b When appropriate, modification of the relative treatment effect by genotype status has been considered under Key Questions 3 and 4 of this report.

loss-of-function alleles than those with no loss-of-function alleles. Only a minority of studies reported analyses under different genetic models, and it was often impossible to reconstruct such analyses from the reported data. Because of the extensive differences among studies of either type of reactivity outcome and the often incomplete reporting of numerical information, we did not perform meta-analyses for studies using reactivity as the outcome of interest.

Key Question 1c: Factors Affecting the Predictive Value of Genetic Testing for CYP2C19 Variants

We reviewed studies to identify any evidence that patient- or system-level factors or test characteristics could modify the prognostic ability of genetic testing for CYP2C19 variants. We considered both within-study information (e.g., studies in which the predictive effect of phenotypic testing was evaluated in two or more patient subgroups) and information across studies (through metaregression analyses on study-level factors).

Effect Modification Within Studies

Twenty studies reported information on modification of the prognostic effect of the genetic test by various factors; 5 studies assessed more than two potential effect modifiers. Only two of the effects assessed were statistically significant, each in a single study: a multiplicative interaction between the *2 and *3 CYP2C19 alleles on on-clopidogrel platelet reactivity and an interaction between clinical presentation and loss-of-function CYP2C19 allele carriership (comparing a cohort of patients with myocardial infarction, ischemic stroke, or peripheral arterial disease vs. asymptomatic patients at high risk for atherothrombotic events). The following nonstatistically significant comparisons were also reported: effect modification by proton-pump inhibitors (five studies) and by ancestry (three studies), gene-gene interactions (four studies). The following modifiers were also evaluated in one study each: indication for clopidogrel use (acute coronary syndromes vs. stable angina), whether patients were clopidogrel pretreated or naïve upon study entry, whether patients required a loading dose or not (because they were on chronic clopidogrel therapy), the duration of clopidogrel therapy, smoking status (number of cigarettes per day), body-mass index (≥ 25 kg/m² vs. < 25 kg/m²), stent type (bare metal vs. drug eluting), myocardial infarction subtype (ST elevation or non-ST elevation), history of PCI (yes vs. no), interactions with a large set of clinical and procedural factors, administration of polyunsaturated fatty acids, whether patients were on calcium channel blockers or their combination with proton pump inhibitors. All results were nonsignificant. Overall,

the reported findings do not provide sufficient evidence to support or exclude a differential effect of CYP2C19 variants across any of the factors assessed in the studies we reviewed. The statistically significant findings should not be overinterpreted, given the number of comparisons performed and the potential for selective reporting across studies.

Effect Modification Across Studies

Potential modifiers of the predictive effect of genetic testing for CYP2C19 that were assessed across studies using subgroup and metaregression analysis were disease subtype (acute coronary syndromes vs. mixed coronary artery disease populations), setting of care (PCI vs. other), race or ethnicity (white vs. East Asian), duration of followup (≤ 30 days vs. > 30 days), and year when enrollment was started (continuous variable). Metaregression analyses, both for stent thrombosis and MACE, suggested that the effect of loss-of-function alleles may be more extreme among individuals of East Asian ethnicity; however, this finding needs to be interpreted with caution, given the relatively small number of publications reporting on individuals of East Asian ethnicity and the potential for confounding by other factors that differ between studies conducted in populations of different ethnicities.

Effect Modification Summary

In general, considering both analyses within studies and across studies, there is insufficient information to support or exclude the presence of substantial modification of the prognostic effect of CYP2C19 variants by any of the investigated factors because most modifiers were evaluated in a single study (in the majority of cases producing nonsignificant results) and because metaregression analyses (nonsignificant for all but one of the factors explored) may be confounded by study-level characteristics.

Key Question 2a: Analytic Validity of Tests for On-Clopidogrel Platelet Reactivity

Eligible Studies

We identified 104 studies reporting information on the analytic validity of assays for measuring platelet reactivity. We also reviewed 20 FDA 510(k) summaries on phenotypic testing assays. All published studies enrolled patients with ischemic cardiovascular disease. The six most commonly assessed assays (with some studies assessing more than one) were LTA, the VerifyNow P2Y12 assay, the VASP assay, the Multiplate analyzer, the PFA device, and thromboelastography. We summarized the reported

information regarding analytic performance, interassay agreement, test reliability and assay variation, and correlations between assays applied to the same sample (by far the most common metric reported). No other aspect of analytic validity was evaluated in the studies.

Summary of Findings

Overall there appeared to be low to moderate agreement between assays. Agreement was generally greater between measurements obtained with the same assay using different agonist concentrations than between different assays.

In the 12 studies providing information on analytic performance, analytic sensitivities ranged between 0.35 and 1.00, and analytic specificities ranged between 0.42 and 0.95. In studies reporting results across multiple cutoff values, a tradeoff between sensitivity and specificity was apparent, as expected. Overall, these results indicate poor agreement in sample classification (e.g., high vs. low reactivity) when one of the two tests compared was considered a gold standard. However, the evidence suggests that no test is a gold standard (i.e., all have measurement error).

Forty studies provided information on interassay agreement. Overall, disagreements were relatively common between measurements obtained by different assays or by using different agonist concentrations within the same assays.

Forty-three studies reported information on assay variability, although more than 90 percent did not describe the methods used in their assessment. One study used the intraclass correlation coefficient for repeat measurements to assess the reliability of measurements using LTA, the VASP assay, the Multiplate analyzer, and the INNOVANCE assay. Variability or coefficient-of-variation results were less than 10 percent in all but two studies. These results need to be interpreted with caution, given the poor reporting of study methods and the fact that multiple studies were published by a limited number of investigative teams. (In most cases, we could not ascertain whether the studied populations overlapped.)

Of the 56 studies reporting correlation values, only 1 used Lin's concordance correlation coefficient (an appropriate metric), reporting a high correlation ($\rho = 0.97$) between observed and estimated platelet inhibition for the VerifyNow assay. The remaining studies used inappropriate metrics (e.g., Spearman or Pearson correlation coefficients or linear regression, which in the simple bivariate case of two measurements is equivalent to the Pearson correlation) or did not report the calculation method used.^{41,42} The results indicated that the association between measurements obtained using different methods is

relatively poor. However, given the inappropriateness of the methods used to assess agreement, even high correlation values would not be considered indicative of good agreement.

None of the 20 FDA 510(k) summaries on phenotypic tests of platelet reactivity reported relevant analyses that met our study selection criteria: either no data were reported or the population or agonist used in testing was not of interest, the analytic validity was not reported for clopidogrel-treated patients, or the sample size was less than 10.

Key Question 2b: Predictive Ability of Phenotypic Testing for Platelet Reactivity

Eligible Studies

Of the 128 studies addressing Key Question 2b, the vast majority (122 studies) were of patients with ischemic heart disease. Four studies enrolled patients with cerebrovascular disease; one study enrolled patients with peripheral arterial disease; and one study enrolled a mix of patients with ischemic heart disease, cerebrovascular disease, and peripheral arterial disease. Studies reported information on a variety of assays for measuring platelet reactivity. Table A summarizes information on the patient populations and outcomes assessed in the 128 studies. Detailed information on each test is presented separately under the discussion of individual assays. For parsimony, we discuss below only results for the test-outcome combinations for which the strength of the overall body of evidence was judged to be at least "low." (Please refer to the Methods section for details on our approach to rating the strength of evidence.) Complete results are available in the full report. The strength of evidence for all patient populations other than those with ischemic heart disease was judged to be insufficient owing to the very few studies and small sample sizes.

Overall, studies were considered to have a moderate risk of bias. All studies used a longitudinal design (not case control, in keeping with our inclusion criteria), and no studies had substantial loss to followup. Inappropriate exclusions were uncommon, but information on blinding was often not reported (particularly for the index test) or not used. It was often unclear whether analyses, including the definitions of increased platelet reactivity and the outcomes assessed, had been prespecified and reported in full. Using the cutoff values based on the number of adequately addressed risk-of-bias items, 36 studies were rated as quality A, 80 studies were rated as quality B, and 12 were rated as quality C. A more detailed discussion of risk of bias, focusing on the individual items assessed, is presented in the full report.

Table A. Populations, outcomes, and strength of evidence in studies for Key Question 2b, according to test used

Test Used (Total Number of Studies; Studies by Patient Population)	All-Cause Death	CV Death	ACS	ST	Stroke	Bleeding	MACE	Other Clinical Outcomes	Platelet Reactivity
LTA (total = 55; IHD = 53; PAD = 1; IHD, CVD, PAD = 1)	IHD = 13 [low]	IHD = 9 [low]	IHD = 18 [low]	IHD = 19 [low]	IHD = 12 [low]	IHD = 7 [insufficient]	IHD = 37 [low] IHD, CVD, PAD = 1 [insufficient]	IHD = 8 [insufficient]	IHD = 11 [insufficient] PAD = 1 [insufficient]
VerifyNow P2Y12 (total = 38; IHD = 35; CVD = 3)	IHD = 10 [low]	IHD = 7 [moderate]	IHD = 19 [low]	IHD = 15 [low]	IHD = 8 [insufficient]	IHD = 12 [low]	IHD = 24 [moderate]	IHD = 7 [insufficient] CVD = 3 [insufficient]	IHD = 4 [insufficient]
VASP (total = 19; IHD = 18; IHD, CVD, PAD = 1)	IHD = 4 [insufficient]	IHD = 6 [insufficient]	IHD = 6 [low]	IHD = 10 [low]	IHD = 1 [insufficient]	IHD = 1 [insufficient]	IHD = 8 [low] IHD, CVD, PAD = 1 [insufficient]	IHD = 4 [insufficient]	IHD = 7 [insufficient]
Multiplate analyzer (total = 18; IHD = 17; CVD = 1)	IHD = 6 [insufficient]	IHD = 5 [insufficient]	IHD = 9 [insufficient]	IHD = 10 [insufficient]	IHD = 3 [insufficient]	IHD = 9 [insufficient]	IHD = 13 [insufficient] CVD = 1 [insufficient]	IHD = 6 [insufficient]	IHD = 2 [insufficient]
TEG (total = 6; IHD = 6)	IHD = 2 [insufficient]	No studies	IHD = 2 [insufficient]	IHD = 1 [insufficient]	IHD = 1 [insufficient]	IHD = 3 [insufficient]	IHD = 4 [insufficient]	No studies	No studies
PFA-100 (total = 11; IHD = 10; IHD, CVD, PAD = 1)	IHD = 2 [insufficient]	IHD = 2 [insufficient]	IHD = 5 [insufficient]	IHD = 3 [insufficient]	IHD = 1 [insufficient]	IHD = 1 [insufficient]	IHD = 9 [low] IHD, CVD, PAD = 1 [insufficient]	IHD = 2 [insufficient]	IHD = 1 [insufficient]
Other (total = 9; IHD = 9)	IHD = 3 [insufficient]	No studies	IHD = 3 [insufficient]	IHD = 1 [insufficient]	IHD = 1 [insufficient]	IHD = 2 [insufficient]	IHD = 6 [insufficient]	IHD = 3 [insufficient]	IHD = 3 [insufficient]

Note: Numbers indicate the number of available studies for each test-outcome combination in the population specified. Studies could have involved more than 1 combination. Text in brackets reflects our assessment of the strength of evidence for each test-outcome association.

ACS = acute coronary syndromes; CV = cardiovascular; CVD = cerebrovascular disease; IHD = ischemic heart disease; LTA = light-transmission aggregometry; MACE = major adverse cardiovascular events; PAD = peripheral arterial disease; PFA = Platelet Function Analyzer; ST = stent thrombosis; TEG = thromboelastography; VASP = vasodilator-stimulated phosphoprotein.

LTA in Ischemic Heart Disease

Fifty-three studies included patients with ischemic heart disease and reported information on the predictive value of LTA for clinical outcomes (47 studies) and platelet reactivity (11 studies). Four studies reported both clinical and intermediate outcomes. Thirty-eight of the 53 studies enrolled patients with chronic stable coronary artery disease, 12 enrolled patients with acute coronary syndromes, and 3 enrolled mixed populations with chronic and acute presentations. Most studies used ADP as the agonist to assess reactivity but a few used ADP in combination with arachidonic acid (AA) to assess the response to both clopidogrel and aspirin. The strength of evidence for the prognostic effect of high platelet reactivity as measured by LTA on the following outcomes was considered low on the basis of clinical heterogeneity, variation in the metrics and cutoffs used to define reactivity, and imprecision of the study-level estimates of effect.

All-Cause Mortality (13 Studies; 12 ADP, 1 ADP + AA)

Studies did not suggest an association between increased platelet reactivity as measured by LTA and increased all-cause mortality in patients with ischemic heart disease.

Cardiovascular Mortality (9 Studies; 8 ADP, 1 ADP + AA)

Studies suggested an association between increased platelet reactivity as measured by LTA and cardiovascular mortality in patients with ischemic heart disease.

Acute Coronary Syndromes (18 Studies; 17 ADP, 1 ADP + AA)

Overall, results provided some evidence of an association between increased platelet reactivity as measured by LTA and increased risk of acute coronary syndromes in patients with ischemic heart disease.

Stent Thrombosis (19 Studies; 17 ADP, 2 ADP + AA)

Nineteen studies reported information on the ability of platelet reactivity as measured by LTA to predict stent thrombosis. Three publications reported data from the same population. Taken together, the studies suggested an association between increased platelet reactivity and increased risk of stent thrombosis in patients with ischemic heart disease.

