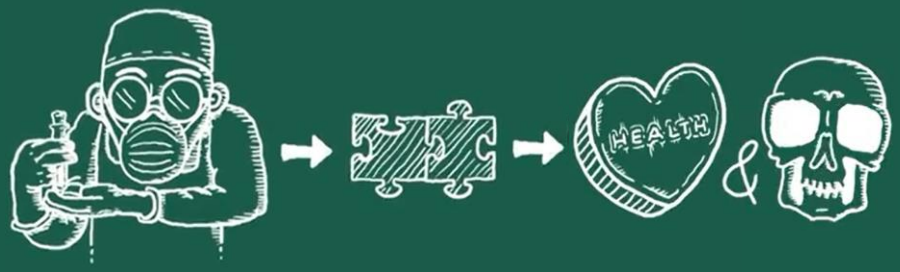
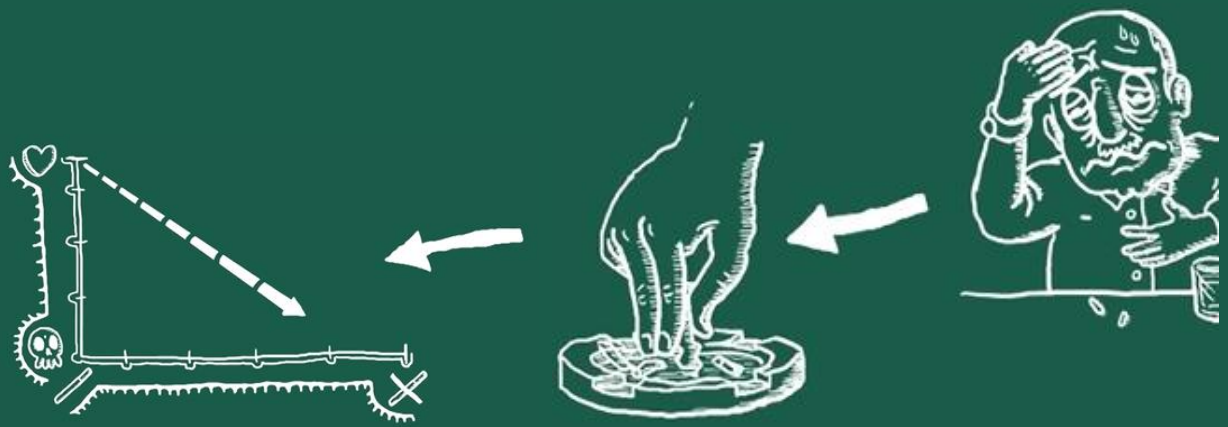
Mendelian randomization

Soyeon Kim, Wonlab

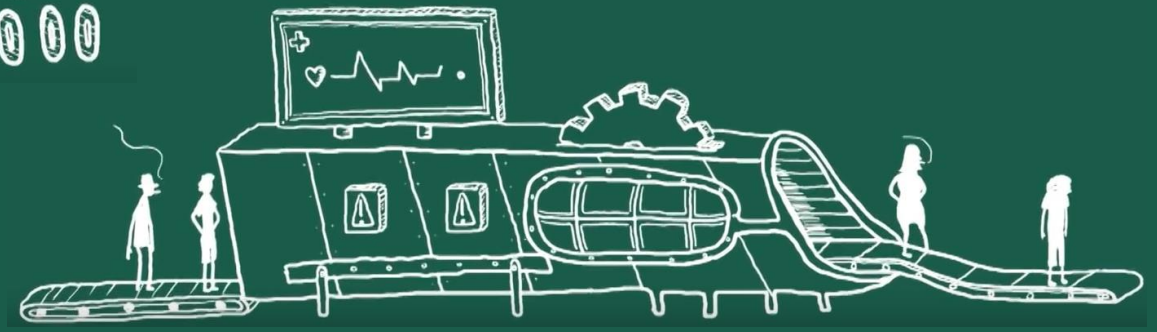
November 17, 2018


MENDELIAN
RANDOMISATION



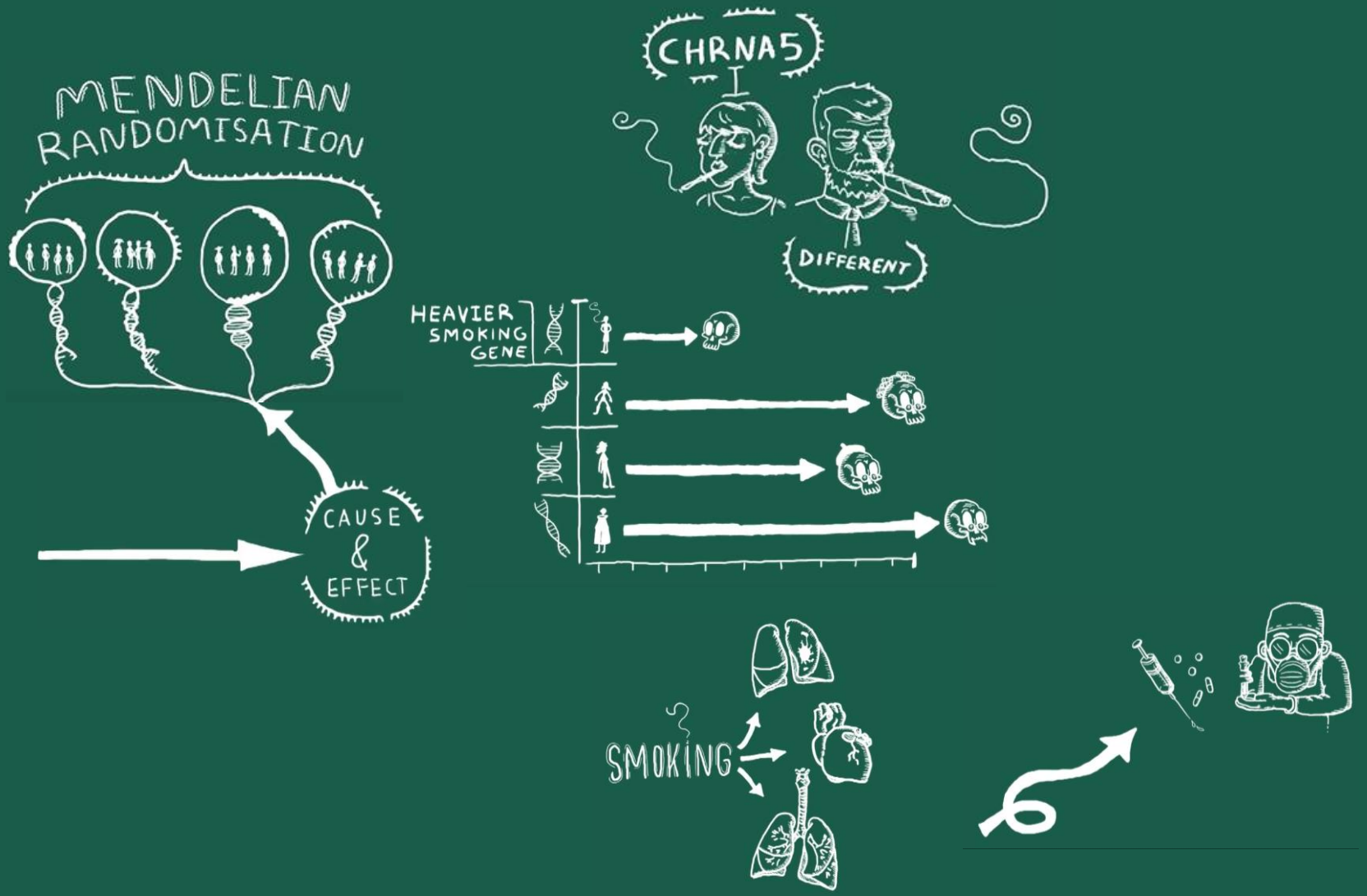


https://www.youtube.com/watch?v=LoTgfGotaQ4



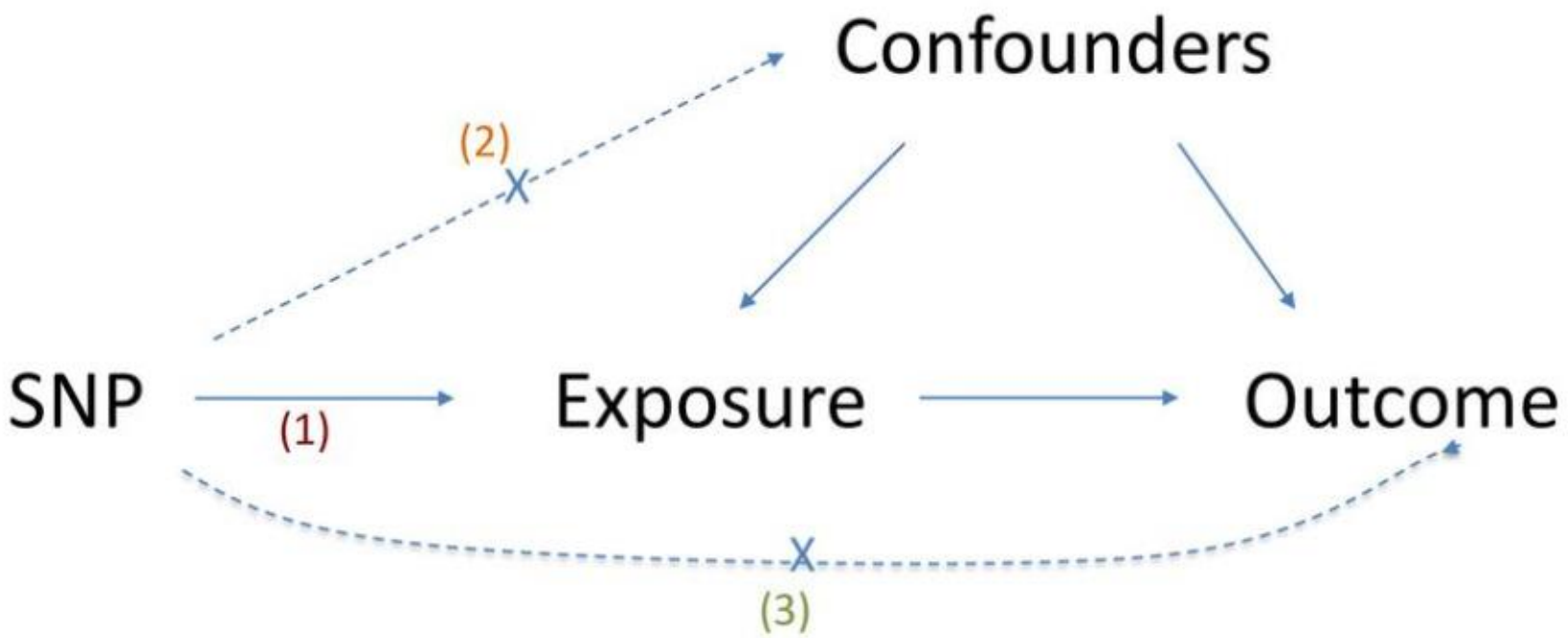
https://www.youtube.com/watch?v=LoTgfGotaQ4

