SUPPLEMENTAL MATERIAL
Data S1. Supplemental Methods.

PubMed Search Strategy

Estimation of Event Rate Curves for Individual Events within a Composite

Breaking Broader Events into Endpoints of Interest

Full Mathematical Specification of the Model

Estimation of Model Parameters

Clinical Benefit Calculator Prototype

Additional Details Regarding Scenario Analysis with ASCEND

Table S1. List of excluded trials and reason for exclusion

Table S2. Parameters retained in the model

Figure S1. Selection criteria of trials considered in model development

Figure S2. Estimated and trial-reported hazard ratios: comparison of final model and model without use of parameter for high baseline high-sensitivity C-reactive protein

Figure S3. Model-predicted vs. trial-reported hazard ratios for unstable angina requiring hospitalization and coronary revascularization

Figure S4. Estimated instantaneous relative risk reduction, $\alpha$, over time for non-fatal MI by lipid-lowering therapy type and 1 mmol/L reduction in LDL-C

Figure S5. Prototype of clinical benefit calculator
Data S1. Supplemental Methods

PubMed Search Strategy

363 results:

"lipid regulating agents"[Pharmacological Action] AND "randomized controlled trial"[Publication Type] AND "cholesterol"[MeSH Terms] AND "LDL"[All Fields] AND "cardiovascular"[All Fields] AND ("anacetrapib"[All Fields] OR "ezetimibe"[All Fields] OR "PCSK9"[All Fields] OR "statin"[All Fields]) AND ("1995"[PDAT] : "2019"[PDAT])

Estimation of Event Rate Curves for Individual Events within a Composite

The following approach was adopted for the estimation of event rates for an endpoint of interest, $E_i(t)$, from the composite event rate curve, $E_T(t)$, and additional data reported in the publication. We first invoke the identity that instantaneous hazard rates for parts of the composite are additive:

$$\lambda_T(t) = \lambda_1(t) + \cdots + \lambda_N(t)$$

Integrating both sides and invoking the identity $\int_0^t \lambda(\tau)d\tau = -\ln (1 - E(t))$ we get:

$$\ln \left(1 - E_T(t) \right) = \ln \left(1 - E_1(t) \right) + \cdots + \ln \left(1 - E_N(t) \right)$$

For a specific time $T$ (e.g. median follow-up or total trial duration) this becomes:

$$\ln \left(1 - E_T(T) \right) = \ln \left(1 - E_1(T) \right) + \cdots + \ln \left(1 - E_N(T) \right)$$

We can define a ratio $P_i$ as:

$$P_i = \frac{\ln \left(1 - E_i(T) \right)}{\ln \left(1 - E_T(T) \right)}$$

The value $P_i$ can be estimated from additional data reported in publications and then utilized to estimate the event rate curves for an endpoint of interest, $E_i(t)$, from the composite event rate curve, $E_T(t)$, as follows:
\[ E_i(t) = 1 - \exp\left( P_i \ln\left( 1 - E_T(t) \right) \right) \]

If applicable, \( P_i \)'s were renormalized to ensure the estimated \( E_i(t) \)'s of interest were equal to the data reported in the publication (e.g. the nonfatal myocardial infarction [MI] endpoint in ODYSSEY OUTCOMES\textsuperscript{46}) at the follow-up time. If a trial had coronary revascularization as a part of its composite endpoint, these individual event rates were re-estimated to reflect elective revascularizations not as a part of other cardiovascular (CV) events. Once event rates for the individual endpoints of interest, \( E_i(t) \), were estimated in this manner, we confirmed that we were able to replicate reported composite curves, \( E_T(t) \), via a simulation of first events from estimated \( E_i(t) \).