MACE (37 Studies; 35 ADP, 2 ADP + AA)

Three of the 37 studies reported data from the same population. We evaluated data for the longest followup time available. All studies used ADP as the agonist to measure platelet reactivity; two studies used ADP in combination with AA to assess the response to both clopidogrel and aspirin. The majority of reviewed studies suggested an association between increased platelet

reactivity as measured by LTA and increased risk of MACE.

Stroke (12 Studies; 11 ADP, 1 ADP + AA)

The 12 reviewed studies did not suggest an association between increased platelet reactivity as measured by LTA and increased stroke in patients with ischemic heart disease.

VerifyNow P2Y12 in Ischemic Heart Disease

Thirty-five studies included patients with ischemic heart disease and reported information on the predictive value of the VerifyNow P2Y12 assay. Of these, 33 assessed the value of the test for predicting clinical outcomes, and 3 for predicting platelet reactivity during followup. Two studies reported both clinical and platelet reactivity outcomes. Of the 35 studies, 21 enrolled patients with chronic stable coronary artery disease, 12 enrolled patients with acute coronary syndromes, and 2 enrolled a mixed population with chronic and acute presentations.

All-Cause Mortality (10 Studies; 9 ADP, 1 ADP + AA)

A meta-analysis of three studies that used ADP as the agonist and defined high platelet reactivity on the basis of platelet reactivity units found a summary RR of 1.21 (95% CI, 0.83 to 1.77; $p=0.313$), indicating a nonsignificant association between high platelet reactivity and all-cause mortality. There was little evidence of between-study heterogeneity ($P_O=0.902$; $I^2=0\%$). Meta-analysis was not performed for the four other studies, which either used percent platelet inhibition to define reactivity or defined reactivity using a different cutoff.

Cardiovascular Mortality (7 Studies, All ADP)

A meta-analysis of the four studies that used cutoff values based on platelet reactivity units found a summary RR of 2.50 (95% CI, 1.28 to 4.87; $p=0.007$), indicating a significant association between high platelet reactivity and cardiovascular mortality. There was little evidence of between-study heterogeneity ($P_O=0.527$; $I^2=0\%$). The three studies not included in the meta-analysis did not report a significant association between higher platelet reactivity and increased cardiovascular mortality.

Acute Coronary Syndromes (19 Studies; 16 ADP, 1 ADP + AA)

Nineteen studies reported information on the ability of the VerifyNow P2Y12 assay to predict myocardial infarction in patients receiving clopidogrel-based treatment. Taken together, the studies suggested an association between increased platelet reactivity as measured by VerifyNow and increased rates of both periprocedural and nonperiprocedural acute coronary syndromes in patients with ischemic heart disease. However, the strength of evidence for this association was considered low on the

basis of variability in the metrics and thresholds used to define reactivity and heterogeneity of the included patient populations.

Stent Thrombosis (15 Studies; All ADP)

Fifteen studies reported information on the ability of the VerifyNow P2Y12 assay to predict stent thrombosis in patients receiving clopidogrel-based treatment. Of these, 11 did not report statistically significant results and produced relatively wide CIs, indicating substantial uncertainty around estimates of the RR; 4 studies reported statistically significant associations between high reactivity with risk of stent thrombosis. Because of heterogeneity in the metrics used to define platelet reactivity, meta-analysis was possible only for six studies that used the same metrics and cutoffs for reactivity. The summary RR was 1.67 (95% CI, 0.80 to 3.47; $p=0.172$). There was some evidence of between-study heterogeneity ($P_{\phi}=0.159$; $I^2=37\%$). Considering all studies, there was weak evidence to support an association between increased platelet reactivity as measured by VerifyNow and stent thrombosis in patients with ischemic heart disease.

MACE (24 Studies; 23 ADP, 1 ADP + AA)

One study that used both ADP and AA to identify a population of responders to both clopidogrel and aspirin reported significantly higher odds of MACE in those who were clopidogrel nonresponders (irrespective of aspirin response status) compared with responders. A meta-analysis was done of 13 of the 23 remaining studies that enrolled nonoverlapping patient populations and used cutoff values for platelet reactivity based on platelet reactivity units. The summary RR was 2.48 (95% CI, 1.86 to 3.32; $p<0.001$) and there was evidence of moderate heterogeneity ($P_{\phi}=0.045$; $I^2=44\%$) and statistically significant small-study effects. Ten studies were not included in the meta-analysis due to differing definitions of reactivity, poor reporting, and patient overlap. Specifically, five studies used percentage of platelet inhibition to define platelet reactivity; two used a different cutoff to define reactivity; two studies did not provide adequate data for inclusion; and one overlapped with another publication that had larger sample size. Among the five studies that used percentage of platelet inhibition to define platelet reactivity, three studies reported significantly higher rates of MACE and one study reported nonsignificantly higher rates of MACE at 6 months or 1 year in those with a low response to clopidogrel. In contrast, one study reported lower rates of MACE at 30 days in those with a low response to clopidogrel.

Bleeding Events (12 Studies; All ADP)

A meta-analysis of six nonoverlapping studies with similar reactivity cutoffs found no significant difference by reactivity status for either all bleeding events (4 studies; RR=1.09; 95% CI, 0.88 to 1.37; $p=0.421$) with little evidence of heterogeneity ($P_{\phi}=0.738$; $I^2=0\%$) or severe bleeding events (4 studies; RR=0.85; 95% CI, 0.32 to 2.25; $p=0.738$) with evidence of moderate heterogeneity ($P_{\phi}=0.074$; $I^2=57\%$). Four other studies with different cutoff values reported lower but not statistically significantly different rates of major and minor bleeding for patients with a low response to clopidogrel (compared with responders), and a fifth study did not report any bleeding events.

VASP Assay With Flow Cytometry in Ischemic Heart Disease

Eighteen studies included patients with ischemic heart disease and reported information on the predictive value of the VASP assay. Of these, 13 assessed the value of the test for predicting clinical outcomes, 6 assessed the value for predicting platelet reactivity during followup, and 1 reported both clinical and platelet reactivity outcome. Eight studies enrolled patients with acute coronary syndromes, five enrolled patients with chronic stable coronary artery disease, and five enrolled mixed populations with chronic and acute presentations.

Acute Coronary Syndromes (6 Studies; All ADP)

One study reported that no events were observed regardless of platelet reactivity status and thus was not included in meta-analysis. Of the remaining five studies, two were nonoverlapping. The other three had overlapping study populations and enrollment periods, so in meta-analysis we used data from the publication reporting the largest number of events. A meta-analysis of the three studies, all of which used cutoff values based on the platelet reactivity index, found a summary RR of 1.47 (95% CI, 0.77 to 2.794; $p=0.246$). There was little evidence of between-study heterogeneity ($P_{\phi}=0.372$; $I^2=0\%$).

Stent Thrombosis (10 Studies; All ADP)

Two studies reported that no events were observed regardless of platelet reactivity status and thus were not included in meta-analysis. Of the remaining eight studies, three were nonoverlapping. The other five had overlapping study populations and enrollment periods, so in meta-analysis we used data from the publication reporting the largest number of events. A meta-analysis of the four studies found a summary RR of 3.37 (95% CI, 1.59 to 7.1; $p=0.015$), indicating a statistically significant association between high platelet reactivity and stent thrombosis.

There was no evidence of between-study heterogeneity ($P_0=0.487$; $I^2=0\%$).

MACE (8 Studies; All ADP)

Two publications involved overlapping study populations; in meta-analysis we included data from the publication reporting the largest total number of cardiovascular events. A meta-analysis of the six nonoverlapping studies found a summary RR of 2.57 (95% CI, 1.21 to 5.47; $p=0.015$), indicating a statistically significant association between high platelet reactivity measured by the VASP assay and MACE. There was evidence of moderate between-study heterogeneity ($P_0=0.044$; $I^2=56\%$).

Multiplate Analyzer, Thromboelastography, and PFA-100 in Ischemic Heart Disease

The strength of evidence was insufficient for all outcomes for these three tests.

Comparative Studies of Test Performance Among Platelet Reactivity Assays

Twelve studies reported extractable information on clinical outcomes for at least two of the assessed tests. We focused on outcomes that were addressed by at least 3 comparative studies: major adverse cardiovascular events (composite outcome, 10 studies) and stent thrombosis (4 studies). The data could not be quantitatively synthesized because the studies involved several assays being applied to the same patient population, in which case results are likely to be correlated because the population is shared and assays done on samples of the same blood will yield correlated, if not identical, results.

MACE (Comparative Studies)

Ten studies reported comparative information regarding the ability of assays to predict MACE. The most commonly used test was LTA, which was compared with various tests (most often thromboelastography and VerifyNow P2Y12). Overall, point estimates were similar between alternative test methods within each study and CIs were overlapping, suggesting that the predictive ability of the compared tests is fairly similar.

Stent Thrombosis (Comparative Studies)

Four studies reported comparative information regarding the ability of assays to predict stent thrombosis for patients undergoing PCI with stent implantation. LTA was used in three studies, and the VerifyNow P2Y12 and PFA-100 assays were each used in two studies. Point estimates were variable within each study. However, CIs were extremely wide and overlapping, suggesting that there is substantial uncertainty regarding the relative predictive ability of the compared tests for stent thrombosis and that there is

substantial uncertainty regarding comparative test performance for this outcome.

Comparative Studies of Test Performance of Genetic Testing for CYP2C19 Variants and Phenotypic Testing for Platelet Reactivity

Four studies reported information on the prognostic value of genetic and phenotypic tests for MACE and three for stent thrombosis. For each of the four studies, we plotted the points corresponding to each assay's sensitivity and specificity in the receiver operating characteristic space. Points were often close to the chance diagonal, indicating that test performance was generally poor. However, the paucity of data did not allow firm conclusions.

Key Question 2c: Factors Affecting the Predictive Value of Phenotypic Testing for Platelet Reactivity

We reviewed studies to identify any evidence that patient- or system-level factors or test characteristics could modify the predictive ability of phenotypic testing for platelet reactivity. As for Key Question 1c, we considered both within-study information (e.g., studies where the predictive effect of phenotypic testing was evaluated in two or more patient subgroups) and information across studies (through metaregression analyses on study-level factors).

Effect Modification Within Studies

In total, seven studies reported information on effect modification of the predictive effect of platelet reactivity. All studies reported information on clinical outcomes. Only a small subset of the eligible studies provided information adequate to statistically assess effect modification, and selective reporting was highly likely. Studies assessed the following factors as potential modifiers: the use of glycoprotein IIb/IIIa inhibitors as an adjunct treatment for PCI (two studies), diabetes mellitus (two studies), and chronic kidney disease (one study). Two studies used the VASP assay to assess platelet reactivity; two used the VerifyNow P2Y12 assay (one of which also used the VerifyNow ASA, which uses arachidonic acid as the agonist to measure "aspirin resistance"); and one used LTA (with ADP as the agonist). Statistically significant interaction effects were reported only in the study that assessed whether coexisting chronic kidney disease in patients with coronary artery disease modified the predictive value of the VASP assay. The study found that high on-clopidogrel platelet reactivity had statistically significantly greater effects on several clinical outcomes (all-cause mortality, cardiac death, and a composite outcome of all-cause mortality, myocardial infarction, or

target-lesion revascularization) in patients with chronic kidney disease than in those without chronic kidney disease.

Effect Modification Across Studies

In analyses across studies (metaregression) the following factors did not statistically significantly modify the prognostic value of the VerifyNow P2Y12 assay on MACE (the only test-outcome combination with 10 or more available studies): disease subtype (acute coronary syndromes vs. coronary artery disease); duration of followup (≤ 30 days vs. >30 days); and year when enrollment was started (continuous variable).

Effect Modification Summary

In general, information on effect modification was limited, both within and across studies. Few studies reported information on the same potential effect modifiers, results were imprecise, and selective reporting was highly likely. Information across studies was also limited by the number of available studies on each test and outcome of interest. It is unclear whether the predictive effect of phenotypic testing differs across patient subgroups.

Key Question 3a: Comparative Effectiveness of Alternative Test-and Treat Strategies

We grouped the studies we identified for this Key Question into three categories:

1. *Randomized trials of test-and-treat strategies:* These studies randomize patients to alternative management strategies, at least one of which is based on a test of interest. Patients are then followed up for intermediate or clinical outcomes.
2. *Randomized treatment trials that evaluate treatment-effect modification:* These are randomized studies in which patients in all groups undergo the test of interest at baseline. Treatment assignment is based on randomization and thus is independent of test results. Because these studies include both test-positive and test-negative patients in each treatment arm, they can be used to assess test result \times treatment interactions.
3. *Randomized trials with test-based selection:* These studies select patients on the basis of baseline test results and then randomize them into non-test-based treatment groups. When properly randomized and conducted, these studies can provide unconfounded estimates of the treatment effect conditional on a particular test result.

Genetic Testing for CYP2C19 Variants

Studies of Test-and-Treat Strategies

We identified a single-center pilot study with low risk of bias that compared a strategy of testing for CYP2C19 variants versus no testing to guide treatment decisionmaking in a predominantly white population (95%). The study randomized 200 adult patients undergoing PCI for the treatment of non-ST-segment-elevation acute coronary syndrome or stable CAD to a treatment group guided by CYP2C19 genotyping or a control group with no testing.