https://www.youtube.com/watch?v=LoTgfGotaQ4



https://www.youtube.com/watch?v=LoTgfGotaQ4

- **Mendelian randomization**

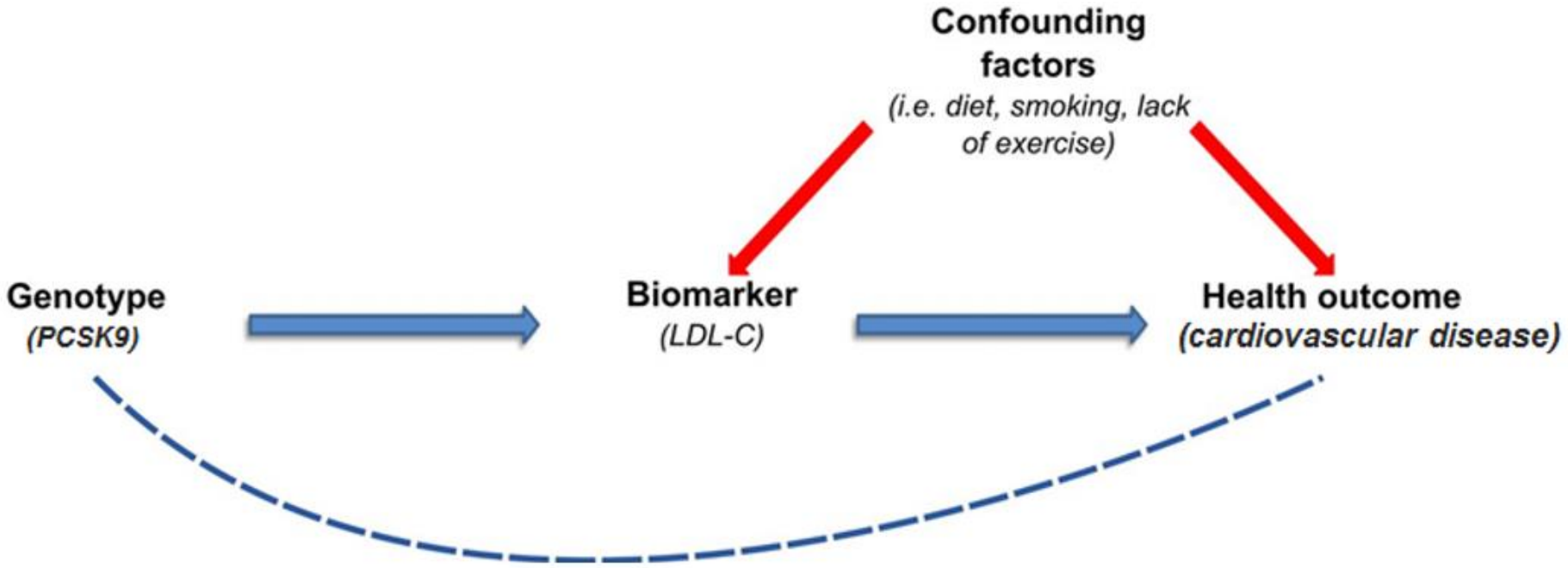


- **Instrumental variable assumptions**

(1) SNP is associated with the exposure

(2) SNP is NOT associated with confounding variables

(3) SNP ONLY associated with outcome through the exposure


Confounding factors
(i.e. diet, smoking, lack of exercise)
Genotype
(PCSK9)
Biomarker
(LDL-C)
Health outcome
(cardiovascular disease)

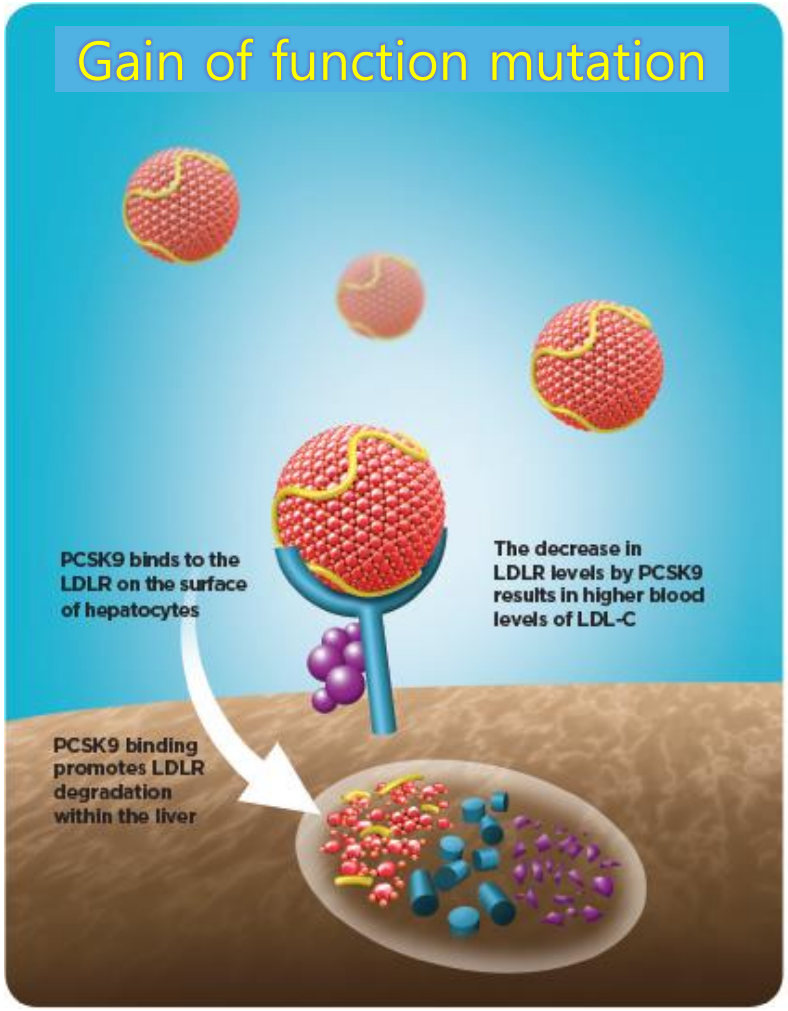
The NEW ENGLAND JOURNAL of MEDICINE

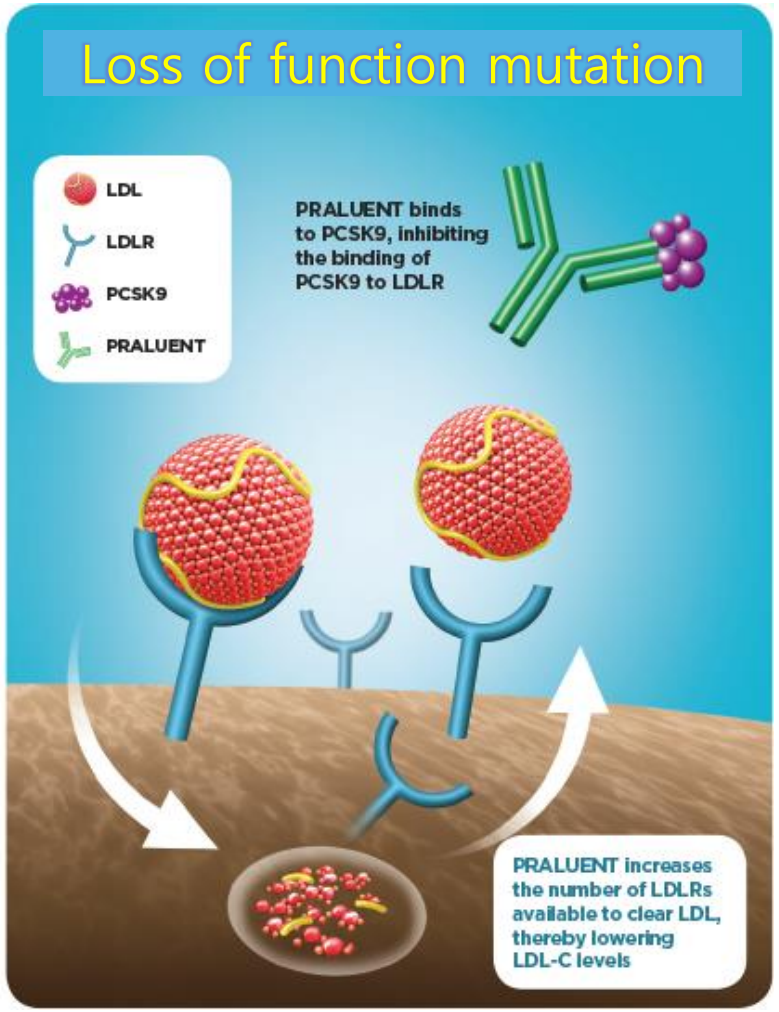
ORIGINAL ARTICLE

Variation in *PCSK9* and *HMGCR* and Risk of Cardiovascular Disease and Diabetes

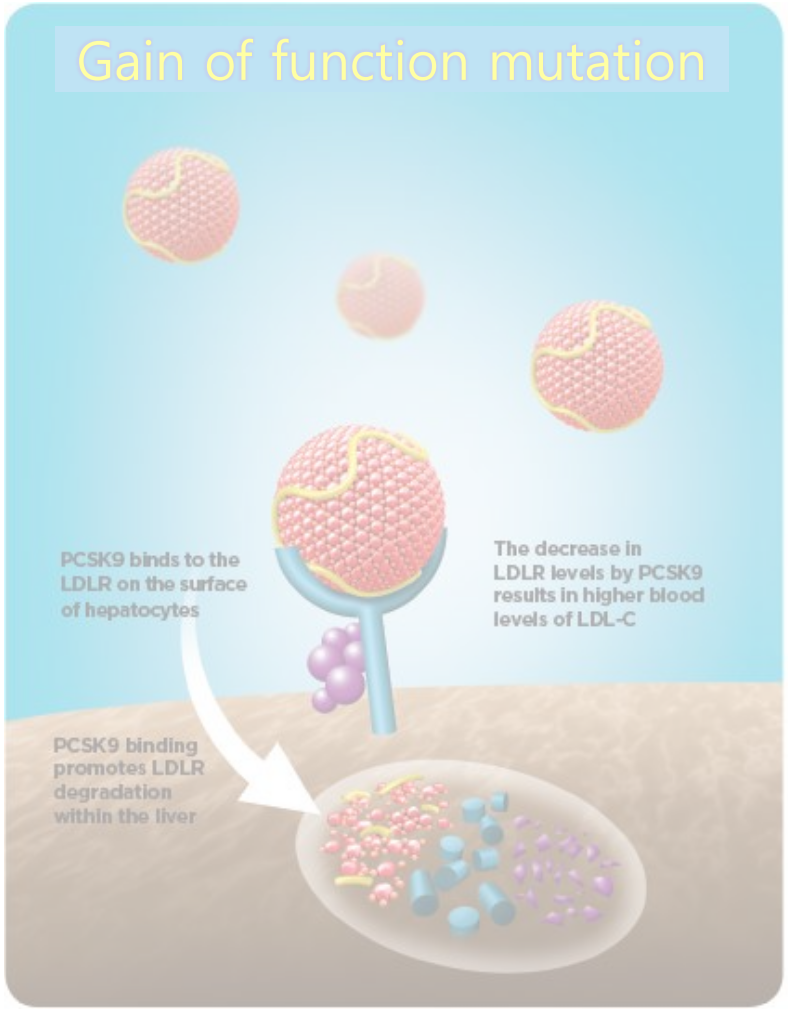
Brian A. Ference, M.D., Jennifer G. Robinson, M.D., M.P.H., Robert D. Brook, M.D., Alberico L. Catapano, Ph.D., M. John Chapman, Ph.D., David R. Neff, D.O., Szilard Voros, M.D., Robert P. Giugliano, M.D., George Davey Smith, M.D., D.Sc., Sergio Fazio, M.D., Ph.D., and Marc S. Sabatine, M.D., M.P.H.