As an example, in the IMPROVE-IT trial,\textsuperscript{42} the event rate for nonfatal MI in the simvastatin monotherapy arm was 14.4\%, which translates to \( \ln\left( 1 - E_i(T) \right) = -0.155 \). The concurrent estimates for \( \ln\left( 1 - E_i(T) \right) \) for CV death, nonfatal stroke, coronary revascularization, and unstable angina [UA] requiring hospitalization were –0.070, –0.043, –0.267, and –0.019, respectively. Using these numbers and an estimate of \( \ln\left( 1 - E_T(T) \right) = 0.426 \), coronary revascularization was re-estimated as –0.138. Thus, \( P_i \) for the nonfatal MI endpoint was estimated as 0.155/0.426 = 0.365, and the event rates were estimated as \( E_i(t) = 1 - \exp\left( 0.365 \ln\left( 1 - E_T(t) \right) \right) \), where \( E_T(t) \) is composite event rate curve reported in IMPROVE-IT.

**Breaking Broader Events into Endpoints of Interest**

A trial may not report the exact event of interest. For example, in some trials the risks for coronary heart disease (CHD) death and ischemic stroke are not reported as a separate event, but instead they are part of a broader CV event or stroke event. It was essential in model development to convert these into a consistent set of endpoints such as CHD death and ischemic stroke. To facilitate this, we estimated the ratio of rates \( P \) for a specific event type A (e.g. CHD death or ischemic stroke) to a broader event type B (e.g. death from other causes or any stroke) as \( P = \ln(1 - E_A)/\ln(1 - E_B) \) by utilizing the data from Cholesterol Treatment Trialists’ (CTT) 2010 meta-analysis.\textsuperscript{5}
| Event Type                              | Estimated Ratio of Rates (P) |
|-----------------------------------------|-----------------------------|
|                                        | Placebo | Treated |
| CHD Death / Cardiac Death               | 0.544   | 0.554   |
| CHD Death / CV Death                    | 0.460   | 0.415   |
| CHD Death / Any Death                   | 0.259   | 0.236   |
| Nonfatal Ischemic Stroke / Any Stroke   | 0.624   | 0.666   |

As an example, in the JUPITER trial, the number of patients experiencing any stroke in the rosuvastatin and placebo arms was 33 and 64, respectively. Using the table above, we re-estimated these numbers as nonfatal ischemic stroke in the rosuvastatin and placebo arms as 21 and 43, respectively.

**Full Mathematical Specification of the Model**

We define the concept of an instantaneous relative risk reduction ($\alpha$), which represents percent reduction in events with treatment at a particular moment in time, $t$, (e.g. at 3 years) over a small incremental follow-up, $dt$. If $dE(t)$ and $dE_c(t)$ denote the incremental number of events in treatment and control populations, respectively, at time $t$, then $\alpha(t) = (dE_c(t) - dE(t))/dE_c(t)$. We similarly define an instantaneous risk ($\lambda$) as percent of population experiencing an event over a small incremental follow-up of $dt$. Thus $\lambda(t) = (dE(t)/S(t))/dt$, where $S(t) = 1 - E(t)$ represents the size of population at risk of events at time $t$ (normalization by $dt$ is required for consistency in mathematical framing), and the idea behind $\lambda(t)$ is identical to that of an instantaneous hazard rate. It follows from these definitions that $\lambda(t) = \lambda_c(t) (1 - \alpha(t))$. As a generalization for the setting of lipid-lowering therapies (LLTs), we specify the instantaneous relative risk reduction, $\alpha$, to depend on time since initiation of LLT ($t$), the magnitude of low-density lipoprotein cholesterol (LDL-C) reduction in mmol/L ($\Delta L$), and additional patient characteristics ($X$). The full functional form for $\alpha$ was specified as:

$$\alpha(t, \Delta L, X) = 1 - (1 - \theta(t, \Delta L, X))^{\Delta L}$$
This functional form for $\alpha$ exhibits the correct limiting behavior with regards to the magnitude of reduction in LDL-C, $\Delta L$ (with $\theta$ between 0 and 1), meaning that as when $\Delta L = 0$, then $\alpha = 0$, and as $\Delta L$ increases, $\alpha$ increases, approaching a limiting value of 1. The parameter $\theta$ can be interpreted as the instantaneous risk reduction per 1 mmol/L reduction in LDL-C. As an example, if $\theta = 0.25$, and $\Delta L = 0.5$, then $\alpha = 0.13$. In other words, if the instantaneous risk reduction per 1 mmol/L reduction in LDL-C is estimated as 25%, then with an LDL-C reduction of 0.5 mmol/L, the instantaneous risk reduction would be estimated as 13%. The overall functional form for $\theta$ has to be chosen such that it lies between 0 and 1. We postulate the parameter $\theta$ can be partitioned into multiplicative parts $\theta_1$ and $\theta_2$ (with both $\theta_1$ and $\theta_2$ bound between 0 and 1), as follows:

$$\theta(t, \Delta L, X) = \theta_1(X) \theta_2(t, \Delta L, X)$$

Where $\theta_1(X)$ captures the saturation effect with long duration of treatment and $\theta_2(t, \Delta L, X)$ captures the transient effect leading to the saturation effect. The parameter $\theta_1(X)$ can be modelled via a generic logistic transformation as:

$$\theta_1(X) = \frac{1}{1 + \exp(-X\beta)}$$

Where $X\beta$ represents the linear combination $\beta_0 + \beta_1 X_1 + \cdots + \beta_N X_N$. The parameter $\theta_2(t, \Delta L, X)$ capturing the time-dependent part can be modelled via a blend of exponential terms:

$$\theta_2(t, \Delta L, X) = \pi (1 - \exp(- t \exp(X\gamma))) + (1 - \pi)(1 - \exp(-t\phi))$$

Where the second term represents a generic growth (or decay) model with a saturation value of 1, and the first term enables modeling of a change in the steepness of initial growth. The parameter $\pi$ is bound between 0 and 1. To enforce this condition, a logistic transformation was again used such that $\pi = 1/(1 + \exp(\omega))$, where $\omega$ is a global parameter that is free of constraints, and $\phi$ is an additional global parameter. The term $X\gamma$ represents the linear combination $\gamma_0 + \gamma_1 X_1 + \cdots + \gamma_M X_M$. Thus, the final formulation of $\alpha(t, \Delta L, X)$ is:
\[
\alpha(t, \Delta L, X) = 1 - \left( 1 - \frac{(1 - \exp(-t \exp(X \gamma))) + \exp(\omega)(1 - \exp(-t \phi))}{(1 + \exp(\omega))(1 + \exp(-X \beta))} \right)^{\Delta L}
\]

**Estimation of Model Parameters**

A cost function \( F \) was defined as the sum of squares of the error between actual event rate curves with treatment from randomized controlled trials (RCTs) and the model-predicted event rate curves for individual endpoints, averaged over the trial duration. The cost function \( F \) was minimized via the Broyden-Fletcher-Goldfarb-Shanno (BFGS) optimization algorithm in Python, which returned a set of estimated model parameters, \( \beta \).\(^{47}\) The estimated parameters with confidence intervals (CIs) for the \( \alpha(t, \Delta L, X) \) function are summarized in Table S2. When these confidence intervals were considered, five trials had better predictions with the model, and one had better prediction with the CTT estimation for a 3-part composite of nonfatal MI, ischemic stroke, and CHD death.

The covariates considered were individual endpoint types (indicator variables for non-fatal MI, ischemic stroke, CHD death, UA hospitalization, and coronary revascularization), LLT type (indicator variables for statin, ezetimibe, proprotein convertase subtilisin/kexin type 9 [PCSK9] inhibitor, and anacetrapib), established atherosclerotic cardiovascular disease (ASCVD) status, diabetes status, trial mean age, baseline LDL-C level, trial proportion female, high baseline high-sensitivity C-reactive protein (hsCRP) levels (variable relevant only for the JUPITER trial\(^{40}\)), and established ischemic cerebrovascular disease (variable relevant only for the SPARCL trial\(^{39}\)).