Clinical Outcomes

The study reported information on a composite outcome of cardiovascular death, nonfatal myocardial infarction, readmission to hospital, and stent thrombosis. Twenty-three (25%) of 91 patients assigned to the rapid genotyping group were CYP2C19*2 carriers (4 were homozygotes); 23 (24%) of 96 in the standard therapy group were CYP2C19 *2 carriers (3 were homozygotes). No clinical adverse ischemic outcomes were observed in either group at 7 or 20 days of followup.

Intermediate Outcome: Platelet Reactivity During Followup

Intermediate outcomes were assessed with the VerifyNow P2Y12 assay. The primary study endpoint was the proportion of CYP2C19*2 carriers with a P2Y12 reactivity unit (PRU) value of more than 234 after 1 week of dual antiplatelet therapy. The results indicated that platelet reactivity at the last followup assessment was lower in the groups that received test-based treatment than in those that did not undergo testing.

Randomized Trials Reporting Information on Treatment-Effect Modification by CYP2C19 Genotype Status

We identified 13 publications (reporting on 12 study populations) describing randomized controlled trials that provide information on effect modification by CYP2C19 variants. Six studies provided information on clinical outcomes, five on intermediate outcomes (platelet reactivity during followup), and one on both types of outcome.

Clinical Outcomes

Six studies (reported in seven publications^c) provided clinical outcome information. Five of the six studies were large ($>1,000$ participants), multicenter, randomized trials of clopidogrel-based treatment versus alternative treatments and had at least one outcome event. The sixth

^c Five publications reported information on a single population each and one publication reported information on two independent populations.

study, a smaller, single-center trial with a followup of 30 days, reported that no clinical outcomes of interest were observed. Studies used robust methods for randomization and allocation concealment.

The five larger studies were the CURE (Clopidogrel in Unstable Angina to Prevent Recurrent Events) trial, which included patients with non–ST-elevation acute coronary syndromes; the PLATO (Platelet inhibition and patient Outcomes) trial, which involved patients with ST-elevation or non–ST-elevation acute coronary syndromes; the TRITON-TIMI 38 (Trial to Assess Improvement in Therapeutic Outcomes by Optimizing Platelet Inhibition with Prasugrel–Thrombolysis in Myocardial Infarction 38), which included those with moderate- to high-risk acute coronary syndromes who were undergoing PCI; the CHARISMA (Clopidogrel for High Atherothrombotic Risk and Ischemic Stabilization, Management and Avoidance) trial, which included a mixed population of patients with manifest thrombotic disease (coronary, cerebrovascular, and peripheral arterial disease) along with individuals at high risk for developing atherothrombotic disease; and ACTIVE A, which enrolled patients with atrial fibrillation who were not candidates for vitamin K antagonist therapy. CURE, CHARISMA, and ACTIVE A compared aspirin plus clopidogrel (at standard doses) with aspirin monotherapy, TRITON-TIMI 38 compared aspirin plus clopidogrel versus aspirin plus prasugrel, and the PLATO trial compared aspirin plus clopidogrel versus aspirin plus ticagrelor. All trials were designed and powered to detect the main effect of antiplatelet therapy but were not specifically powered to detect heterogeneity of treatment effects and typically included only a subsample of the overall trial population.

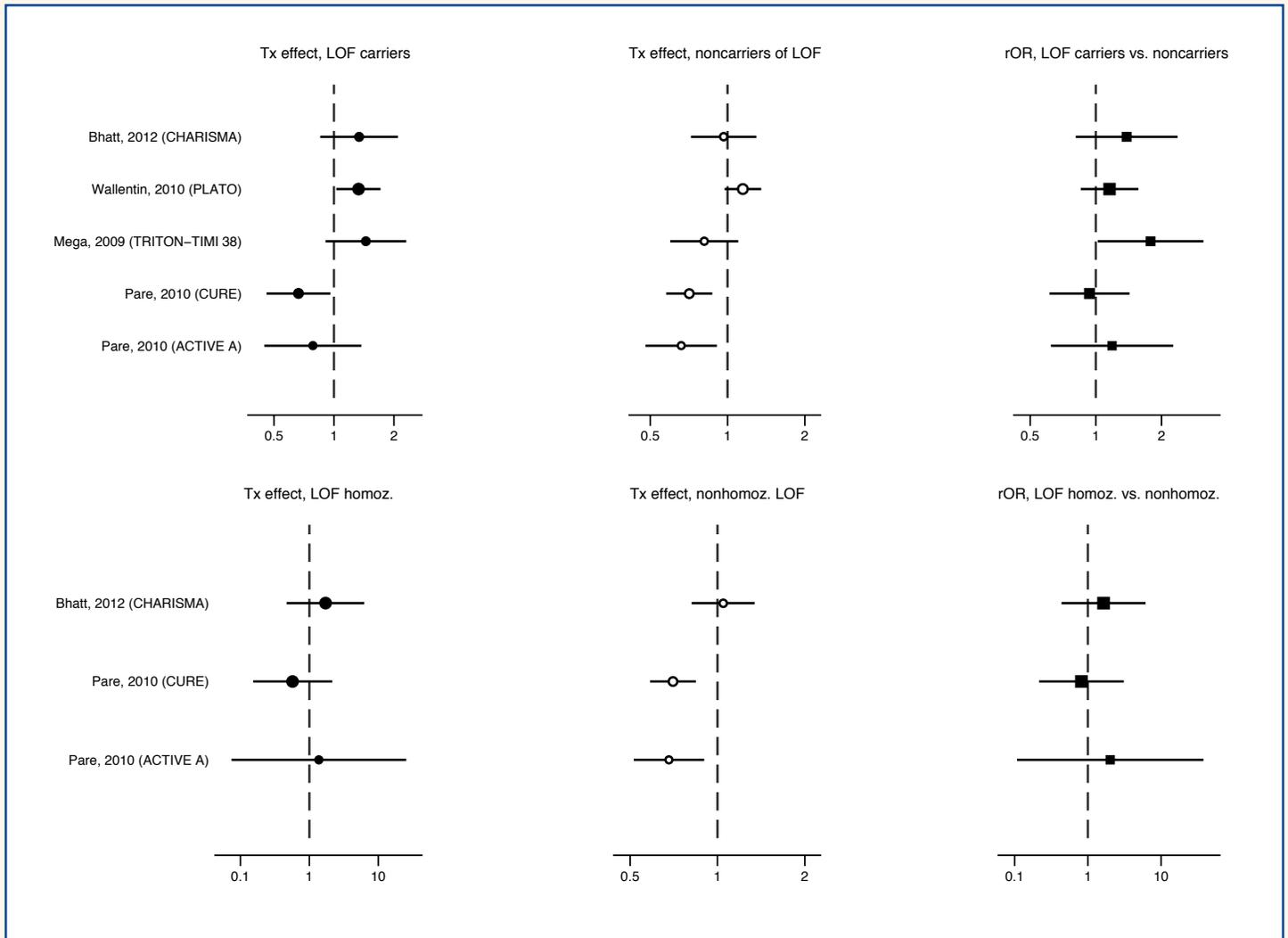
The CURE, PLATO, CHARISMA, and ACTIVE A trials did not find statistically significant effect modification by CYP2C19 genotype for any of their efficacy outcomes.

The genetic substudy of TRITON-TIMI 38 reported statistically significant treatment-effect heterogeneity

among genotype groups (at least one loss-of-function allele vs. none; $p=0.046$), with prasugrel being superior to clopidogrel among carriers of loss-of-function CYP2C19 alleles. Overall the available studies do not suggest that CYP2C19 genotype status is a strong modifier of the treatment effects evaluated in the studies. However, these studies included only small subsets, 15 to 40 percent of the original trial populations, suggesting that selection bias may have affected their results. This was a concern particularly for the CHARISMA trial, in which differences in baseline characteristics and outcome rates were observed between the patients included and those not included in the genetic substudy. Furthermore, details were not provided regarding the timing of obtaining samples for genetic analyses, but samples were generally not obtained at the trial baseline. In such cases, survivor bias, another form of selection bias, may also affect study results.

Because of the large differences in included populations, treatments compared, and exposure and outcome definitions among studies reporting on treatment-effect modification by CYP2C19 variants on clinical outcomes, we did not perform a meta-analysis. Given that comparators (placebo, prasugrel, or ticagrelor) differed across studies, it is plausible that interaction effects could have different magnitudes or directions. For purposes of illustration, we used the counts reported in the studies to compare the treatment effect among carriers of CYP2C19 loss-of-function alleles versus noncarriers (i.e., those with normal or gain-of-function alleles), as shown in Figure B.

Figure B. Results from large randomized trials assessing effect modification by CYP2C19 variants on MACE



Note: The top set of panels presents forest plots of treatment effects (odds ratios) on MACE among carriers of at least 1 LOF allele (top left panel), treatment effects among noncarriers of LOF alleles (top middle panel), and relative effects (rOR) comparing the treatment effect among LOF carriers and LOF noncarriers (top right panel). The bottom set of panels presents forest plots of treatment effects on MACE among homozygotes for 2 LOF alleles (bottom left panel), treatment effects among nonhomozygotes of LOF alleles (bottom middle panel), and relative effects (rOR) comparing the treatment effect among homozygotes and nonhomozygotes of LOF alleles (bottom right panel). Two studies did not provide adequate data for the comparisons of homozygotes and nonhomozygotes. The CURE, CHARISMA, and ACTIVE A trials compared aspirin plus clopidogrel vs. aspirin monotherapy; the TRITON-TIMI 38 trial compared aspirin plus clopidogrel vs. aspirin plus prasugrel; the PLATO trial compared aspirin plus clopidogrel vs. aspirin plus ticagrelor. Point estimates for treatment effects are shown as black circles (carriers) or white circles (noncarriers); point estimates for relative treatment effects are shown as black squares. For all symbols, size is inversely proportional to the standard error of each estimate. Horizontal lines denote 95% confidence intervals for all estimates. Vertical dashed lines denote no effect. Please see Tables 18 and 19 in the full report for definitions of the genotype categories and outcomes reported by each study. References to individual studies are provided in Table 5 of the main report.

ACTIVE A = Atrial Fibrillation Clopidogrel Trial with Irbesartan for Prevention of Vascular Events A; CHARISMA = Clopidogrel for High Atherothrombotic Risk and Ischemic Stabilization, Management and Avoidance; CURE = Clopidogrel in Unstable Angina to Prevent Recurrent Events trial; homoz. = homozygotes; LOF = loss of function; MACE = major adverse cardiovascular events; nonhomoz. = nonhomozygotes; PLATO = Platelet inhibition and patient Outcomes trial; rOR = relative odds ratio; TRITON-TIMI 38 = Trial to Assess Improvement in Therapeutic Outcomes by Optimizing Platelet Inhibition with Prasugrel–Thrombolysis in Myocardial Infarction; Tx = treatment.

Intermediate Outcome: Platelet Reactivity During Followup

Seven studies assessing treatment-effect modification by CYP2C19 variants provided information on platelet reactivity during followup as an intermediate outcome. All seven were based on randomized trials comparing clopidogrel-based treatment with alternative therapies, had small to moderate sample sizes (range, 60 to 474 participants), and enrolled heterogeneous populations of patients with acute or chronic coronary artery disease. In this group of studies, 79 to 100 percent of the patients enrolled were included in the genetic substudies, suggesting that selection bias was unlikely. All had short followup periods (<7 days to 6 weeks), and they generally provided adequate descriptions of the methods used for generating the randomization sequence but did not provide sufficient information to assess methods of allocation concealment

The studies differed in the alleles genotyped and the genotype groupings used. Platelet reactivity during followup was assessed by the VerifyNow P2Y12 assay in all seven studies, as well as by LTA in four studies and the VASP assay (based on flow cytometry) in two studies. Because of the differences in designs, populations, treatments compared, and followup durations among the included studies, we did not perform a meta-analysis. The overall results were variable and no conclusions could be drawn.

Studies With Genetic Test–Based Selection of Patients

We identified a single multicenter trial (ELEVATE-TIMI 56) that used genetic–test-based selection of patients and then randomized them to alternative antiplatelet treatments. The study enrolled 335 patients with known cardiovascular disease (57.1% with a history of myocardial infarction; 97.3% with a history of PCI) on maintenance clopidogrel therapy (75 mg daily). The trial was well conducted, with centralized randomization and blinding of both patients and outcome assessors (both for clinical and intermediate outcomes) to the treatment assessment. The sample size was based on a priori power analysis for platelet reactivity outcomes and the recruitment target was attained. There were minimal dropouts and losses to followup.

Clinical Outcomes

The study reported no deaths or cerebrovascular events. However, it was not powered to provide robust evidence on clinical outcomes and did not have adequate followup to do so. Among CYP2C19*2 noncarriers, two patients had cardiac ischemic events while taking the 75 mg dose and

three with the 150 mg dose. Among carriers of a CYP2C19*2 allele, one patient experienced a cardiac ischemic event while taking the 75 mg clopidogrel dose.

Intermediate Outcome: Platelet Reactivity During Followup

Intermediate outcomes were assessed with the VASP assay (primary analysis) and the VerifyNow assay (secondary analysis). When treated with a standard clopidogrel maintenance dose of 75 mg/day, both CYP2C19*2 heterozygotes and homozygotes had significantly higher on-treatment platelet reactivity than did noncarrier patients ($p<0.001$ for both pairwise comparisons). Among CYP2C19*2 heterozygotes, higher clopidogrel maintenance doses up to 300 mg produced significant reductions in platelet reactivity ($p<0.001$ for trend). Results with the VerifyNow assay were similar to the VASP data across dose and genotype. Among CYP2C19*2 homozygotes, there was a trend toward less platelet reactivity with higher maintenance doses of clopidogrel; however, even with 300 mg daily of clopidogrel, these individuals had increased reactivity as measured by VASP and VerifyNow. In CYP2C19*2 heterozygotes, 150 mg resulted in platelet reactivity that tended to be higher than that seen in noncarrier patients treated with 75 mg daily. For CYP2C19*2 homozygotes, even 300 mg daily of clopidogrel did not result in platelet reactivity levels similar to those with standard clopidogrel dosing in noncarriers.