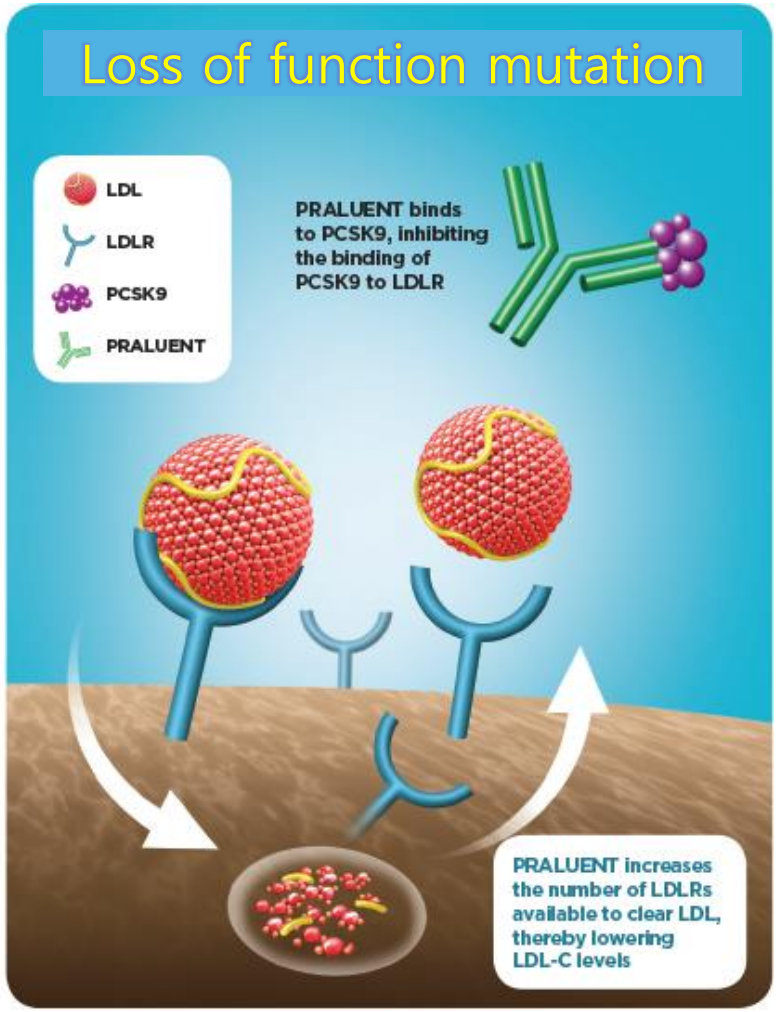
How PCSK9 and inhibitors work??





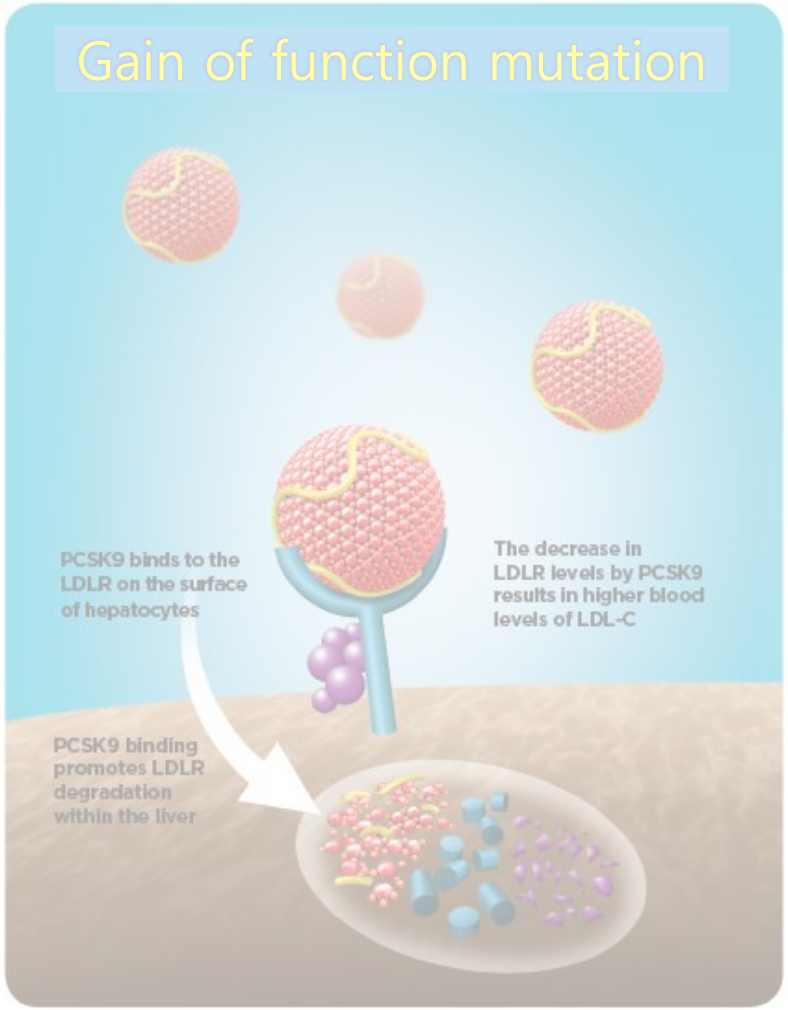
How PCSK9 and inhibitors work??

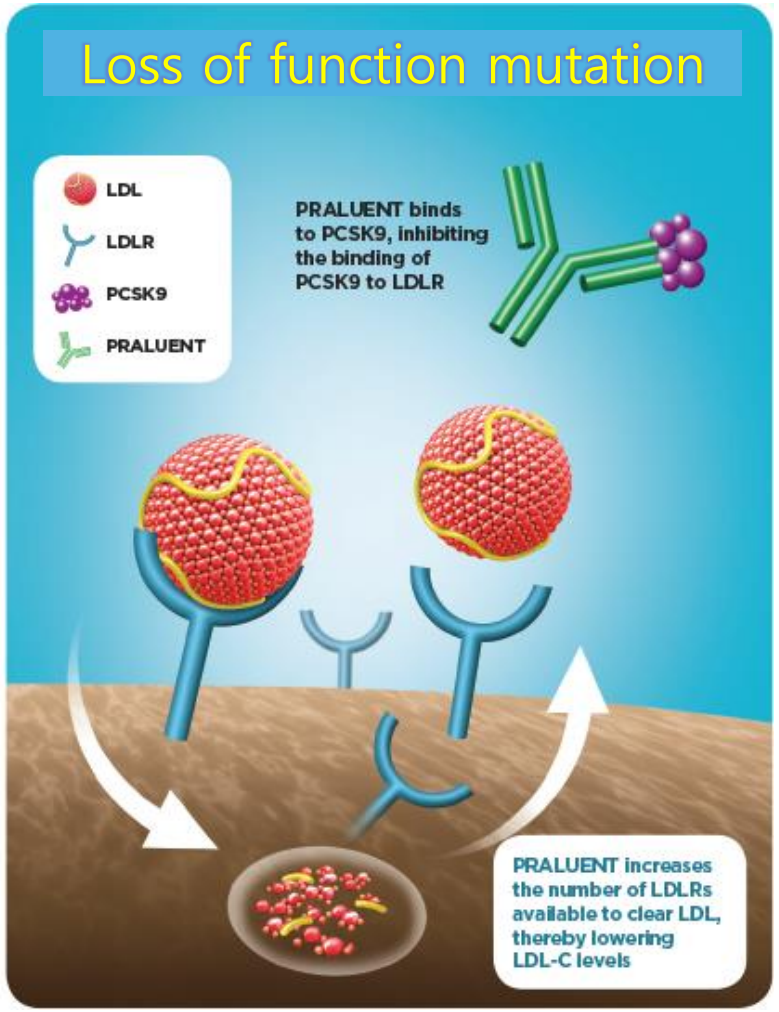




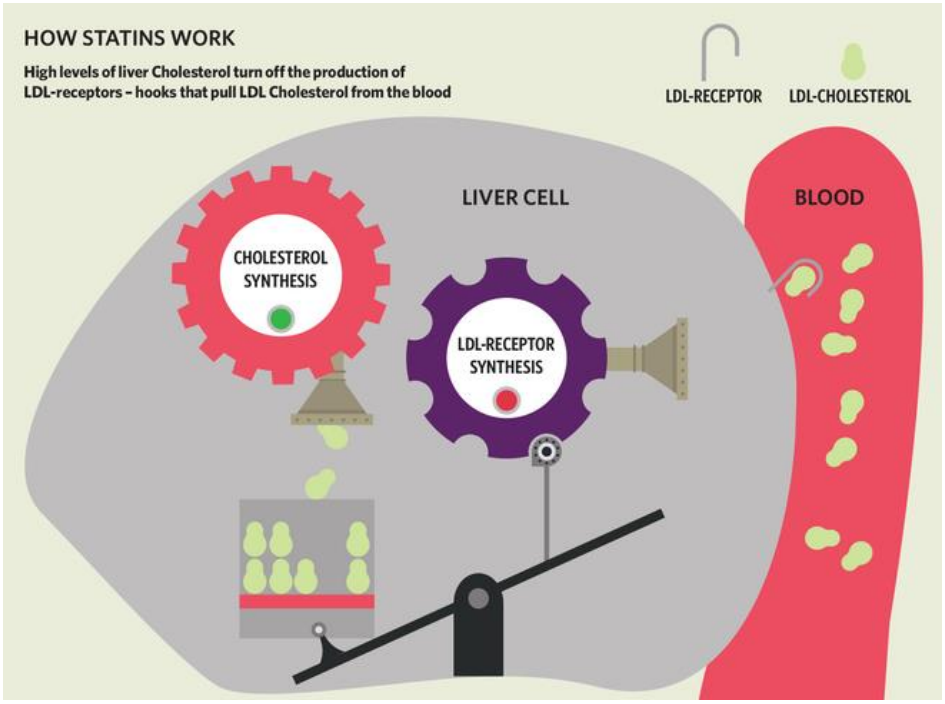
How PCSK9 and inhibitors work??