For each covariate of interest, we tried both \( X \beta \) and \( X \gamma \) terms and retained the covariate in the term that maximized model performance. For example, LLT types distinguishing PCSK9 inhibitors were retained in the \( X \gamma \) term. The CIs for retained model parameters and model-predicted hazard ratios were generated via the bootstrap method where hazard ratios were probabilistically sampled 1000 times from reported CIs in selected RCTs. Note the CIs for some parameter estimates overlap zero. In light of the mathematical framing, zero values for several model parameters such as \( \beta_0, \gamma_0, \gamma_2, \omega, \) and \( \phi \) convey a non-zero effect on the model.
Hence, the estimates of these parameters are valid even if the CIs overlap zero. For other parameters, retaining them in the model was critical regardless of whether or not the CIs overlapped zero in order to ensure the overall model performed as described in the methods section (we confirmed dropping any of the variables in Table S2 deteriorated the overall model performance). This approach is consistent with the principles of model development when the aim is to maximize the overall model performance. Figure S4 provides the summary behavior of instantaneous risk reduction function, $\alpha$, over time via the estimated model parameters for the non-fatal MI endpoint by LLT types, and 1 mmol/L reduction in LDL-C.

We have included the data point corresponding to the evidence from Mendelian randomization analysis from Ference et al.\textsuperscript{24} in this Figure, which illustrates that the behavior of $\alpha$ with long duration of treatment is in excellent agreement with the Mendelian randomization data. Finally, the cumulative event rates curves over time (corresponding to the Kaplan-Meier curves) can be estimated from $\alpha(t, \Delta L, X)$ and the control population risk $\lambda_c(t)$ as:

$$E(t) = 1 - \exp \left( - \int_0^t \left( 1 - \alpha(t, \Delta L, X) \right) \lambda_c(\tau) \, d\tau \right)$$

A simple way of modeling the control population risk, $\lambda_c(t)$, is via a constant hazard model, in which case it can be estimated as $-\ln(1 - E(t))/t$. As an example, if the 2-year risk is 10%, then $\lambda = -\ln(1 - 0.1)/2 = 0.05268$. In the case that a constant hazard model does not adequately describe the risk over time (i.e. the risk changes over time), a function of type $\lambda_c(t) = A + B \exp(-Ct)$ can be used, and is the one we have utilized in modeling the control population risk in RCTs. This function covers a range of behaviors, such as initial elevation of risk (e.g. in trials including recent acute coronary syndrome [ACS] population) that gradually declines over time to a more constant risk (e.g. representing a stable CHD risk profile). It also has the flexibility to model a constant risk over time (i.e. $B = 0$) to capture a stable risk profile.
Clinical Benefit Calculator Prototype

A treatment benefit calculator based on the estimated model can be developed and implemented via an online tool or Microsoft Excel to provide an easy-to-use interface. An example of a prototype is provided in Figure S5. The calculator can rely on five inputs from the user: (1) time frame for which the treatment benefit estimate is desired; (2) estimated risk before treatment; (3) estimated reduction in LDL-C via LLT (this can be estimated from published evidence on LDL-C lowering efficacy from a given LLT); (4) LLT type; and (5) high hsCRP status. The outputs of the calculator include an estimate of risk over time with and without treatment, absolute risk reduction, and number needed to treat.