Phenotypic Testing for Platelet Reactivity

Studies of Test-and-Treat Strategies

We identified seven studies directly comparing alternative test-and-treat strategies. Six of the seven studies had a randomized design, and one was a nonrandomized comparative study of test-and-treat strategies. Generally, the randomized trials had moderate risk of bias, were prospectively conducted, and performed phenotypic testing immediately after sample collection, without knowledge of clinical or intermediate outcomes. However, information to judge whether outcomes were assessed without knowledge of the index-test result was often not reported. Subjects and personnel were not blinded, and reporting was incomplete regarding the methods of generating the randomization sequence and concealing allocation. The single nonrandomized comparative study had high risk of bias because the two groups being compared (test-guided treatment and non–test guided treatment) were enrolled in different research institutions, increasing the probability that results were affected by confounding or selection bias. Four studies evaluating the use of the VASP assay were of

moderate size (the smallest enrolling 153 patients; the largest, 429 patients); three were multicenter studies; and 1 was a single-center investigation. One randomized controlled trial (RCT) evaluating the use of the Multiplate analyzer enrolled 192 patients and one nonrandomized comparative study evaluating the same assay enrolled 798 patients. The single study assessing VerifyNow was smaller (60 patients) and was conducted in a single research center.

The six RCTs directly comparing alternative test-and-treat strategies assessed patients undergoing PCI. Four enrolled patients with stable coronary artery disease or acute coronary syndromes, and one enrolled exclusively patients undergoing elective stenting. The experimental groups in five studies (three using the VASP assay, one using the VerifyNow assay, and one using the Multiplate analyzer) employed repeat reactivity monitoring at multiple time points with modification of the administered clopidogrel dose on the basis of test results. The other two studies performed only a single assessment of platelet reactivity, with subsequent treatment modification in patients found to have reactivity values above a predefined threshold. Control groups were given clopidogrel-based therapy at standard doses. Four studies reported a prospective power calculation and enrollment goal, which was met in all cases.

Clinical Outcomes

All seven studies comparing alternative test-and-treat strategies reported information on cardiovascular mortality. In addition, six reported on MACE (composite outcomes), five on stent thrombosis, three on acute coronary syndromes (myocardial infarction or unstable angina), three on myocardial infarction alone, two on all-cause mortality, and two on repeat revascularization. Overall, the studies had short followup durations and included moderate numbers of participants; thus, the outcome rates were low, and relative effect estimates (when possible to calculate) were often extreme (e.g., odds ratios <0.5) and had substantial uncertainty (wide CIs). Studies generally indicated that the groups with test-based monitoring had better outcomes (lower event rates) than the groups without test-based monitoring; however, the differences were often not statistically significant. Meta-analyses were not performed, owing to the differences in the populations included, interventions compared, and durations of followup used.

Intermediate Outcome: Platelet Reactivity During Followup

Four of the seven studies directly comparing alternative test-and-treat strategies reported information on platelet reactivity as an intermediate outcome. Although results generally indicated that platelet reactivity at the last followup assessment was lower in the groups that received test-based treatment than in those that received standard treatment, reporting was often incomplete and precluded statistical comparisons between groups. Furthermore, studies had short followup periods, and it was unclear whether the observed differences in reactivity affected clinical outcomes.

Studies of Treatment-Effect Modification by Baseline Platelet Reactivity

We identified three studies reporting information on effect modification by baseline platelet reactivity in patients randomized to alternative antiplatelet therapies.

Clinical Outcomes

One study, a platelet reactivity substudy of the ISAR-REACT 4 trial (Intracoronary Stenting and Antithrombotic Regimen: Rapid Early Action for Coronary Treatment-4), reported information on clinical outcomes. In the platelet sub-cohort of this trial, 205 patients (36%) had high on-clopidogrel platelet reactivity at the time of PCI (35.0% in the abciximab plus heparin group vs. 37.6% in the bivalirudin group). A significant interaction was observed between study treatment arm and platelet aggregation regarding the combined efficacy endpoint (death, myocardial infarction, or urgent target vessel revascularization; P for interaction = 0.037). This study was considered to have moderate risk of bias because the patient population was not representative of the parent trial population, suggesting the possibility of selection bias.

Intermediate Outcome: Platelet Reactivity During Followup

Two studies reported information on platelet reactivity during followup. The first study was a post hoc evaluation based on a crossover RCT comparing triple therapy (aspirin + clopidogrel + cilostazol) with double therapy (aspirin + clopidogrel + placebo) for patients with stable coronary artery disease. Based on the study results, we estimated that baseline platelet reactivity did not modify the effect of cilostazol on subsequent measurements. This study was considered to have high risk of bias because it was a post hoc assessment based on a convenience sample enrolled in a crossover trial and because the parent trial had a large withdrawal rate (23%). The second study reported the response rate among “poor responders” to the

clopidogrel loading dose during prasugrel-based therapy and during clopidogrel-based therapy. Generally the response rates were higher during prasugrel therapy, regardless of the assay used to assess platelet reactivity. However, the study did not report the response status during followup for patients who were “responders” to the clopidogrel loading dose. Thus, the interaction between post-loading-dose response to clopidogrel and treatment assignment could not be assessed. This study was considered to have a high risk of bias because of incomplete outcome reporting and because information on the generation of the randomized sequence and allocation concealment was unclear.

Studies With Phenotypic Test–Based Selection of Patients

Fourteen studies met our inclusion criteria and reported information on the comparative effectiveness of treatments administered to patients selected on the basis of baseline platelet reactivity. The sample sizes ranged from 21 to more than 2,000 participants, and all 14 studies were relatively recent (published in 2008–2012). Only two trials, the GRAVITAS multicenter trial (Gauging Responsiveness with a VerifyNow Assay—Impact on Thrombosis and Safety) and the TRIGGER PCI trial (Testing Platelet Reactivity In Patients Undergoing Elective Stent Placement on Clopidogrel to Guide Alternative Therapy With Prasugrel), reported data from more than 100 randomized patients. Eleven studies were performed mainly or exclusively in the PCI setting, two studies included patients with stable coronary artery disease (noninterventional setting), and one study enrolled patients on chronic hemodialysis receiving clopidogrel treatment. On-clopidogrel platelet reactivity was used as a selection criterion in all studies; it was assessed using the VerifyNow P2Y12 assay in nine studies, LTA in three studies, the VASP assay with flow cytometry in two studies, and other assays in two studies. (One study combined measurements from three assays to define high on-treatment reactivity.) The treatment comparisons were between standard-dose clopidogrel-based therapy and high-dose clopidogrel in six studies, prasugrel in four studies, ticagrelor in two studies, and addition of a glycoprotein IIb/IIIa inhibitor in two studies.

Overall, the risk of bias varied across studies. The GRAVITAS trial had low risk of bias, both regarding aspects related to the index test of interest and regarding general aspects of randomized trial design (e.g., generation of the randomization sequence and allocation concealment). The TRIGGER-PCI trial did not provide adequate information about the randomization procedure and allocation concealment or blinding of patients to

treatment assignment; however outcomes assessors were blinded to treatment assignment. Smaller studies (typically with short-term followup) were generally considered to have a higher risk of bias, owing to problems in the application of the tests of interest (e.g., an unclear rationale for the thresholds used) or incomplete reporting of outcomes. Furthermore, these studies often did not provide information sufficient to judge their risk of bias regarding general aspects of randomized trial design.

Clinical Outcomes

Clinical-outcome comparisons between the randomized treatment groups were reported in 10 of the 14 studies. Here, we discuss only the results of the two larger trials (GRAVITAS and TRIGGER-PCI). The remaining 12 studies had smaller sample sizes, ranging from 21 to 159 patients, and also had short followup durations. Information on these trials is presented in the full report.

The GRAVITAS trial (2,214 randomized patients) included patients who had undergone PCI for stable coronary artery disease or non-ST-segment-elevation acute coronary syndrome and showed increased on-clopidogrel reactivity on the VerifyNow P2Y12 assay. The patients were randomized to high-dose clopidogrel or standard-dose clopidogrel, both in combination with aspirin. After 6 months of followup, there was no statistically significant difference between the randomized groups in the rate of cardiovascular death, nonfatal myocardial infarction, stent thrombosis, all-cause mortality, or composite cardiovascular outcomes, either (1) cardiovascular death or nonfatal myocardial infarction or (2) cardiovascular death, nonfatal myocardial infarction, or stent thrombosis. The study also included followup information for a randomly selected group of patients with low platelet reactivity at baseline who were treated with standard-dose clopidogrel. (See the Results section for Key Question 1b for details.)

The TRIGGER-PCI study compared prasugrel versus standard-dose clopidogrel in 423 patients with high on-clopidogrel platelet reactivity as measured by the VerifyNow P2Y12 assay. After 236 patients had completed the planned 6-month followup, a blinded interim review identified a single primary endpoint event. Because of the very low event rate, the trial was terminated early for futility. As such, for all outcomes, event rates were very low and differences in event rates between groups were not statistically significant. Across all 10 studies reporting data on clinical outcomes, patient populations were heterogeneous, selected on the basis of different inclusion criteria, and assessed using different therapeutic regimens. For these reasons, we did not perform meta-analyses for any of the clinical outcomes reported.

Intermediate Outcome: Platelet Reactivity During Followup

Ten studies reported information on intermediate outcomes. Eight studies had a total duration of 3 months or less; five studies had a crossover design. The outcomes were assessed using different assays and were heterogeneously reported. Generally, patients on higher dose clopidogrel regimens and those receiving prasugrel showed greater responses in platelet reactivity compared with those receiving standard-dose clopidogrel regimens.

Combined Genetic Testing for CYP2C19 Variants and Phenotypic Testing To Guide Antiplatelet Treatment

We identified four studies providing information on test-based treatment strategies that also provided information on the CYP2C19 genotype of participants. All four studies reported genetic analyses based on randomized trials that had enrolled patients on the basis of baseline platelet reactivity testing. Briefly, two of the studies were conducted in the setting of small (21 and 64 patients) crossover RCTs of short duration (30 and 60 days); one was based on a short-term parallel-arm trial (2 weeks of followup); and one study (GIFT—Genotype Information and Functional Testing) was conducted in the setting of the large GRAVITAS trial with a followup of 6 months. Analyses stratified by treatment and genotype status were not reported for clinical outcomes, and all four studies reported results for the intermediate outcome of platelet reactivity. Studies did not report significant effect modification by genotype for this outcome. (All analyses assumed a dominant model for loss-of-function alleles; analyses under an alternative model were not possible.) In general, results were inconclusive because studies were small and none had been prospectively powered specifically to assess effect modification by genotype.

Key Question 3b: Factors Modifying the Comparative Effectiveness of Alternative Test-and-Treat Strategies

Only four of the studies considered relevant to Key Question 3a provided information about the use of testing for clinical decisionmaking with data stratified by patient characteristics: ancestry in two, baseline percent inhibition of on-clopidogrel reactivity in one, diabetes status in one, and history of PCI and symptomatic atherothrombosis on trial entry in one. None of the factors appeared to statistically significantly affect study results relevant to the use of testing to guide antiplatelet therapy.

Key Question 4: Harms of Testing and of Test-Directed Treatment

Harms of Test-Directed Treatment

All studies addressing Key Question 4 were also included in Key Question 3a; assessment of the risk-of-bias of individual studies is addressed in that section. We discuss studies belonging to each of three designs—studies of test-and-treat strategies, studies of treatment-effect modification, and studies with test-based selection—separately for genetic testing (for CYP2C19 variants) and for phenotypic testing (of platelet reactivity).

Genetic Testing for CYP2C19 Variants

We identified a single study comparing testing for CYP2C19 variants against a no-testing strategy to guide treatment decisionmaking. The study monitored major and minor bleeding using the thrombolysis in myocardial infarction (TIMI) classification over 30 days of followup. The frequency of minor and major bleeding was not different between the study groups.