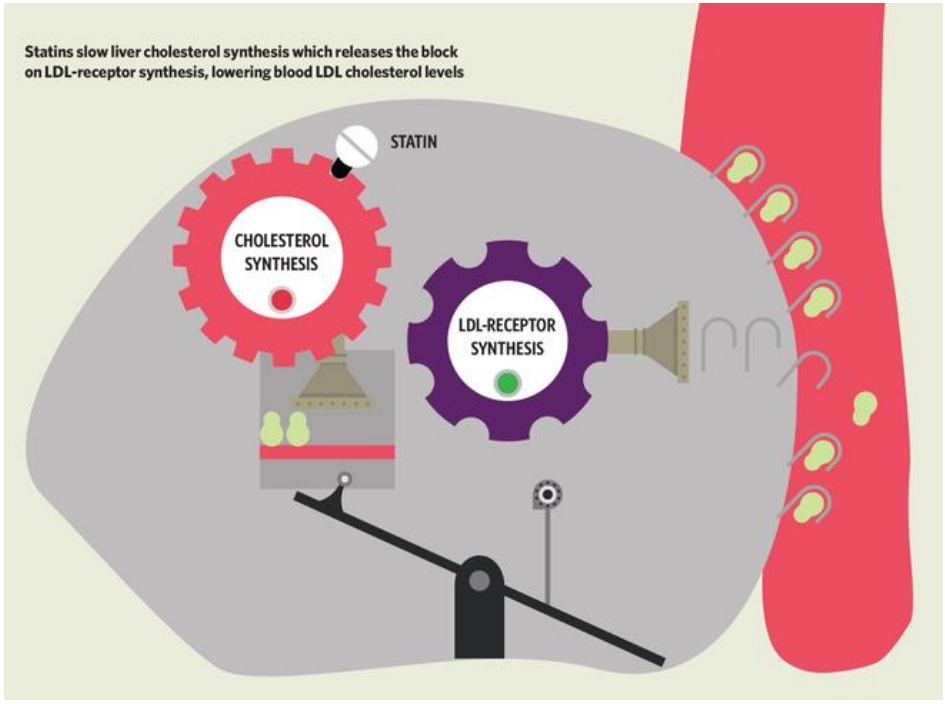
PCSK9 variant = PCSK9 inhibitor



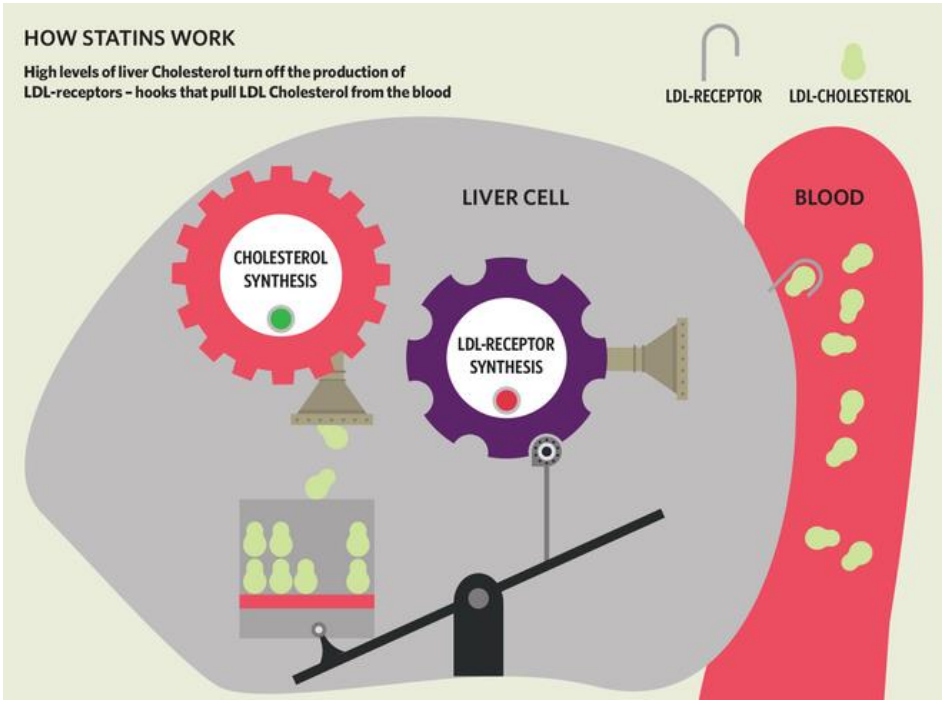


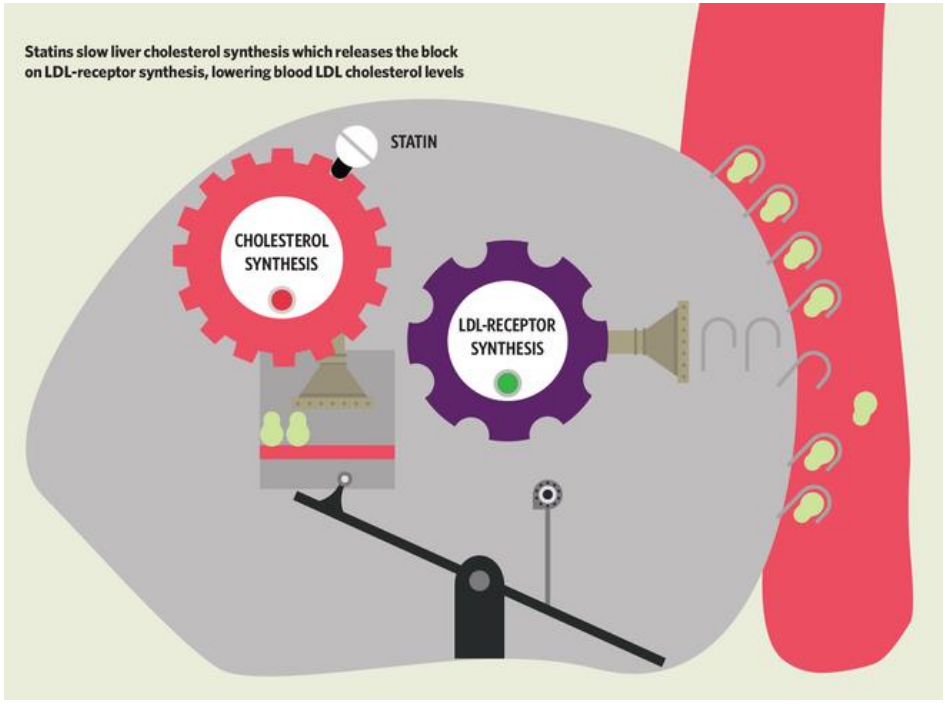
How HMGCR and statin work??





How HMGCR and statin work??

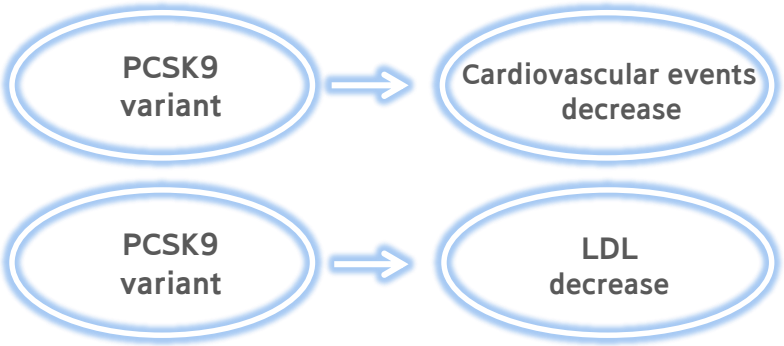




Statin = HMGCR variant

Already Known and unknown yet

- Inhibitors of PCSK9 are being evaluated in clinical trials for the treatment of cardiovascular disease.



- The effect of lowering LDL cholesterol levels by inhibiting PCSK9 on the risk of cardiovascular events or diabetes is unknown.

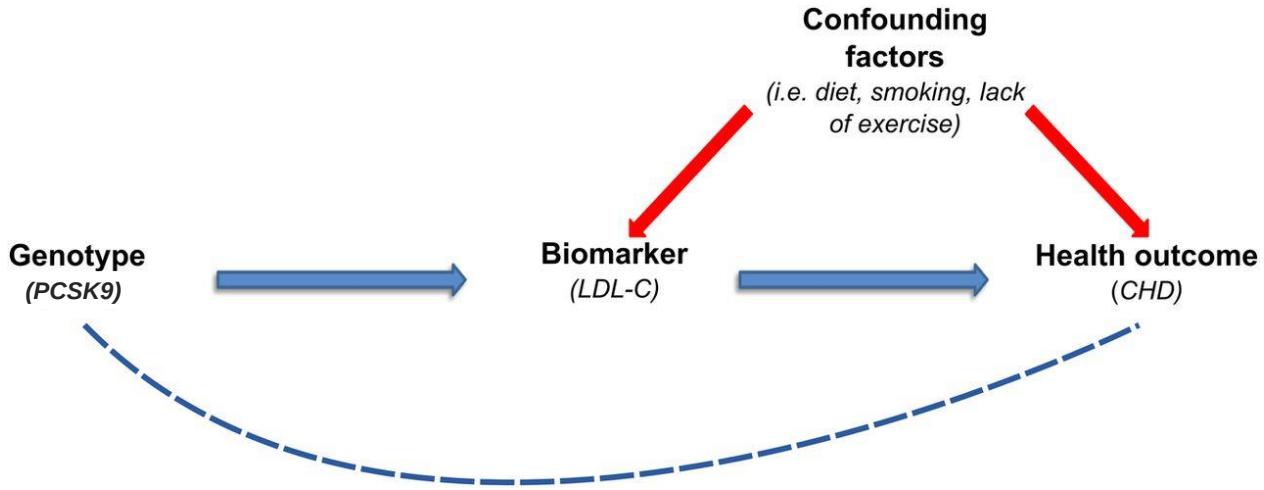
Mendelian Randomization

➢ Why is this method called MR?

1. As stated by **Mendel's second law** (the law of segregation), genetic variants segregate **randomly** and independently of environmental factors.