Additional Details Regarding Scenario Analysis with ASCEND

Complete information regarding baseline LDL-C levels and background LLT were not available for the ASCEND trial. It was reported that 75% of patients in the ASCEND trial were receiving a statin at baseline. We estimated the statin potency for these 75% by utilizing published data on the relative proportion of diabetes without ASCVD patients receiving statins in the UK on low, moderate, and high-intensity statins. We then estimated the overall mean LDL-C for the ASCEND population by utilizing data on achieved LDL-C by these groups (low, moderate, and high-intensity statins, and no statin) from published data.
Table S1. List of excluded trials and reason for exclusion

| Trial    | Year | Treatment          | Comparator | Reason for Exclusion                                              |
|----------|------|--------------------|------------|------------------------------------------------------------------|
| ALLHAT-LLT | 2002 | Pravastatin 40 mg | Placebo    | Open label                                                        |
| ALLIANCE | 2004 | Atorvastatin       | Placebo    | Open label                                                        |
| GISSI-P  | 2000 | Pravastatin 20 mg  | Placebo    | Open label                                                        |
| MEGA     | 2006 | Pravastatin 10-20 mg | Placebo | Open label; data reported as rates instead of KM curves  |
| AFCAPS   | 1998 | Lovastatin 20-40 mg | Placebo    | Data not available to enable breakdown of KM curves by events     |
| AURORA   | 2009 | Rosuvastatin 10 mg | Placebo    | Population with end stage renal disease                          |
| ALERT    | 2003 | Fluvastatin 40 mg  | Placebo    | Population with end stage renal disease                          |
| 4D       | 2005 | Atorvastatin 20 mg | Placebo    | Population with end stage renal disease                          |
| SHARP    | 2011 | Simvastatin 20 mg + ezetimibe | Placebo | High proportion of patients with end stage renal disease |
| CORONA   | 2007 | Rosuvastatin 10 mg | Placebo    | Population with heart failure                                    |
| SEAS     | 2008 | Simvastatin 40 mg + ezetimibe | Placebo | Population with aortic stenosis                                  |
| SPIRE    | 2017 | Bococizumab        | Placebo    | Trial discontinued due to antidrug antibodies                    |
| SSSS     | 1994 | Simvastatin        | Placebo    | Technical issues with digitization of KM curves                  |
| POST-CABG| 1997 | Lovastatin 40-80 mg | Lovastatin 2.5-5 mg | Two-by-two factorial design not conducive for model estimation  |
| GISSI-HF | 2008 | Rosuvastatin 10 mg | Placebo    | Technical issues with digitization of KM curves                  |
| GREACE   | 2002 | Atorvastatin 10-80 mg | Usual Care | Data reported as rates instead of KM curves                     |

KM, Kaplan-Meier; ALLHAT-LLT, Antihypertensive and Lipid-Lowering Treatment to Prevent Heart Attack Trial; ALLIANCE, Aggressive Lipid-Lowering Initiation Abates New Cardiac Events; GISSI-P, Gruppo Italiano per lo Studio della Sopravvivenza nell’Infarto Miocardio; MEGA, Management of Elevated Cholesterol in the Primary Prevention Group of Adult Japanese; AFCAPS/TexCAPS, Air Force/Texas Coronary Atherosclerosis Prevention Study; AURORA, A Study to Evaluate the Use of Rosuvastatin in Subjects on Regular Hemodialysis; ALERT, Assessment of Lescol in Renal Transplantation; 4D, Die Deutsche Diabetes Dialyse Studie; SHARP, Study of Heart and Renal Protection; CORONA, Controlled Rosuvastatin in Multinational Trial in Heart Failure; SEAS, Simvastatin and Ezetimibe in Aortic Stenosis.
SPIRE, Studies of PCSK9 Inhibition and the Reduction of Vascular Events\textsuperscript{19}; SSSS, Scandinavian Simvastatin Survival Study\textsuperscript{20}; POST-CABG, Post-Coronary Artery Bypass Graft\textsuperscript{21}; GISSI-HF, Gruppo Italiano per lo Studio della Sopravvivenza nell’Insufficienza cardiaca\textsuperscript{22}; GREACE, GREek Atorvastatin and Coronary-heart-disease Evaluation.\textsuperscript{23}
| Parameter | Description | Value (95% CI) |
|-----------|-------------|---------------|
| $\beta_0$ | Intercept term for $X\beta$ (MI) | 0.521 (0.163, 0.833) |
| $\beta_1$ | UA requiring hospitalization | -0.731 (-2.035, 0.730) |
| $\beta_2$ | CHD Death | -1.084 (-1.761, -0.478) |
| $\beta_3$ | Ischemic Stroke | -0.986 (-1.821, -0.254) |
| $\beta_4$ | Coronary Revascularization | -0.269 (-0.908, 0.411) |
| $\beta_5$ | High hsCRP levels | 1.687 (0.722, 3.781) |
| $\gamma_0$ | Intercept term for $X\gamma$ (statin) | 0.136 (-1.436, 0.893) |
| $\gamma_1$ | PCSK9 inhibitor | -1.871 (-3.616, -0.745) |
| $\gamma_2$ | Impact of $\Delta L$ on earlier risk reduction (early separation of event rate curves) | 1.197 (-0.523, 2.638) |
| $\omega$ | Global parameter | -0.038 (-0.633, 0.367) |
| $\phi$ | Global parameter | 0.031 (0.016, 0.153) |