Studies of Treatment-Effect Modification by CYP2C19 Genotype Status

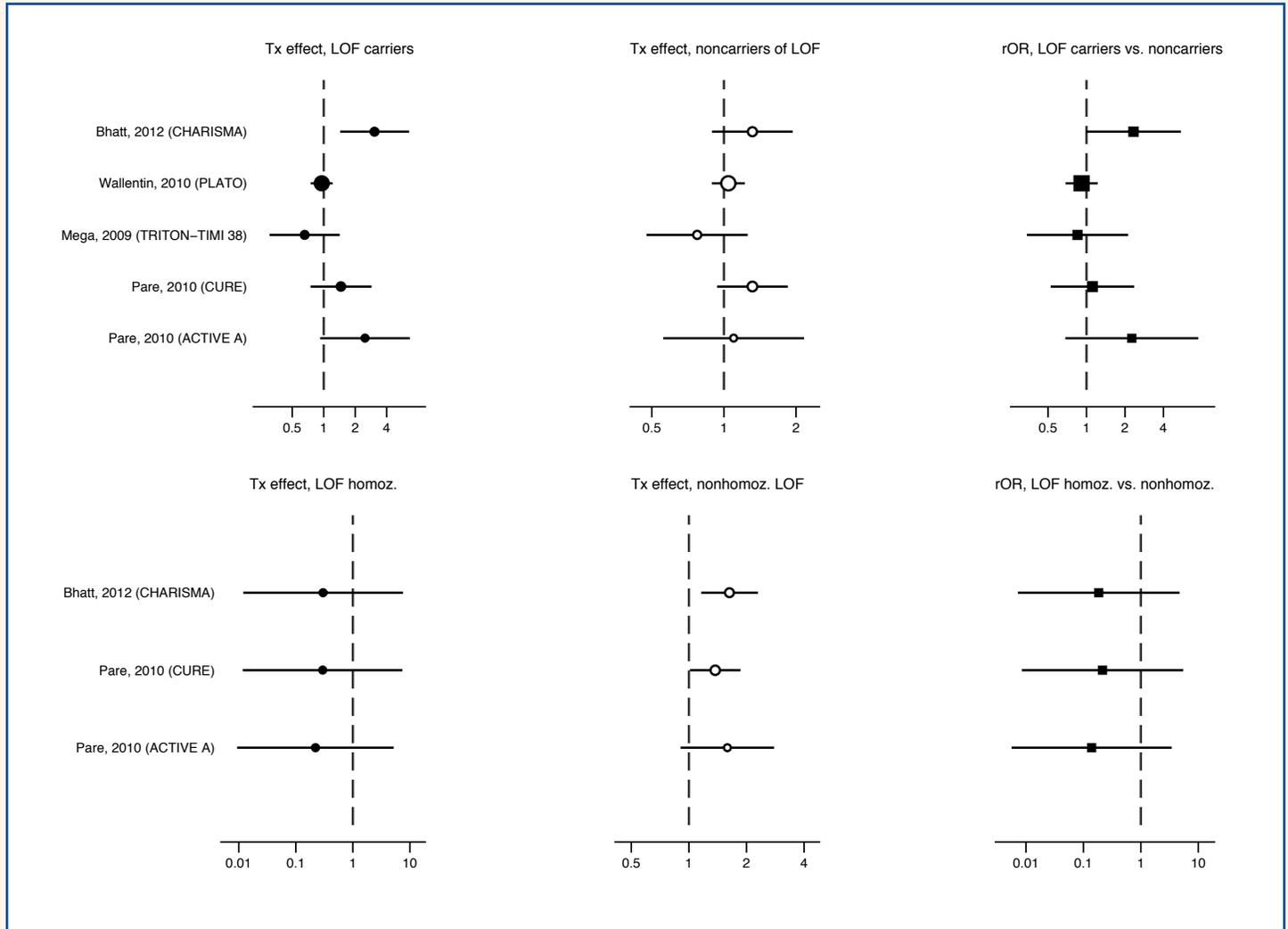
Six studies (reported in five publications) provided information on treatment-effect modification of bleeding outcomes by CYP2C19 status. Five were based on large randomized trials of clopidogrel-based therapy that included more than 1,000 patients in their genetic substudies (the same studies discussed in the corresponding section of Key Question 3a). The sixth study was a small genetic substudy of 126 patients that reported no major bleeding events by TIMI criteria in either group. The five larger studies compared the effect of alternative treatment strategies, stratified by CYP2C19 genotype, on safety outcomes (in all five studies, bleeding events). The test for interaction (a test for heterogeneity of treatment effects across genotype groups) was not statistically significant for any of the reported comparisons, indicating that the impact of the compared treatments on bleeding events was not significantly different across patient groups defined by CYP2C19 genotype.

Because of the large differences in populations included, treatments compared, and exposure and outcome definitions among studies reporting on treatment-effect modification by CYP2C19 variants, we did not perform a meta-analysis. However, we used the counts reported in the studies to compare the treatment effect among carriers of CYP2C19*2 or *3 (loss-of-function alleles) versus noncarriers (i.e., carriers of CYP2C19*1 or *17 [normal and gain-of-function alleles, respectively]). The odds ratios for the treatment effect within each genotype subgroup and

relative odds ratios comparing the treatment effect across genotype groups showed that treatment-effect modification was nonstatistically significant in all five studies (Figure C). Effect modification was also nonstatistically significant

under a recessive genetic model; however, only three studies provided data for this analysis and CIs were wide (reflecting the low number of homozygous individuals in each study).

Figure C. Bleeding events in large randomized trials reporting information on effect modification by CYP2C19 variants



Note: The top set of panels presents forest plots for treatment effects (odds ratios) on bleeding outcomes among carriers of at least 1 LOF allele (top left panel), treatment effects among noncarriers of LOF alleles (top middle panel), and relative effects (rOR) comparing the treatment effect among LOF carriers and LOF noncarriers (top right panel). The bottom set of panels presents forest plots of treatment effects on bleeding outcomes among homozygotes for 2 LOF alleles (bottom left panel), treatment effects among nonhomozygotes of LOF alleles (bottom middle panel), and relative effects comparing the treatment effect among homozygotes and nonhomozygotes of LOF alleles (bottom right panel). Two studies did not provide adequate data for the comparisons of homozygotes and nonhomozygotes. The CURE, CHARISMA, and ACTIVE A trials compared aspirin plus clopidogrel vs. aspirin monotherapy; the TRITON-TIMI 38 trial compared aspirin plus clopidogrel vs. plus prasugrel; the PLATO trial compared aspirin plus clopidogrel vs. aspirin plus ticagrelor. Point estimates for treatment effects are shown as black circles (carriers) or white circles (noncarriers); point estimates for relative treatment effects are shown as black squares. For all symbols, size is inversely proportional to the standard error of each estimate. Horizontal extending lines denote 95% confidence intervals for all estimates. Vertical dashed lines denote no effect. Please see Table 41 in the full report for definitions of the genotype categories and outcomes reported by each study. References to individual studies are provided in Table 5 of the main report.

ACTIVE A = Atrial Fibrillation Clopidogrel Trial with Irbesartan for Prevention of Vascular Events A; CHARISMA = Clopidogrel for High Atherothrombotic Risk and Ischemic Stabilization, Management, and Avoidance; CURE = Clopidogrel in Unstable Angina to Prevent Recurrent Events trial; homoz. = homozygote; LOF = loss of function; nonhomoz. = nonhomozygote; PLATO = PLATElet inhibition and patient Outcomes trial; rOR = relative odds ratio; TRITON-TIMI 38 = Trial to Assess Improvement in Therapeutic Outcomes by Optimizing Platelet Inhibition with Prasugrel–Thrombolysis in Myocardial Infarction; Tx = treatment.

Studies With Genetic Test–Based Selection of Patients

One study, the ELEVATE-TIMI 56 trial, genotyped patients on chronic clopidogrel therapy for the presence of CYP2C19 *2 alleles. Patients with at least one *2 allele were randomized to various sequences of clopidogrel at doses of 75, 150, 225, or 300 mg daily, each for approximately 2 weeks. Noncarriers were randomized to clopidogrel 75 or 150 mg daily, each dose for two periods of approximately 2 weeks. There were no TIMI major or minor bleeding events overall, and there were no significant differences in hematologic, gastrointestinal, or musculoskeletal disorders in CYP2C19*2 carriers across different clopidogrel doses.

Phenotypic Testing for Platelet Reactivity

Studies of Test-and-Treat Strategies

Seven studies comparing alternative test-and-treat strategies provided information on harms of test-directed treatment. The studies had short followup durations (1 year in one study, 6 months in another, and 30 days in the remaining five), and few events were observed, particularly severe or major bleeding outcomes. Consequently, data were sparse and CIs around effect estimates were wide, indicating substantial uncertainty.

Studies of Treatment-Effect Modification by Baseline Platelet Reactivity

Two studies provided information on treatment-effect modification by baseline on-clopidogrel platelet reactivity. One reported no severe bleeding events and the other reported no significant effect modification.

Studies With Phenotypic Test–Based Selection of Patients

Twelve of the 14 randomized trials with phenotypic test–based patient selection reported treatment-related harms. The two larger studies (the GRAVITAS and the TRIGGER-PCI trials) found no statistically significant difference in bleeding events. The remaining 12 small studies had short followup durations (<1 month in 6 of the 12 studies) and generally reported low rates of events.

Combined Testing for CYP2C19 Variants and Phenotypic Testing To Guide Antiplatelet Treatment

Of the four studies providing information on test-based treatment strategies that also provided information on the CYP2C19 genotype of participants, none reported data on treatment-related harms stratified by treatment group and genotype status. Therefore, the interaction of genotype status and treatment could not be assessed.

Harms of Testing Per Se

We found no studies reporting on the harms of the testing process for CYP2C19 genotyping or measuring platelet reactivity in the populations of interest. However, one study comparing VASP-guided therapy with standard clopidogrel dosing in the PCI setting noted that patients in the test-guided arm had a longer time from clopidogrel loading to PCI than patients in the arm that was not test guided ($p < 0.001$). The delay was due to the need for repeat testing and treatment modification until a predefined reactivity threshold was reached in the test-guided group. It is unclear whether this delay resulted in harm to patients.

Discussion

Clopidogrel is used extensively in the interventional management of coronary artery disease and the treatment and secondary prevention of acute coronary syndromes.⁴³ Furthermore, it is used for the management of patients undergoing neurointervention (with stent placement), for the prevention of stroke in patients with atrial fibrillation who are not candidates for vitamin K antagonist therapy, and for the management of selected patients with peripheral arterial disease. However, response to clopidogrel therapy—as assessed by ex vivo studies of platelet function—is variable among patients and over time. Some patients experience little suppression of platelet reactivity despite adhering to treatment, while others experience more profound suppression that may increase their risk of bleeding. Given the availability of several therapeutic options for antiplatelet treatment (e.g., increasing the loading or daily maintenance dose of clopidogrel or using adjunctive or replacement therapies such as prasugrel, ticagrelor, or cilostazol), there is interest in reliably identifying patients who are less likely to respond to standard clopidogrel treatment, as well as those who are most likely to respond to alternative treatments. This report reviewed the evidence of the effectiveness and comparative effectiveness of two types of tests that have been extensively evaluated as biomarkers for outcome prognosis for patients receiving clopidogrel therapy and as biomarkers of treatment response: genetic testing for

CYP2C19 variants and phenotypic testing for on-clopidogrel platelet reactivity.

Key Findings and Assessment of the Strength of Evidence

Table B presents a summary of the report's key findings. When appropriate, results are presented separately for each of the populations and outcomes of interest. We did not assess the strength of evidence for studies of analytic validity because analytic validity is a prerequisite for the clinical use of the tests and because no framework exists

for assessing the strength of evidence for analytic validity studies. We also did not assess the strength of evidence for studies exclusively assessing platelet reactivity as an outcome because platelet reactivity measurements during followup are not usually performed as part of clinical care and because platelet reactivity is not a patient-relevant outcome. Instead, we focus here on the body of evidence pertaining to predictive effects, treatment decisionmaking, and harms as related to patient-relevant clinical outcomes. Please see the Methods section for a detailed discussion of our approach to rating the strength of evidence.

Table B. Key findings from this review and assessment of strength of evidence

Key Question	Population	Test/Assay	Outcome	SOE Summary and Comments
1a: Analytic validity of tests for genotyping CYP2C19 variants	NA	Genotyping for any CYP2C19 variant	NA	<p>SOE = NA</p> <ul style="list-style-type: none"> Few studies provided information on analytic validity specifically using samples obtained from patient populations relevant to this review. When available, data were limited to test-retest reliability or interassay agreement There was limited information comparing the validity of different genetic testing assays. However, it is generally accepted that the analytic validity of genotyping assays is robust.
1b: Predictive value of genetic testing for CYP2C19 variants	Ischemic heart disease	Genotyping for LOF CYP2C19 variants	Stent thrombosis	<p>SOE = Moderate</p> <ul style="list-style-type: none"> Meta-analysis of 17 studies found a statistically significant association under a dominant model RR=1.52; 95% CI, 1.17 to 1.97 There was little evidence of heterogeneity ($I^2 = 0\%$), but the test for small-study effects was statistically significant. Results under additive and recessive models were consistent with a positive association and produced larger effect sizes; however, these analyses were based on a small subset of the available studies. There was some concern about selective outcome reporting. Studies reported few outcome events and the summary estimate was imprecise.
			MACE	<p>SOE = Moderate</p> <ul style="list-style-type: none"> Meta-analysis of 25 studies found a statistically significant association under a dominant model RR=1.20; 95% CI, 1.04 to 1.39 There was some evidence of heterogeneity ($I^2 = 31\%$). Results under additive and recessive models were consistent with a positive association and produced larger effect sizes; however, these analyses were based on a small subset of the available studies. The test for small-study effects was statistically significant.