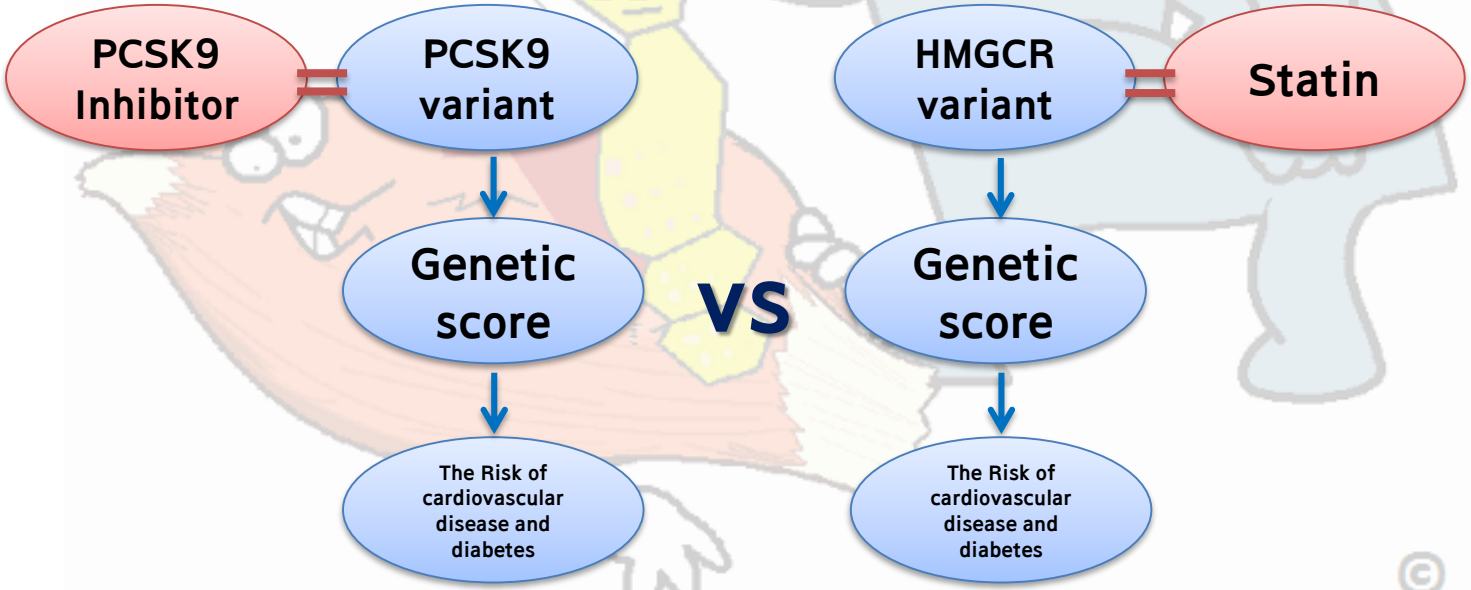
2. **Mendel's third law** (the law of independent assortment) suggests that genetic variants should also segregate independently of other traits

➢ Apply PCSK9,LDL and CHD to MR



The Goal of this study

- Use **LDL-C lowering variants in PCSK9** to estimate the effect of **inhibiting PCSK9** on both the risk of cardiovascular events and the risk of diabetes.
- Construct **genetic scores** that mimic the effect of **PCSK9 inhibitors** and the effect of **statins** (which target HMGCR)
- Compare the effect of these scores on the risk of cardiovascular disease and the risk of diabetes to make inferences about the potential clinical benefit and safety of treatment with a PCSK9 inhibitor **as compared with treatment with a statin.**



Study Population

Baseline Characteristic		Mean (SD or IQR)
Sample Size		112,772
No. Included Studies		14
Age (years)		59.9 (±6.5)
Women (%)		58.2%
LDL-C (mg/dl)	100-130	129.9 (±32.0)
HDL-C (mg/dl)	40-60	52.3 (±15.4)
triglycerides (mg/dl)*	< 200	117.0 (85-162)
total cholesterol (mg/dl)	< 240	207.8 (±36.8)
non-HDL-C (mg/dl)		155.3 (±37.6)
Systolic Blood Pressure (mmHg)		127.0 (±18.7)
Diastolic Blood pressure (mmHg)		75.2 (±10.2)
Weight (lbs)		169.2 (±33.1)
Body mass index (kg/m^2)		27.7 (±5.2)
Prevalent Diabetes (%)		5.7
Prevalent Cardiovascular disease (%)		1.9
Ever smoker (%)		54.3

Lipids

Table 1. 112,772 participants from 14 prospective cohort or case–control studies

Genetic instruments

➢ Selecting polymorphisms & Calculating GS

1. Identifying all polymorphisms within 100Kb of the target gene (PCSK9 or HMGCR).
2. Ranking each polymorphism by its p-value(< $5x10^{-8}$) for the association with LDL-C .
3. Low linkage disequilibrium with all other polymorphisms included in the score (r^2 < 0.2)
4. Confirming that each polymorphism added to a score had an effect on LDL-C using regression
5. To calculate the PCSK9 and HMGCR genetic score for each participant.
6. Sum these values to create a weighted genetic score for each participant.

- ❖ kb =kilo base pair
- ❖ 1kbp = 3.4nm

Genetic instruments

Table S4: PCSK9 polymorphisms included in genetic score and their association with LDL-C in the Global Lipids Genetics Consortium

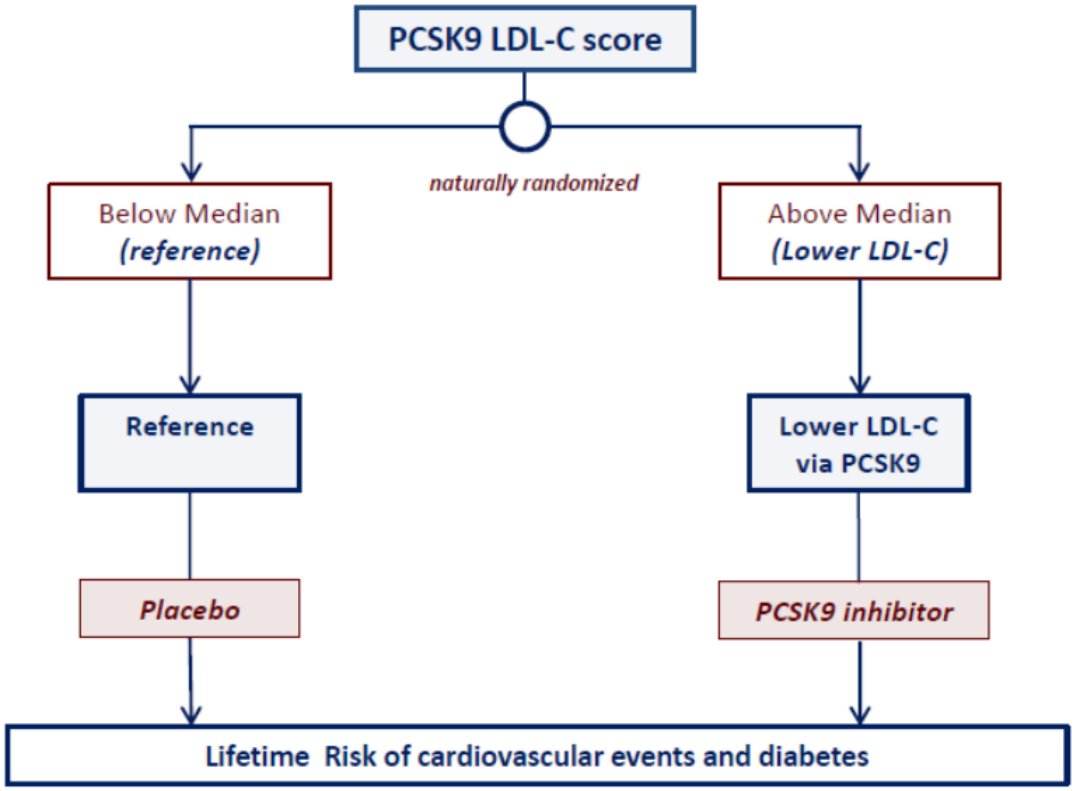
SNP	position	Sample size (N)	Frequency of LDL-C lowering allele	LDL-C effect size (mg/dl)	SE	P
rs11206510	chr1_55268627	172812	0.1544	2.6592	0.005	2.38E-53
rs2479409	chr1_55277238	172970	0.6675	2.0544	0.0041	2.52E-50
rs2149041	chr1_55274725	172903	0.8391	2.0352	0.0049	1.44E-35
rs2479394	chr1_55258652	172953	0.715	1.2352	0.0041	1.58E-19
rs10888897	chr1_55285649	165232	0.3945	1.6224	0.0042	8.43E-31
rs7552841	chr1_55291340	140234	0.6346	1.1776	0.0044	5.40E-15
rs562556	chr1_55296825	99192	0.1939	2.048	0.0066	6.16E-21

Table S6: HMGCR polymorphisms included in genetic score and their association with LDL-C in the Global Lipids Genetics Consortium

SNP	position	Sample size (N)	Frequency of LDL-C lowering allele	LDL-C effect size (mg/dl)	SE	P
rs12916	chr5:74656539	168357	0.5686	2.3456	0.1216	7.79E-78
rs17238484	chr5:74648496	80959	0.7467	2.0064	0.1984	1.35E-21
rs5909	chr5:74656175	89875	0.8984	1.9744	0.2816	4.93E-13
rs2303152	chr5:74641707	160116	0.8799	1.3536	0.2048	1.04E-09
rs10066707	chr5:74560579	89888	0.5831	1.5904	0.1728	2.97E-19
rs2006760	chr5:74562029	89885	0.8140	1.7056	0.2432	1.67E-13

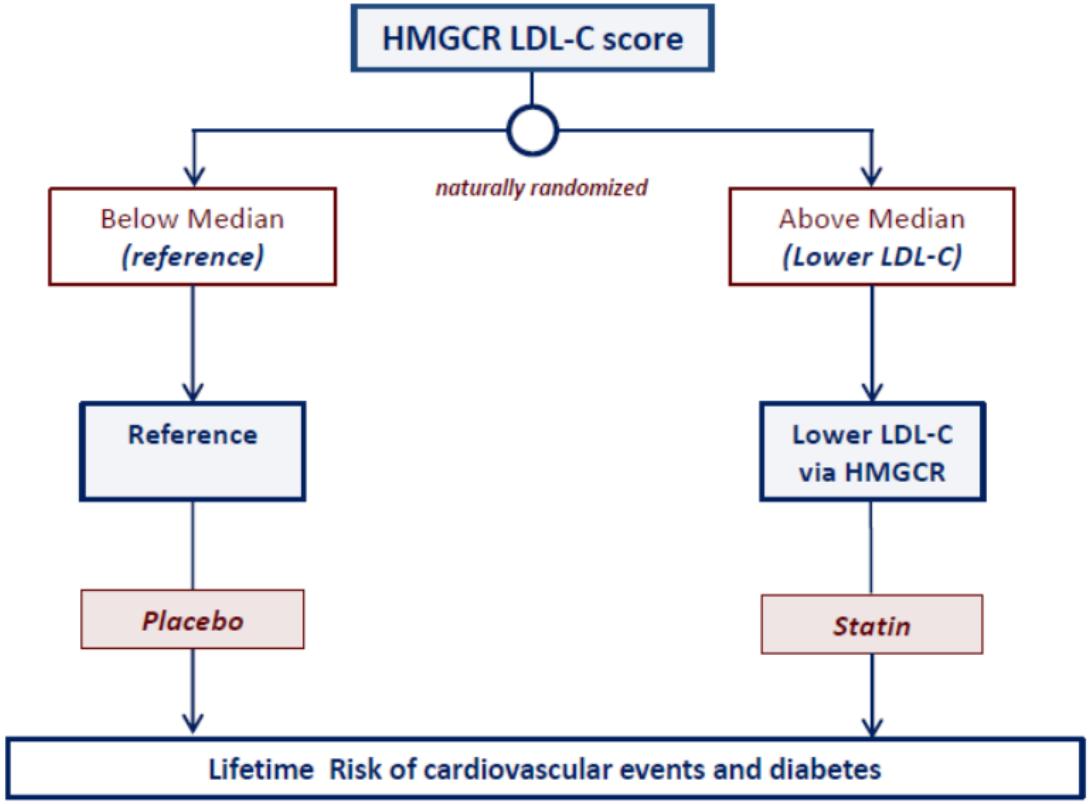
Study Design

PCSK9 genetic score



Study Design

HMGCR genetic score



Statistical analysis

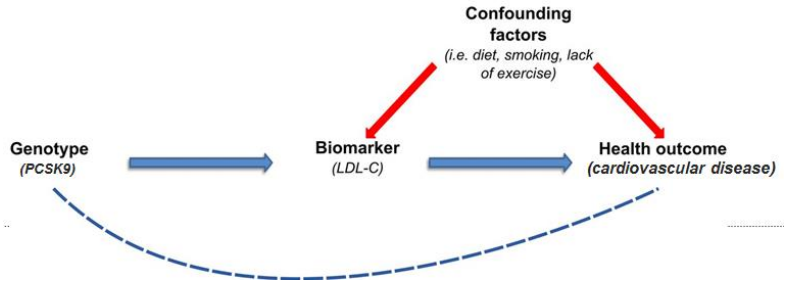
- ✓ **Linear regression**

  → The association between each weighted genetic LDL-C score and plasma LDL-C level