Values listed represent the median based on a sensitivity analysis rather than the point estimated model parameters. PCSK9 inhibitor parameter for PCSK9 inhibitors and anacetrapib therapy.

CHD, coronary heart disease; CI, confidence interval; hsCRP, high-sensitivity C-reactive protein; MI, myocardial infarction; PCSK9, proprotein convertase subtilisin/kexin type 9; UA, unstable angina.
Figure S1. Selection criteria of trials considered in model development

PCSK9; proprotein convertase subtilisin/kexin type 9

*Studies with lipid-lowering therapy (statins, ezetimibe, PCSK9 inhibitors, and anacetrapib), at least 1000 patients, an endpoint of cardiovascular events or mortality. †Data reported as rates and not as Kaplan-Meier curves over time, not possible to digitize published Kaplan-Meier curves, and factorial design resulting in limitations in appropriate data abstraction. ‡Populations with end stage renal disease, heart failure, or aortic stenosis.
Figure S2. Estimated and trial-reported hazard ratios: comparison of final model and model without use of parameter for high baseline high-sensitivity C-reactive protein

Hazard ratios were calculated for a composite of non-fatal myocardial infarction, ischemic stroke, and coronary heart disease death. Letters a to v denote the following trials: a, A to Z; b, ASCOT-LLA; c, ASPEN; d, CARDS; e, CARE; f, FOURIER; g, HOPE; h, HPS; i, IDEAL; j, IMPROVE-IT; k, JUPITER; l, LIPID; m, LIPS; n, MIRACL; o, ODYSSEY OUTCOMES; p, PROSPER; q, PROVE-IT; r, REVEAL; s, SEARCH; t, SPARCL; u, TNT; v, WOSCOPS.

hsCRP, high-sensitivity C-reactive protein
Figure S3. Model-predicted vs. trial-reported hazard ratios for unstable angina requiring hospitalization and coronary revascularization

Coronary revasc., coronary revascularization; HR, hazard ratio; UA hosp., unstable angina requiring hospitalization
Figure S4. Estimated instantaneous relative risk reduction, $\alpha$, over time for non-fatal MI by lipid-lowering therapy type and 1 mmol/L reduction in LDL-C

Instantaneous relative risk reduction is the relative risk reduction at a specific moment in time. Dotted line indicates estimates for PCSK9 inhibitors and anacetrapib. LDL-C, low-density lipoprotein cholesterol; MI, myocardial infarction; PCSK9, proprotein convertase subtilisin/kexin type 9.
ACS, acute coronary syndromes; ARR, absolute risk reduction; ASCVD, atherosclerotic cardiovascular disease; CHD, coronary heart disease; MI, myocardial infarction; NNT, number needed to treat; RRR, relative risk reduction