Table B. Key findings from this review and assessment of strength of evidence (continued)

Key Question	Population	Test/Assay	Outcome	SOE Summary and Comments
			Cardiovascular mortality	<p>SOE = Low</p> <ul style="list-style-type: none"> • Meta-analysis of 7 studies found a statistically significant association under a dominant model • RR=1.98; 95% CI, 1.13 to 3.46 • There was no evidence of heterogeneity ($I^2 = 0\%$). The summary estimate was imprecise. • The test for small-study effects was not statistically significant. • There was some concern about selective outcome reporting.
			All other clinical outcomes	<p>SOE = Insufficient</p> <ul style="list-style-type: none"> • Few studies reported information for noncomposite clinical outcomes other than stent thrombosis. • There was substantial concern about selective outcome reporting. • Study-specific and meta-analysis estimates (when performed) indicated substantial uncertainty.
	Genotyping for GOF CYP2C19 variants		MACE	<p>SOE = Low</p> <ul style="list-style-type: none"> • Meta-analysis of 7 studies found a statistically significant protective effect for carriers vs. noncarriers • RR=0.82; 95% CI, 0.74 to 0.92 • There was substantial concern about selective outcome reporting.
All other clinical outcomes			All other clinical outcomes	<p>SOE = Insufficient</p> <ul style="list-style-type: none"> • Few studies provided relevant information. • There was substantial concern about selective outcome reporting. • Study-specific and meta-analysis estimates (when performed) indicated substantial uncertainty.
Other patient groups who are candidates for clopidogrel therapy	Genotyping for any CYP2C19 variants		All clinical outcomes	<p>SOE = Insufficient</p> <ul style="list-style-type: none"> • Only few studies (often only a single study) were available for patient populations other than those with ischemic heart disease. • Some of the studies did not report information on clinical outcomes.

Table B. Key findings from this review and assessment of strength of evidence (continued)

Key Question	Population	Test/Assay	Outcome	SOE Summary and Comments
1c: Factors affecting the predictive value of genetic testing for CYP2C19 variants	All patient populations	Genotyping for any CYP2C19 variants	All clinical outcomes	<p>SOE = Insufficient</p> <ul style="list-style-type: none"> 20 studies provided information on effect modification; a single study reported a statistically significant interaction effect on the prognostic performance of CYP2C19 variants for a clinical outcome. No factor was assessed by more than 5 studies, giving rise to concerns about selective outcome reporting. In meta-regression analyses (using study-level factors as covariates) we found some evidence of effect modification by ethnicity (East Asians vs. white populations) for MACE and stent thrombosis. However, this result is based on comparisons across studies, which may be confounded by other study characteristics, and was not corroborated by within-study analyses. CIs for all interaction effects were wide for all genotype-outcome pairs assessed, and only a few studies in East Asian populations were available. Estimates of effect modification by study-level variables are susceptible to confounding by other study-level characteristics.
2a: Analytic validity of tests for on-clopidogrel platelet reactivity	NA	All assays used to measure on-clopidogrel platelet reactivity	NA	<p>SOE = NA</p> <ul style="list-style-type: none"> Few studies reported information on analytic sensitivity and specificity, possibly reflecting the research community's belief that there is no good reference standard assay for platelet reactivity. Agreement ranged from poor to moderate and was variable between tests. The highest agreement was observed between applications of the same assay with different concentrations of agonists rather than between different assays.
2b: Predictive ability of phenotypic testing for platelet reactivity	Ischemic heart disease	LTA	All-cause mortality	<p>SOE = Low</p> <ul style="list-style-type: none"> 13 studies using heterogeneous methods to define increased reactivity were available These studies support an association between increased platelet reactivity measured by LTA and mortality. There was some concern about selective outcome reporting.
			Cardiovascular mortality	<p>SOE = Low</p> <ul style="list-style-type: none"> 9 studies using heterogeneous methods to define increased reactivity were available. Studies provided evidence of an association between increased reactivity and cardiovascular mortality; however, clinical heterogeneity precluded firm conclusions. There was some concern about selective outcome reporting.

Table B. Key findings from this review and assessment of strength of evidence (continued)

Key Question	Population	Test/Assay	Outcome	SOE Summary and Comments
			Acute coronary syndromes	<p>SOE = Low</p> <ul style="list-style-type: none"> 18 studies using heterogeneous methods to define increased reactivity were available. Studies often found statistically significant associations between increased reactivity as measured by LTA and clinical events; however, clinical heterogeneity did not allow for stronger conclusions. There was some concern about selective outcome reporting.
			Stent thrombosis	<p>SOE = Low</p> <ul style="list-style-type: none"> 19 studies using heterogeneous methods to define increased reactivity were available. Studies often found statistically significant associations between increased reactivity as measured by LTA and clinical events; however, clinical heterogeneity did not allow for stronger conclusions. There was some concern about selective outcome reporting.
			Stroke	<p>SOE = Low (for lack of association)</p> <ul style="list-style-type: none"> 12 studies using heterogeneous methods to define increased reactivity were available. Studies generally did not report statistically significant associations between increased reactivity as measured by LTA and clinical events; however, clinical heterogeneity did not allow for stronger conclusions or quantitative synthesis to increase precision.
			MACE	<p>SOE = Low</p> <ul style="list-style-type: none"> 37 studies using heterogeneous methods to define increased reactivity were available. The majority of reviewed studies suggested a statistically significant association between increased platelet reactivity measured by LTA and composite cardiovascular events. Definitions of composite outcomes were often heterogeneous.
			All other clinical outcomes	<p>SOE = Insufficient</p> <ul style="list-style-type: none"> Clinical and population heterogeneity or small number of studies limited our ability to draw conclusions.

Table B. Key findings from this review and assessment of strength of evidence (continued)

Key Question	Population	Test/Assay	Outcome	SOE Summary and Comments
	VerifyNow	All-cause mortality	All-cause mortality	<p>SOE = Low (for lack of association)</p> <ul style="list-style-type: none"> 10 studies were available. Meta-analysis of 4 studies did not find an association between increased reactivity measured by VerifyNow and all-cause mortality RR=1.21; 95% CI, 0.83 to 1.76 The summary estimate was imprecise and 95% CI did not rule out clinically meaningful effects.
		Cardiovascular mortality	Cardiovascular mortality	<p>SOE = Moderate</p> <ul style="list-style-type: none"> 7 studies were available. Meta-analysis of 4 studies found a statistically significant association with little evidence of heterogeneity RR=2.50; 95% CI, 1.28 to 4.87 The CI of the summary estimate indicated substantial uncertainty. There was some concern about selective outcome reporting.
		Acute coronary syndromes	Acute coronary syndromes	<p>SOE = Low</p> <ul style="list-style-type: none"> 19 studies using heterogeneous methods to define increased reactivity were available. Studies generally suggested an association between increased reactivity as measured by VerifyNow and acute coronary syndromes, both periprocedurally and during longer followup.
		Stent thrombosis	Stent thrombosis	<p>SOE = Low (for lack of association)</p> <ul style="list-style-type: none"> 15 studies were available. Meta-analysis of 6 studies did not find an association between reactivity measured by VerifyNow and stent thrombosis RR=1.67; 95% CI, 0.80 to 3.47 There was some evidence of heterogeneity (I² = 37%) and the CI of the summary estimate indicated substantial uncertainty. Studies not included in the meta-analysis generally produced nonsignificant results. There was some concern about selective outcome reporting. The test for small-study effects was statistically significant.

Table B. Key findings from this review and assessment of strength of evidence (continued)

Key Question	Population	Test/Assay	Outcome	SOE Summary and Comments
			MACE	<p>SOE = Moderate</p> <ul style="list-style-type: none"> • 24 studies were available. Meta-analysis of 13 studies identified a statistically significant association • RR=2.48; 95% CI, 1.85 to 3.32 • There was moderate statistical heterogeneity (I² = 44%) and studies used fairly similar definitions of increased reactivity. • The test for small-study effects was statistically significant.
			Bleeding events (major and all levels of severity combined)	<p>SOE = Low (for lack of association)</p> <ul style="list-style-type: none"> • 13 studies were available. • Meta-analysis of 4 studies with data on any bleeding event did not find an association between increased reactivity measured by VerifyNow • RR = 1.09; 95% CI, 0.88 to 1.36 • There was little evidence of heterogeneity (I² = 0%). • Meta-analysis of 4 studies with data on major bleeding events did not find an association between increased reactivity measured by VerifyNow • RR=0.85; 95% CI, 0.32 to 2.25 • There was evidence of moderate heterogeneity (I² = 57%). • For major bleeding events the summary estimate was imprecise and the 95% CI did not rule out clinically meaningful effects.
			All other clinical outcomes	<p>SOE = Insufficient</p> <ul style="list-style-type: none"> • Clinical heterogeneity or small number of studies limited our ability to draw conclusions.
	VASP assay		Cardiovascular mortality	<p>SOE = Insufficient</p> <ul style="list-style-type: none"> • 6 studies were available. Meta-analysis of 4 studies did not identify a statistically significant association. • RR=2.42; 95% CI, 0.86 to 6.82 • Although the test for heterogeneity was nonsignificant, point estimates from individual studies ranged from protective effects to strong harmful effects. • The meta-analytic summary point estimate was far from the null and its CI was wide (imprecise). • Clinically significant effects could not be ruled out.

Table B. Key findings from this review and assessment of strength of evidence (continued)

Key Question	Population	Test/Assay	Outcome	SOE Summary and Comments
			Acute coronary syndromes	<p>SOE = Low (for lack of association)</p> <ul style="list-style-type: none"> • 6 studies were available. Meta-analysis of 3 studies did not identify a statistically significant association • RR=1.47; 95% CI, 0.77 to 2.79 • The test for heterogeneity was nonsignificant but point estimates from individual studies were highly variable. • The meta-analytic summary point estimate was far from the null and its CI was wide (imprecise). • Clinically significant effects could not be ruled out.
			Stent thrombosis	<p>SOE = Low</p> <ul style="list-style-type: none"> • 10 studies were available. Meta-analysis of 4 studies identified a statistically significant association • RR=3.37; 95% CI, 1.59 to 7.11 • There was little evidence of statistical heterogeneity and the 4 studies used fairly similar definitions of increased reactivity. • The summary estimate was imprecise but the lower bound was consistent with a 59% increase in risk in the high-reactivity group. • There was some concern about selective outcome reporting.
			MACE	<p>SOE = Low</p> <ul style="list-style-type: none"> • 8 studies were available. Meta-analysis of 6 studies identified a statistically significant association • RR=2.57; 95% CI, 1.21 to 5.47 • There was evidence of statistical heterogeneity. • The summary estimate was imprecise but the lower bound was consistent with a 21% increase in risk in the high-reactivity group. • There was some concern about selective outcome reporting.
	All other clinical outcomes			<p>SOE = Insufficient</p> <ul style="list-style-type: none"> • Few studies reported information. • Clinical heterogeneity or small number of studies limited our ability to draw conclusions.

Table B. Key findings from this review and assessment of strength of evidence (continued)

Key Question	Population	Test/Assay	Outcome	SOE Summary and Comments
		PFA-100	MACE	<p>SOE = Low</p> <ul style="list-style-type: none"> 7 of the 9 studies on this assay reporting information on composite clinical outcomes produced statistically significant results indicating an association between increased reactivity and adverse outcomes. Heterogeneity in the methods used to define increased reactivity precluded definitive conclusions; however, studies generally indicated an association between increased platelet reactivity as measured by the PFA-100 assay and composite clinical outcomes.
			All other clinical outcomes	<p>SOE = Insufficient</p> <ul style="list-style-type: none"> Few of the available studies reported information on other outcomes. There was concern about selective outcome reporting.
		All other assays	All clinical outcomes	<p>SOE = Insufficient</p> <ul style="list-style-type: none"> Few studies were available. When ≥ 2 studies were available for the same outcome, they used heterogeneous metrics or thresholds to define increased reactivity or used different agonists for ex vivo stimulation of platelets.
	Other patient groups who are candidates for clopidogrel therapy	All assays used to measure on-clopidogrel platelet reactivity	All clinical outcomes	<p>SOE = Insufficient</p> <ul style="list-style-type: none"> Only 6 studies, using diverse assays to measure reactivity, were available in clinically heterogeneous populations. Studies were fairly small.
2c: Factors affecting the predictive value of phenotypic testing for platelet reactivity	All patient populations	All assays used to measure on-clopidogrel platelet reactivity	All clinical outcomes	<p>SOE = Insufficient</p> <ul style="list-style-type: none"> 7 studies provided information on effect modification; no factor was assessed by more than 3 studies. Effect modification by study-level factors could not be assessed for most assay-outcome pairs; when such analysis was possible (for VerifyNow MACE), results indicated substantial uncertainty.

Table B. Key findings from this review and assessment of strength of evidence (continued)

Key Question	Population	Test/Assay	Outcome	SOE Summary and Comments
3a: Comparative effectiveness of alternative test-and-treat strategies	Ischemic heart disease	Genetic testing for CYP2C19 variants or phenotypic testing for platelet reactivity (all assays assessed)	All clinical outcomes	<p>SOE = Insufficient</p> <ul style="list-style-type: none"> 1 RCT of testing vs. no testing was identified. The study had short duration and a small sample size; no events were observed in the 2 groups during the study period. 3 studies provided information on treatment-effect modification for clinical outcomes and reported at least 1 outcome event. 1 study randomized patients selected on the basis of genotype status into different clopidogrel doses. No conclusions could be drawn regarding clinical outcomes because of the short duration and small sample size of the study. Studies compared different antiplatelet treatments and produced heterogeneous results. Study-specific estimates were imprecise.
		Phenotypic testing for platelet reactivity	All clinical outcomes	<p>SOE = Insufficient</p> <ul style="list-style-type: none"> The 6 RCTs of testing strategies were small, had different designs, and produced extreme results with considerable statistical uncertainty. 1 NRCS was judged to be at high risk of bias on the basis of study design (patients in each of the 2 compared arms were enrolled at different centers). 3 studies of effect modification were identified; studies evaluated heterogeneous interventions and used different methods to assess reactivity. Studies of test-based patient selection assessed different treatments, enrolled heterogeneous patient populations, and did not provide robust evidence on clinical outcomes.
	Atrial fibrillation	Genetic testing for CYP2C19 variants or phenotypic testing for platelet reactivity (all assays assessed)	All clinical outcomes	<p>SOE = Insufficient</p> <ul style="list-style-type: none"> Only 1 study providing information on effect modification by CYP2C19 status was identified The study did not find evidence of effect modification by genotype status, but there was considerable statistical uncertainty in the study estimates.
		Phenotypic testing for platelet reactivity	All clinical outcomes	<p>SOE = Insufficient</p> <ul style="list-style-type: none"> No studies were identified.