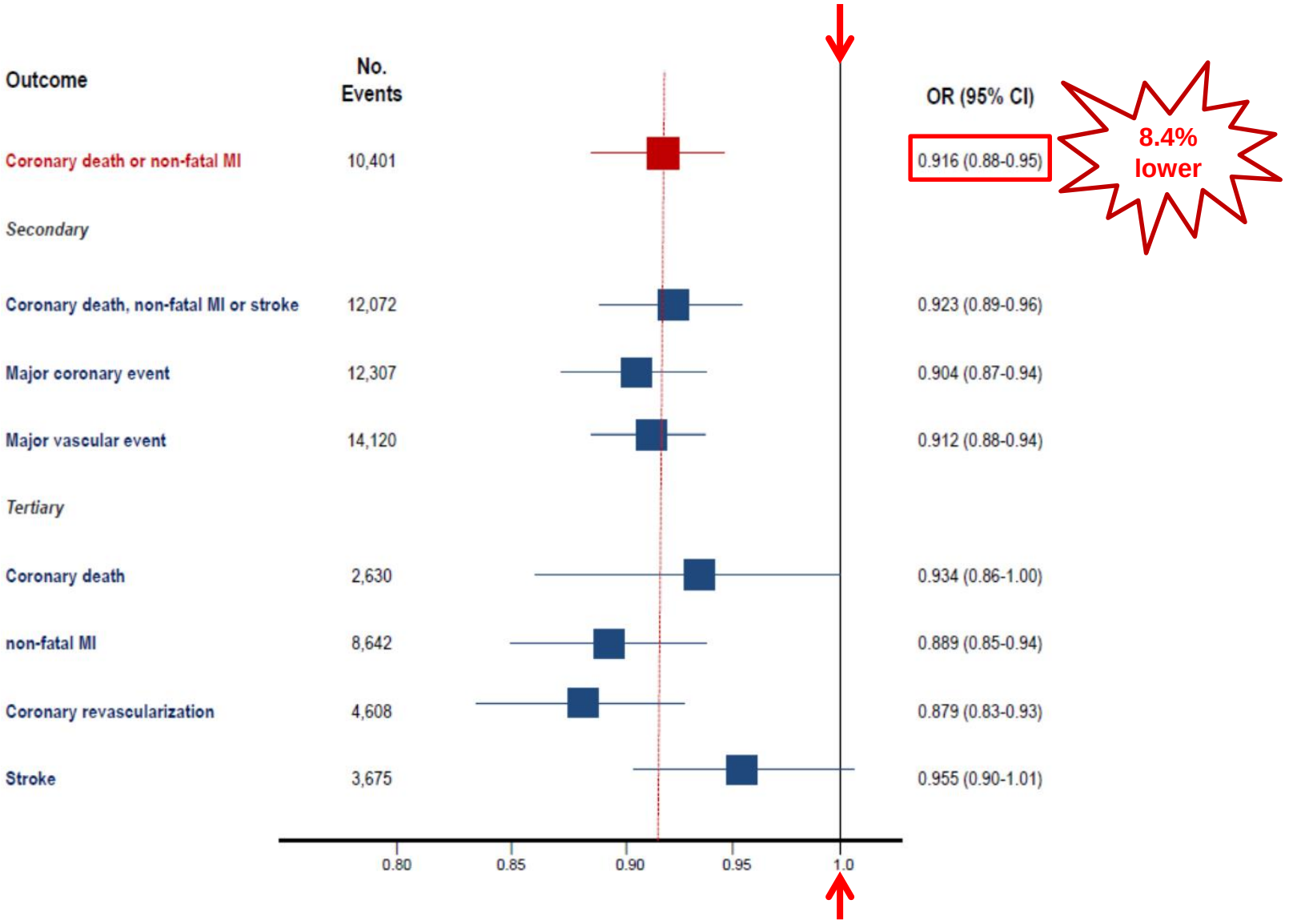
- ✓ **Logistic regression**

  → Compare the risk of cardiovascular events or diabetes with genetic score

- ✓ **All analyses were adjusted for age and gender.**
- ✓ **All analyses were conducted separately in each of the 14 studies**
- ✓ **Using a fixed-effects inverse variance-weighted meta-analysis to produce summary estimates of effect.**



With higher PCSK9 Genetic scores

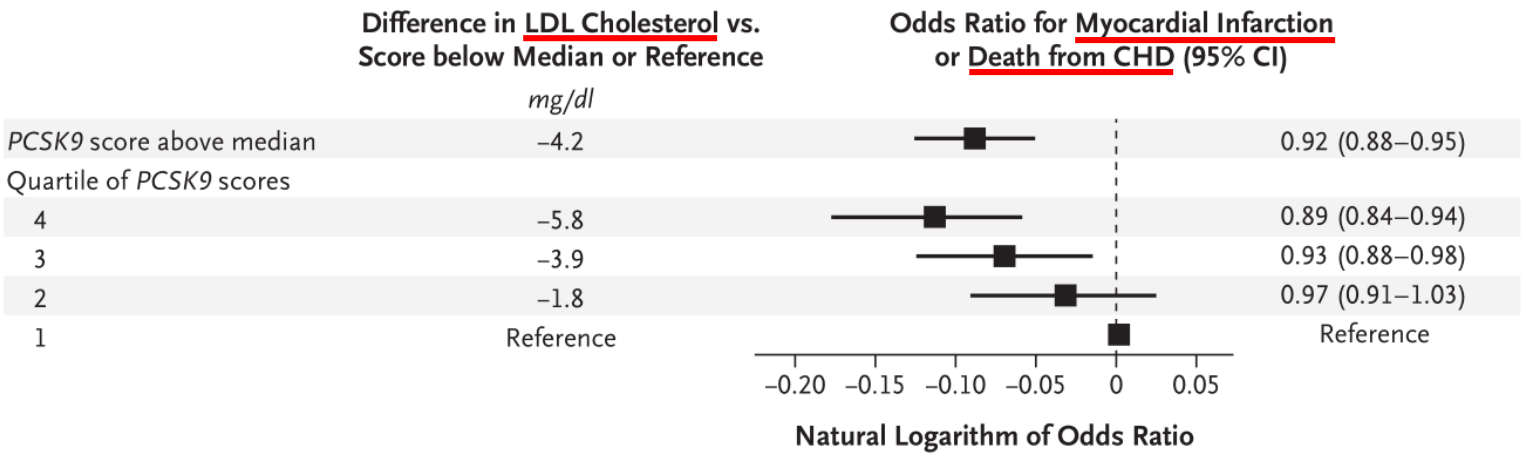


In dose-response analyses

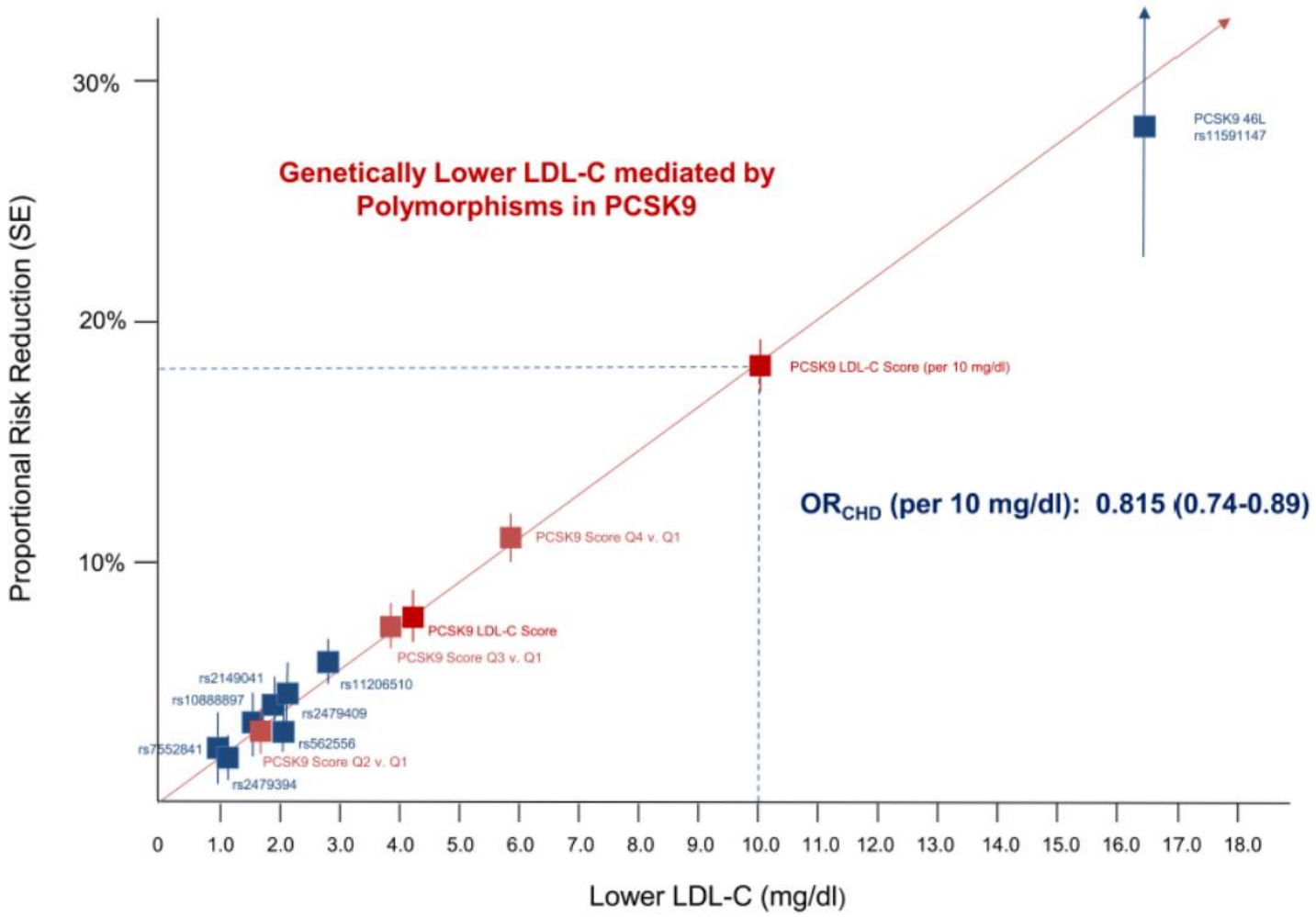
Effect of PCSK9 Genetic Scores on the Risk of MI or death from CHD