Table B. Key findings from this review and assessment of strength of evidence (continued)

Key Question	Population	Test/Assay	Outcome	SOE Summary and Comments
	Other patient populations	Genetic testing for CYP2C19 variants or phenotypic testing for platelet reactivity (all assays assessed)	All clinical outcomes	SOE = Insufficient <ul style="list-style-type: none"> • 1 study provided information on treatment-effect modification in a mixed population of patients with atherosclerotic disease (cardiovascular, cerebrovascular, or peripheral arterial) along with asymptomatic individuals at risk for atherosclerotic disease. • The study did not provide robust evidence of effect modification.
		Phenotypic testing for platelet reactivity	All clinical outcomes	SOE = Insufficient <ul style="list-style-type: none"> • No studies were identified.
3b: Factors modifying the comparative effectiveness of alternative test-and-treat strategies	All patient populations	Genetic testing for CYP2C19 variants or phenotypic testing for platelet reactivity (all assays assessed)	All clinical outcomes	SOE = Insufficient <ul style="list-style-type: none"> • 4 studies provided information on effect modification. Each assessed different effect modifiers; no statistically significant interactions were reported.
4: Harms of testing and of test-directed treatment	All patient populations	Genetic testing for CYP2C19 variants	All clinical outcomes	SOE = Insufficient <ul style="list-style-type: none"> • 1 RCT of testing vs. no testing was identified. The study had short duration and a small sample size; few events were observed in the 2 groups during the study period. • 5 studies assessed treatment-effect modification by genotype status and reported at least 1 outcome event. • 1 study randomized patients selected on the basis of genotype status; it did not provide robust evidence regarding harms due to the relatively small sample size and short followup. • Studies compared different antiplatelet treatments and had heterogeneous results. No studies provided direct information on the harms of testing per se.

Table B. Key findings from this review and assessment of strength of evidence (continued)

Key Question	Population	Test/Assay	Outcome	SOE Summary and Comments
	Phenotypic testing for platelet reactivity (all assays assessed)	All clinical outcomes	<p>SOE = Insufficient</p> <ul style="list-style-type: none"> The 6 randomized studies of testing strategies were small, had different designs, produced extreme results with considerable statistical uncertainty, and in some cases did not report any outcome events. 1 NRCS was judged to be at high risk of bias on the basis of study design (patients in each of the 2 compared arms were enrolled at different centers). 2 studies of effect modification were identified. In 1 study safety outcomes either did not occur (regardless of reactivity status) or results were not stratified by reactivity group; the second study did not identify a significant effect but had short-term followup and reported too few outcome events to allow any robust conclusions to be drawn. Studies of test-based patient selection assessed different treatments. 	

Note: CI = confidence interval; GOF = gain-of-function; LOF = loss-of-function; LTA = light-transmission aggregometry; MACE = major adverse cardiovascular events; NA = not applicable; NRCS = nonrandomized controlled study; PFA = Platelet Function Analyzer; RCT = randomized controlled trial; RR = relative risk; SOE = strength of evidence; VASP = vasodilator-stimulated phosphoprotein.

In summary, the analytic validity for genotyping appears well established. In contrast, the relatively poor agreement among phenotypic tests suggests that more work is needed to specify which phenotypic tests provide measurements that are usable for clinical decisionmaking. Both genetic testing for CYP2C19 variants and assays for measuring platelet reactivity appear to predict adverse cardiovascular outcomes. However, the evidence is weakened by a substantial concern about selective reporting, publication bias, and concerns about risk of bias in individual studies. Evidence of the utility of these tests to guide treatment is still inconclusive due to the small number of available studies, as well as heterogeneity in the included populations, tests used, and interventions compared. Evidence directly comparing the two testing approaches is totally lacking.

Our review has synthesized more publications than previous reviews have, with generally similar findings. Regarding the predictive effects of CYP2C19 genotype status, existing systematic reviews have reached similar conclusions to ours, both in magnitude and direction. Also consistent with our findings, previous analyses have suggested that selective outcome reporting and publication bias may have affected meta-analytic estimates.^{16,44}

Compared with previous systematic reviews regarding platelet reactivity assays, our review includes a much larger number of studies and considers multiple assays assessing on-clopidogrel platelet reactivity using agonists to stimulate platelets *ex vivo*. In contrast to previous meta-analyses, we did not combine results across different assays (i.e., across tests using different measurement principles), different agonist concentrations, or different calculation methods or cutoff values for defining high reactivity. We believe that this choice is supported by our review of analytic validity, which found low to moderate agreement between different assays. Of note, our analyses relevant to the VerifyNow P2Y12 assay include almost double the number of studies included in a recently published meta-analysis of individual data on the same assay.⁴⁵ Despite differences in selection criteria and analysis methods, our results were similar, identifying a large effect size for the association between platelet reactivity as measured by the VerifyNow P2Y12 assay and adverse cardiovascular outcomes.

To our knowledge, this is the first review to comprehensively evaluate the use of genetic and phenotypic testing to guide clinical decisionmaking. We developed a structured approach that considered different experimental designs (randomized trials of alternative test-and-treatment strategies, randomized treatment trials

assessing effect modification by biomarkers, and randomized treatment trials using the biomarkers to select patients for inclusion). Although the studies we identified were too diverse to support firm conclusions on the value of the tests of interest, we believe that our methodological approach will be helpful as the evidence base continues to grow. For example, it will be applied in our updated literature review.

Applicability

The vast majority of included studies enrolled patients with ischemic heart disease. Acute or chronic coronary disease represented almost all available studies for all Key Questions. Other populations who are potential candidates for antiplatelet therapy (e.g., patients with cerebrovascular disease, peripheral arterial disease, or atrial fibrillation) were included in a minority of studies only. This imbalance is not unexpected, given that clopidogrel's primary indications pertain to ischemic heart disease. However, it is probably not prudent to extrapolate findings from studies of ischemic heart disease to other patient populations. Given that a large number of studies included patients undergoing PCI, these findings are most applicable to interventional settings.

For CYP2C19 variants, we found limited evidence that prognostic effects were different in subgroup and metaregression analysis by ethnicity (East Asian vs. white). More evidence is needed to validate this finding and to obtain information on patient populations underrepresented in this review (e.g., blacks). Patient race or ethnicity may be an important effect modifier because the prevalence of variant alleles is substantially different among racial and ethnic groups. For example, *2 variants are much more common in East Asian populations than others.

The majority of studies were conducted in tertiary (usually academic) medical centers. Studies of treatment-effect modification by CYP2C19 genotype were based on large randomized trials, and findings may not be generalizable to everyday care settings. Because patient information on preexisting vascular disease in studies of predictive effects was generally incompletely reported, it is unclear whether patients in the included studies are representative of those seen in clinical practice. Nonetheless, the distribution of risk factors for ischemic vascular disease (male sex, hyperlipidemia, diabetes, hypertension, smoking, etc.) appeared to be representative of contemporary patient populations, and the majority of studies were conducted in recent years.

Implications for Clinical and Policy Decisionmaking

Despite the availability of a large literature on the use of genetic testing of CYP2C19 variants and phenotypic testing of platelet reactivity for predicting outcomes in patients receiving clopidogrel-based therapy, studies provided limited information on the value added by these tests over ascertainment of conventional risk factors in the populations of interest (e.g., clinical or laboratory information or disease-specific predictive scores). The data suggest that both test methods can provide prognostic information for some important clinical outcomes. However, selective outcome reporting for both types of tests, uncertainty about the underlying genetic model for CYP2C19 variants, and heterogeneity across studies in the metrics used to assess reactivity undermine certainty regarding this prognostic effect. Furthermore, there is little comparative evidence on the prognostic utility of individual tests or combinations of tests. These and other limitations of the existing literature may reduce the potential for clinical application of the tests reviewed here as prognostic markers for patients on clopidogrel-based antiplatelet therapy. The available evidence was insufficient for determining the utility of either type of testing for guiding the choice of antiplatelet therapy.

Limitations of the Evidence

On the basis of the large number of reviewed studies, we believe that the evidence regarding genetic testing for CYP2C19 variants and phenotypic testing for platelet reactivity for guiding antiplatelet treatment and predicting outcomes in patients who receive such treatment is limited in the following ways:

- Despite the large number of available studies providing information on analytic validity, most studies used inappropriate statistical methods to assess interassay agreement.
- There was a lack of comparative studies evaluating the relative predictive ability of alternative assays for measuring platelet reactivity, genetic testing of CYP2C19 variants, or combinations of these tests.
- Development (“training”) and assessment (“test”) samples were not separated when developing predictive markers.
- Selective outcome reporting was a concern regarding the association between test results and several clinical outcomes. Most studies reported information on composite clinical outcomes, but often they did not provide results for the component clinical events.

- There was uncertainty about the genetic model for CYP2C19 variants. Poor reporting of primary study results precluded the assessment of alternative genetic models (e.g., results were often reported only for collapsed genotype categories).
- Exposure definitions were heterogeneous because not all studies genotyped the same CYP2C19 variants and because studies used different assays, metrics, and cutoff values to define increased platelet reactivity.
- There was a paucity of studies evaluating the impact of test-guided treatment selection on the basis of CYP2C19 genotyping or reactivity measurements.
- The number of studies providing information on treatment-effect modification by CYP2C19 genotype status or baseline on-clopidogrel platelet reactivity was limited. Investigations based on completed randomized trials (repurposed RCTs) were not powered to detect treatment-effect modification and were susceptible to selection bias because included patients represented only a minority of the populations included in the parent trials.

Future Research

This review identified substantial gaps in the literature on genetic testing for CYP2C19 variants and phenotypic testing for platelet reactivity, both as biomarkers of future outcomes among patients who are receiving clopidogrel therapy and, more importantly, as tests for guiding treatment selection for patients who are candidates for antiplatelet treatment. We believe that the following evidence gaps may represent fruitful areas for future research:

- *Analytic validity of phenotypic testing:* Future studies using rigorous methods to inform the analytic validity of tests for measuring platelet reactivity are needed, particularly with regard to test-retest reliability, interassay agreement, and analytic performance.
- *Prognostic accuracy, with a focus on comparative prognostic performance:* Large-scale prospectively designed studies of the tests of interest are needed to derive reliable estimates of prognostic performance. Studies should focus on the relative performance of competing tests, prespecify “positive” and “negative” test results, and report complete data for all outcomes assessed.
- *Direct comparisons of methods for test-guided treatment selection:* Even if the predictive value of tests were established, this information is inadequate as a basis for treatment decisionmaking. The most

promising tests could be prioritized for assessment in directly comparative studies of testing versus no testing for guiding treatment choice. Such studies could provide unconfounded estimates of the relative benefits and harms of the compared strategies.⁴⁶ However, randomized comparisons of alternative testing strategies are costly and time consuming. Furthermore, recruitment may be challenging or impossible if one of the treatment groups is standard clopidogrel-based therapy, in view of the current FDA-approved labeling and recent results from studies using pharmacodynamic endpoints. Still, such designs may be appropriate when comparing antiplatelet therapies other than standard clopidogrel dosing (including high-dose clopidogrel treatment). When experimental studies are not considered logistically or ethically feasible, observational data may be useful, especially given that CYP2C19 testing is not universally implemented.

- *“Repurposing” completed randomized trials to assess effect modification:* An alternative to direct comparative studies of testing strategies is to assess effect modification by genotype status by repurposing already completed randomized trials, in which the drugs of interest were tested against a suitable comparator, by genotyping samples from enrollees. Results of genetic analyses could be associated with the prospectively recorded clinical outcomes.^{47,48} Although this approach did not provide definitive answers in this review due to limitations of the existing studies, future repurposed trials could yield more informative results if they were properly planned. Such planning must include a strategy for obtaining samples from all participants (or a random sample thereof), acquiring specimens prior to treatment, and using appropriate methods to control for multiple testing. When randomized trials are not available for repurposing, a similar approach can be implemented in the setting of registries linking DNA information to electronic health records. Patients receiving different antiplatelet therapies whose choice of treatment was not based on CYP2C19 status, but for whom material for genotyping is available, are candidates for such research.
- *Monitoring of platelet reactivity to guide treatment:* Strategies of monitoring platelet reactivity can be conceptualized as “dynamic treatment regimes”⁴⁹⁻⁵¹ (i.e., rules for sequential decisionmaking based on the evolution of reactivity measurements over time). With these methods, the impact of alternative monitoring strategies on clinical outcomes can be evaluated using observational data. The most promising monitoring

strategies can then be evaluated in randomized comparative studies.

Conclusions

In summary, we found limited evidence on the analytic validity of genetic testing for platelet reactivity. However, using evidence from other populations and genetic variants, we believe that the available assays for CYP2C19 genotyping have adequate technical test performance. In contrast, we found a large body of evidence on the analytic validity of assays for measuring platelet reactivity suggesting that interassay agreement is only poor to moderate. No phenotypic assays can be considered a “gold standard” test.