- ✓ MI = Myocardial Infarction
- ✓ CHD = Coronary Heart Disease

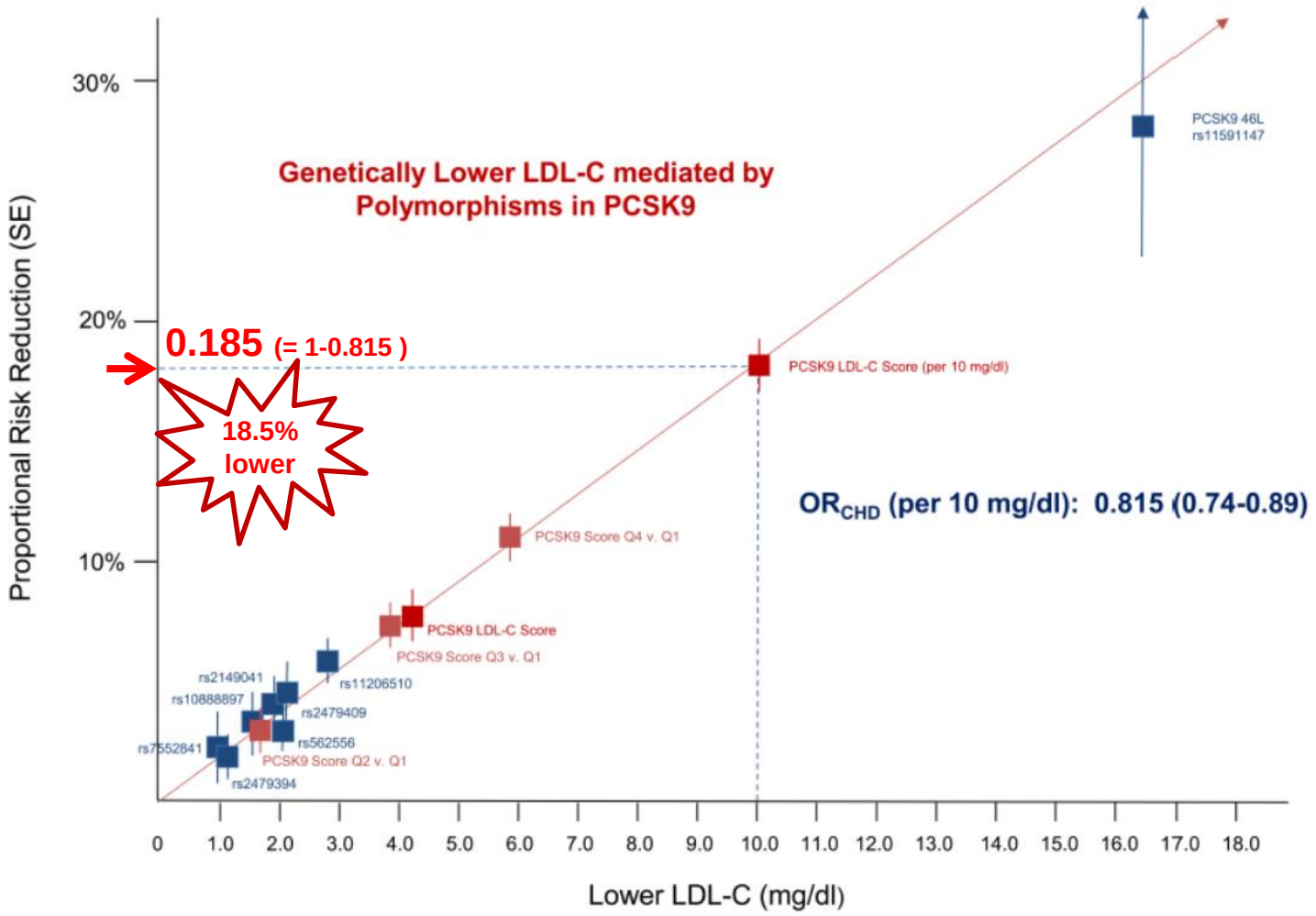
A ***PCSK9* Score**



Log-linear association between lower LDL-C mediated by PCSK9 polymorphisms and the risk of coronary death or MI

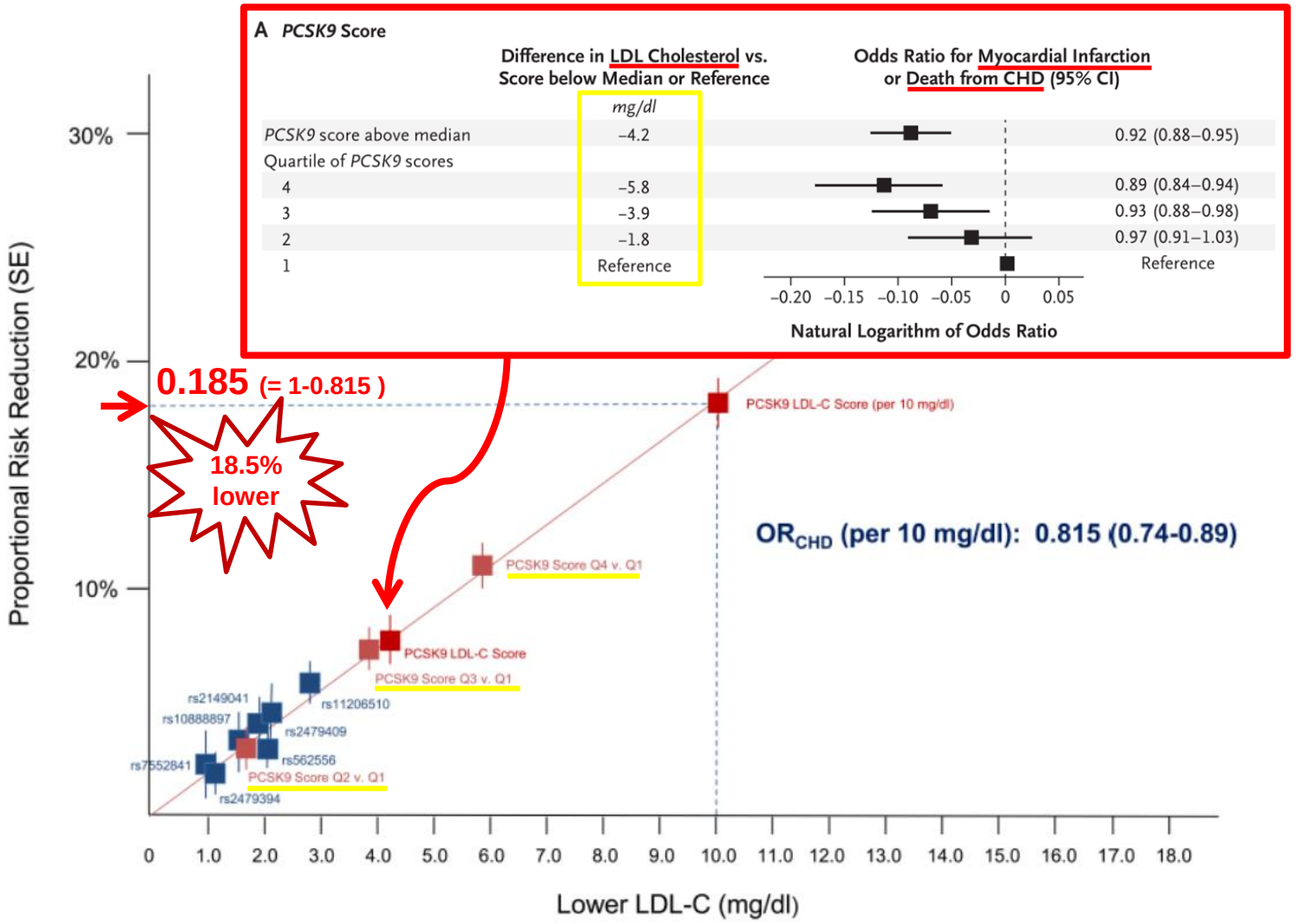


Log-linear association between lower LDL-C mediated by PCSK9 polymorphisms and the risk of coronary death or MI



4.Results

Log-linear association between lower LDL-C mediated by PCSK9 polymorphisms and the risk of coronary death or MI