We found some evidence supporting a significant association between loss-of-function CYP2C19 variants and increased risk of stent thrombosis, cardiovascular mortality, and MACE. We also found a significant association between gain-of-function alleles and reduced risk of MACE. The interpretation of these associations should be cautious, given the potential for selective reporting and small-study effects to have affected study results. Furthermore, the applicability of findings to patient populations other than those with ischemic coronary artery disease, particularly those undergoing revascularization procedures, was limited. We also found evidence supporting an association between high on-clopidogrel platelet reactivity as measured by various assays (particularly LTA, VerifyNow P2Y12, and the VASP assay) and adverse cardiovascular events. Our confidence in these findings is limited by the relatively small number of studies available for each test-outcome combination, the potential for selective outcome reporting, and the common lack of separation between the populations used to derive test thresholds of optimal predictive value and those used to assess predictive value at these thresholds.

The evidence on the use of testing to guide treatment choice was insufficient. A single randomized trial of CYP2C19 testing versus no testing provided limited evidence on clinical outcomes. Subanalyses of five well-conducted randomized controlled trials generally did not find strong evidence of effect modification by CYP2C19 status. However, concern regarding selection bias in the genetic substudies and the heterogeneity of patient populations and treatments rendered the evidence inconclusive. Similarly, the short followup periods and low numbers of outcome events in trials of platelet reactivity-guided treatment versus standard antiplatelet therapy did not offer a firm base for conclusions. No studies comparing

genetic and phenotypic testing strategies were identified.

Additional research is needed to better establish the prognostic value and clinical utility for treatment decisionmaking of both genetic testing for CYP2C19 variants and phenotypic testing for platelet reactivity, focusing on standardizing testing methods and assessing the relative impact of testing strategies on patient-relevant clinical outcomes in large well-conducted clinical trials.

References

1. Writing Group Members, Roger VL, Go AS, et al. Executive Summary: Heart Disease and Stroke Statistics—2012 Update. *Circulation*. 2012 Jan 3;125(1):188-97.
2. Keenan NL, Shaw KM. Coronary heart disease and stroke deaths - United States, 2006. *MMWR Surveill Summ*. 2011 Jan 14;60(Suppl):62-6.
3. Squizzato A, Keller T, Romualdi E, et al. Clopidogrel plus aspirin versus aspirin alone for preventing cardiovascular disease. *Cochrane Database Syst Rev*. 2011 Jan;(1):CD005158.
4. Writing Committee, Levine GN, Bates ER, et al. 2011 ACCF/AHA/SCAI Guideline for Percutaneous Coronary Intervention. *Circulation*. 2011 Dec 6;124(23):e574-e651.
5. Bowry AD, Brookhart MA, Choudhry NK. Meta-analysis of the efficacy and safety of clopidogrel plus aspirin as compared to antiplatelet monotherapy for the prevention of vascular events. *Am J Cardiol*. 2008 Apr 1;101(7):960-6.
6. Sabatine MS, Hamdalla HN, Mehta SR, et al. Efficacy and safety of clopidogrel pretreatment before percutaneous coronary intervention with and without glycoprotein IIb/IIIa inhibitor use. *Am Heart J*. 2008 May;155(5):910-7.
7. Sudlow CL, Mason G, Maurice JB, et al. Thienopyridine derivatives versus aspirin for preventing stroke and other serious vascular events in high vascular risk patients. *Cochrane Database Syst Rev*. 2009 Oct 7;(4):CD001246.
8. Antithrombotic Trialists' Collaboration. Collaborative meta-analysis of randomised trials of antiplatelet therapy for prevention of death, myocardial infarction, and stroke in high risk patients. *BMJ*. 2002 Jan 12;324(7329):71-86.
9. ACTIVE Investigators, Connolly SJ, Pogue J, et al. Effect of clopidogrel added to aspirin in patients with atrial fibrillation. *N Engl J Med*. 2009 May 14;360(20):2066-78.
10. Bellemain-Appaix A, Brieger D, Beygui F, et al. New P2Y12 inhibitors versus clopidogrel in percutaneous coronary intervention: a meta-analysis. *J Am Coll Cardiol*. 2010 Nov 2;56(19):1542-51.
11. Wallentin L, Becker RC, Budaj A, et al. Ticagrelor versus clopidogrel in patients with acute coronary syndromes. *N Engl J Med*. 2009 Sep 10;361(11):1045-57.
12. Wiviott SD, Braunwald E, McCabe CH, et al. Prasugrel versus clopidogrel in patients with acute coronary syndromes. *N Engl J Med*. 2007 Nov 15;357(20):2001-15.
13. Navarese EP, Verdoia M, Schaffer A, et al. Ischaemic and bleeding complications with new, compared to standard, ADP-antagonist regimens in acute coronary syndromes: a meta-analysis of randomized trials. *QJM*. 2011 Jul 1;104(7):561-9.
14. Geiger J, Brich J, Honig-Liedl P, et al. Specific impairment of human platelet P2Y(AC) ADP receptor-mediated signaling by the antiplatelet drug clopidogrel. *Arterioscler Thromb Vasc Biol*. 1999 Aug;19(8):2007-11.
15. Kazui M, Nishiya Y, Ishizuka T, et al. Identification of the human cytochrome P450 enzymes involved in the two oxidative steps in the bioactivation of clopidogrel to its pharmacologically active metabolite. *Drug Metab Dispos*. 2010 Jan;38(1):92-9.
16. Bauer T, Bouman HJ, van Werkum JW, et al. Impact of CYP2C19 variant genotypes on clinical efficacy of antiplatelet treatment with clopidogrel: systematic review and meta-analysis. *BMJ*. 2011;343:d4588.
17. Chang MH, Lindgren ML, Butler MA, et al. Prevalence in the United States of selected candidate gene variants: Third National Health and Nutrition Examination Survey, 1991-1994. *Am J Epidemiol*. 2009 Jan 1;169(1):54-66.
18. Chan MY, Tan K, Tan HC, et al. CYP2C19 and PON1 polymorphisms regulating clopidogrel bioactivation in Chinese, Malay and Indian subjects. *Pharmacogenomics*. 2012 Apr;13(5):533-42.
19. Yin SJ, Ni YB, Wang SM, et al. Differences in genotype and allele frequency distributions of polymorphic drug metabolizing enzymes CYP2C19 and CYP2D6 in mainland Chinese Mongolian, Hui and Han populations. *J Clin Pharm Ther*. 2012 Jun;37(3):364-9.
20. Veiga MI, Asimus S, Ferreira PE, et al. Pharmacogenomics of CYP2A6, CYP2B6, CYP2C19, CYP2D6, CYP3A4, CYP3A5 and MDR1 in Vietnam. *Eur J Clin Pharmacol*. 2009 Apr;65(4):355-63.
21. Shuldiner AR, O'Connell JR, Bliden KP, et al. Association of cytochrome P450 2C19 genotype with the antiplatelet effect and clinical efficacy of clopidogrel therapy. *JAMA*. 2009 Aug 26;302(8):849-57.
22. Bonello L, Tantry US, Marcucci R, et al. Consensus and future directions on the definition of high on-treatment platelet reactivity to adenosine diphosphate. *J Am Coll Cardiol*. 2010 Sep 14;56(12):919-33.
23. Agency for Healthcare Research and Quality. Methods Guide for Effectiveness and Comparative Effectiveness Reviews. AHRQ Publication No. 10(11)-EHC063-EF. Rockville (MD): Agency for Healthcare Research and Quality; March 2011. Chapters available at www.effectivehealthcare.ahrq.gov.
24. Viswanathan M, Ansari MT, Berkman ND, et al. Assessing the risk of bias of individual studies in systematic reviews of health care interventions. In: *Methods Guide for Effectiveness and Comparative Effectiveness Reviews*. AHRQ Publication No. 10(11)-EHC063-EF. Rockville (MD): Agency for Healthcare Research and Quality; March 2011. Chapters available at www.effectivehealthcare.ahrq.gov.
25. Agency for Healthcare Research and Quality. Methods Guide for Medical Test Reviews. AHRQ Publication No. 12-EC017. Rockville, MD: Agency for Healthcare Research and Quality; June 2012.

26. Whiting PF, Rutjes AWS, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med.* 2011 Oct 18;155(8):529-36.
27. Whiting P, Rutjes AW, Reitsma JB, et al. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Med Res Methodol.* 2003 Nov 10;3:25.
28. Whiting P, Rutjes AW, Dinnes J, et al. Development and validation of methods for assessing the quality of diagnostic accuracy studies. *Health Technol Assess.* 2004 Jun;8(25):iii, 1-iii234.
29. Whiting PF, Westwood ME, Rutjes AWS, et al. Evaluation of QUADAS, a tool for the quality assessment of diagnostic accuracy studies. *BMC Med Res Methodol.* 2006;6:9.
30. Higgins JPT, Douglas GA, Peter CG, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ.* 2011 Oct 18;343.
31. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials.* 1986 Sep;7(3):177-88.
32. Cochran WG. The combination of estimates from different experiments. *Biometrics.* 1954 Mar;10(1):101-29.
33. Higgins JPT, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Statist Med.* 2002;21(11):1539-58.
34. Pearce N. What does the odds ratio estimate in a case-control study? *Int J Epidemiol.* 1993 Dec;22(6):1189-92.
35. Knol MJ, Vandenbroucke JP, Scott P, et al. What do case-control studies estimate? Survey of methods and assumptions in published case-control research. *Am J Epidemiol.* 2008 Nov 1;168(9):1073-81.
36. Miettinen OS. *Theoretical Epidemiology: Principles of Occurrence Research in Medicine.* New York: Wiley; 1985.
37. Egger M, Davey SG, Schneider M, et al. Bias in meta-analysis detected by a simple, graphical test. *BMJ.* 1997 Sep 13;315(7109):629-34.
38. Lau J, Ioannidis JP, Terrin N, et al. The case of the misleading funnel plot. *BMJ.* 2006 Sep 16;333(7568):597-600.
39. Owens DK, Lohr KN, Atkins D, et al. AHRQ series paper 5: grading the strength of a body of evidence when comparing medical interventions--Agency for Healthcare Research and Quality and the effective health-care program. *J Clin Epidemiol.* 2010 May;63(5):513-23.
40. Atkins D, Chang SM, Gartlehner G, et al. Assessing applicability when comparing medical interventions: AHRQ and the Effective Health Care Program. *J Clin Epidemiol.* 2011;64:1198-207.
41. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet.* 1986 Feb 8;1(8476):307-10.
42. Lin LI. A concordance correlation coefficient to evaluate reproducibility. *Biometrics.* 1989;45(1):255-68. PMID: 2720055.
43. Holmes DR Jr, Dehmer GJ, Kaul S, et al. ACCF/AHA clopidogrel clinical alert: approaches to the FDA "boxed warning": a report of the American College of Cardiology Foundation Task Force on clinical expert consensus documents and the American Heart Association endorsed by the Society for Cardiovascular Angiography and Interventions and the Society of Thoracic Surgeons. *J Am Coll Cardiol.* 2010 Jul;56(4):321-41.
44. Holmes MV, Perel P, Shah T, et al. CYP2C19 genotype, clopidogrel metabolism, platelet function, and cardiovascular events. *JAMA.* 2011 Dec 28;306(24):2704-14.
45. Brar SS, ten Berg J, Marcucci R, et al. Impact of platelet reactivity on clinical outcomes after percutaneous coronary intervention: a collaborative meta-analysis of individual participant data. *J Amer Coll Cardiol.* 2011 Nov 1;58(19):1945-54.
46. Lord SJ, Irwig L, Simes RJ. When is measuring sensitivity and specificity sufficient to evaluate a diagnostic test, and when do we need randomized trials? *Ann Intern Med.* 2006;144(11):850-5.
47. Dahabreh IJ, Terasawa T, Castaldi PJ, et al. Systematic review: anti-epidermal growth factor receptor treatment effect modification by KRAS mutations in advanced colorectal cancer. *Ann Intern Med.* 2011 Jan 4;154(1):37-49.
48. Simon RM, Paik S, Hayes DF. Use of archived specimens in evaluation of prognostic and predictive biomarkers. *J Natl Cancer Inst.* 2009 Nov 4;101(21):1446-52.
49. Cain LE, Robins JM, Lanoy E, et al. When to start treatment? A systematic approach to the comparison of dynamic regimes using observational data. *Int J Biostat.* 2010;6(2):18.
50. Hernan MA, Lanoy E, Costagliola D, et al. Comparison of dynamic treatment regimes via inverse probability weighting. *Basic Clin Pharmacol Toxicol.* 2006;98(3):237-42.
51. Orellana L, Rotnitzky A, Robins JM. Dynamic regime marginal structural mean models for estimation of optimal dynamic treatment regimes, Part I: main content. *Int J Biostat.* 2010;6(2):8.

Full Report

This executive summary is part of the following document: Dahabreh IJ, Moorthy D, Lamont JL, Chen ML, Kent DM, Lau J. Testing of CYP2C19 Variants and Platelet Reactivity for Guiding Antiplatelet Treatment. Comparative Effectiveness Review No. 125. (Prepared by Tufts Evidence-based Practice Center under Contract No. 290-2007-10055-I.) AHRQ Publication No. 13-EHC117-EF. Rockville, MD: Agency for Healthcare Research and Quality; September 2013. www.effectivehealthcare.ahrq.gov/reports/final/cfm.