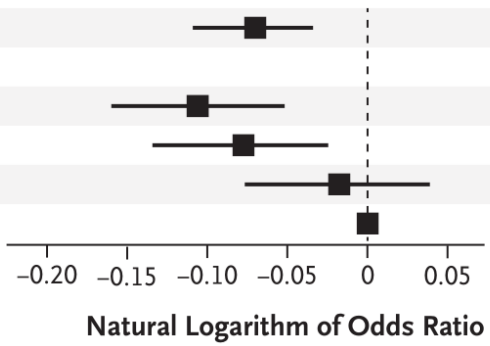


In similar analyses using the HMGCR

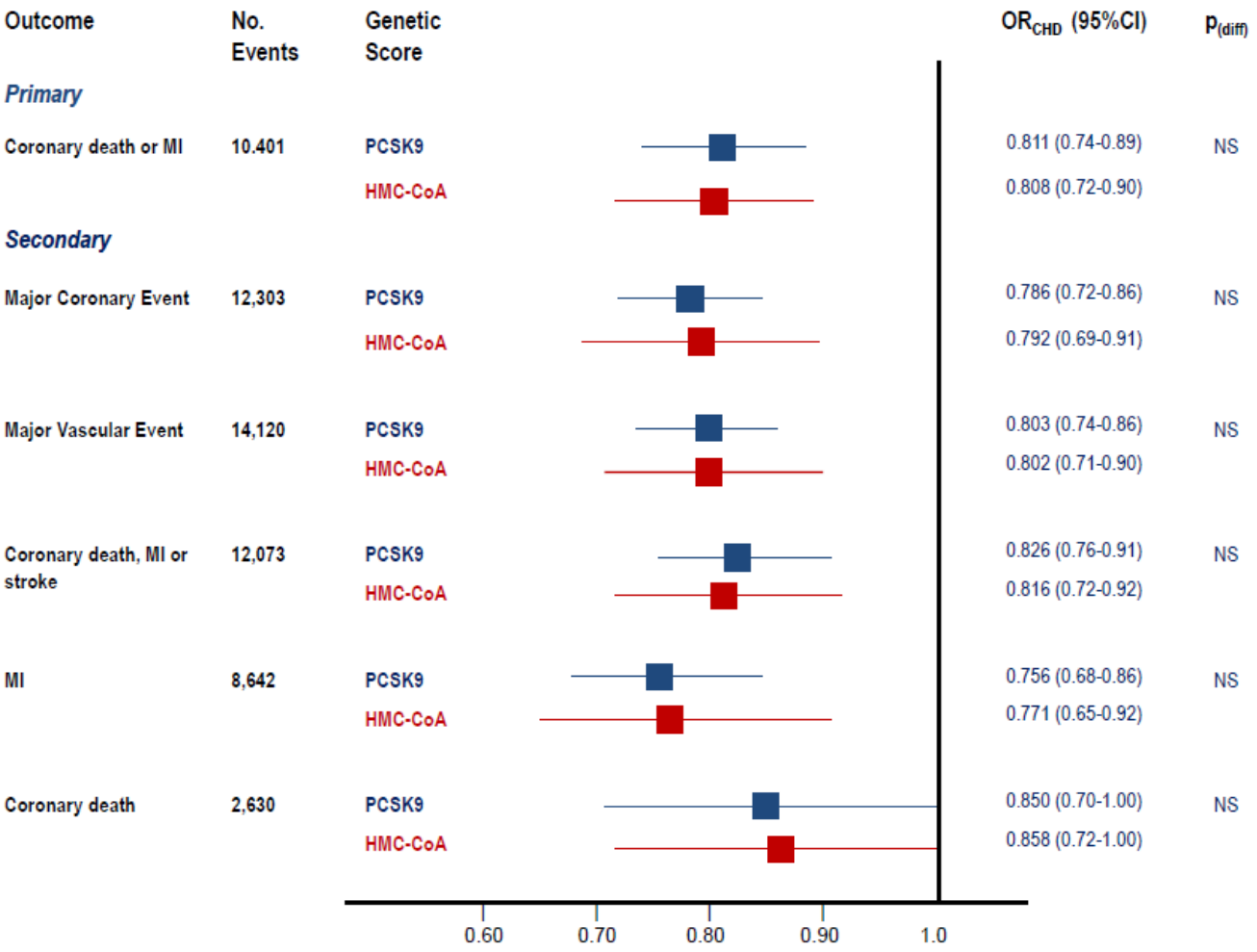
Effect of HMGCR Genetic Scores on the Risk of MI or death from CHD

B ***HMGCR* Score**

	Difference in LDL Cholesterol vs. Score below Median or Reference	Odds Ratio for Myocardial Infarction or Death from CHD (95% CI)
	mg/dl	
HMGCR score above median	−3.2	0.93 (0.90–0.97)
Quartile of *HMGCR* scores		
4	−4.6	0.90 (0.85–0.95)
3	−3.1	0.93 (0.88–0.98)
2	−1.2	0.98 (0.92–1.04)
1	Reference	Reference



The effects of the PCSK9 and HMGCR scores were very similar for all of the cardiovascular outcomes



Risk of Diabetes

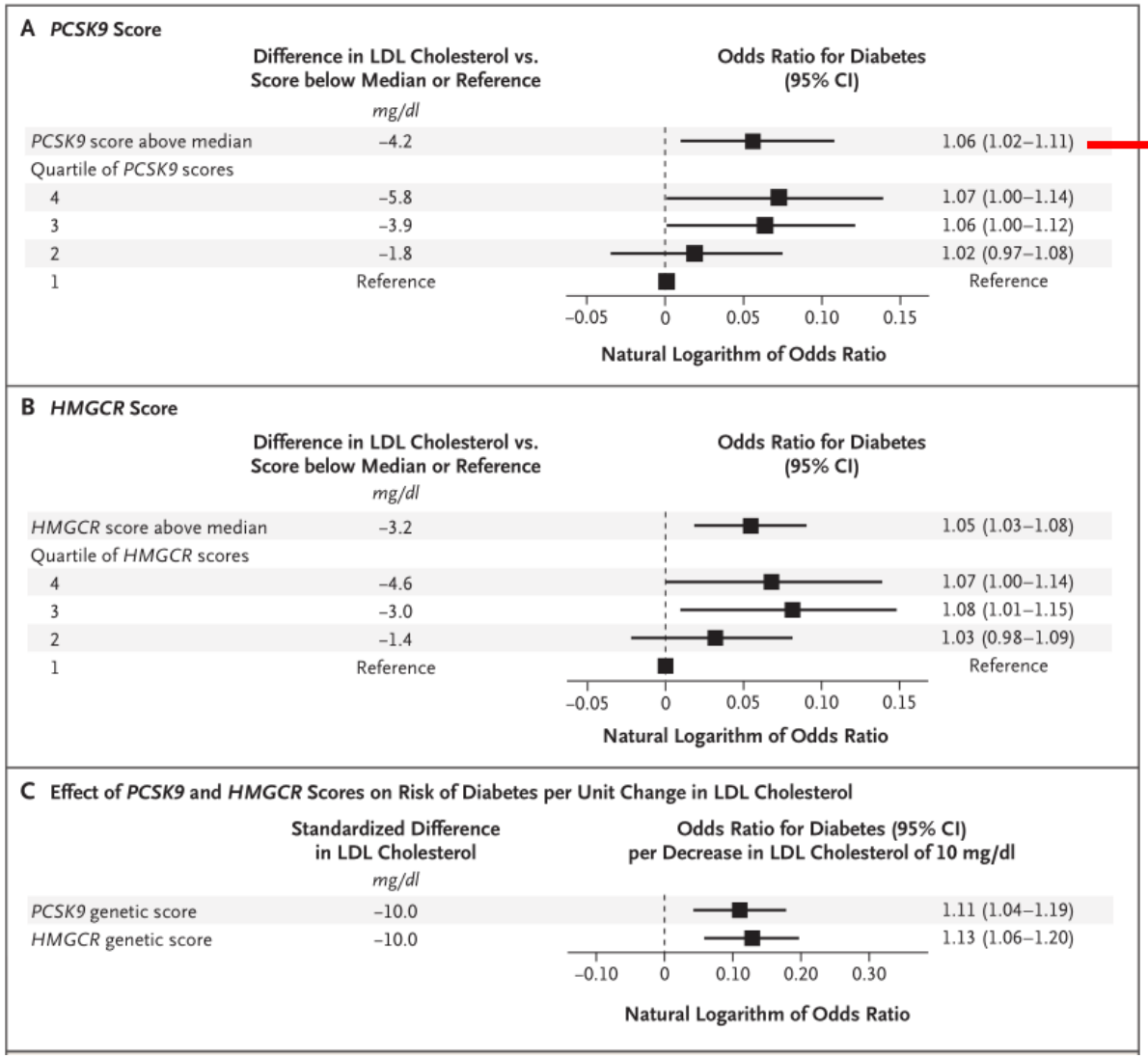
PCSK9 & HMGCR

The Risk of cardiovascular disease

PCSK9 & HMGCR

The Risk of diabetes

Risk of Diabetes (PCSK9 & HMGCR)



In the 2×2 factorial analysis

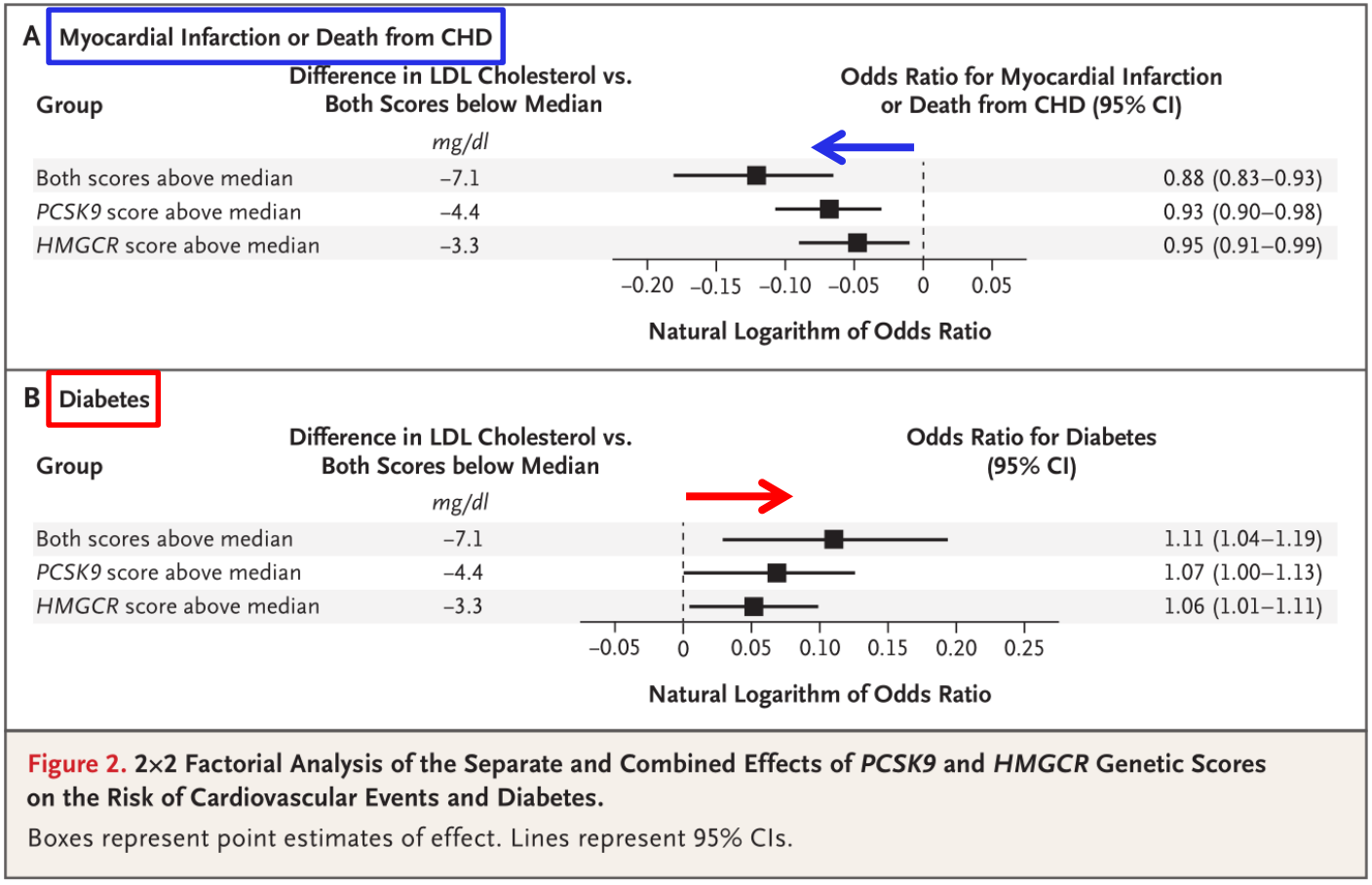
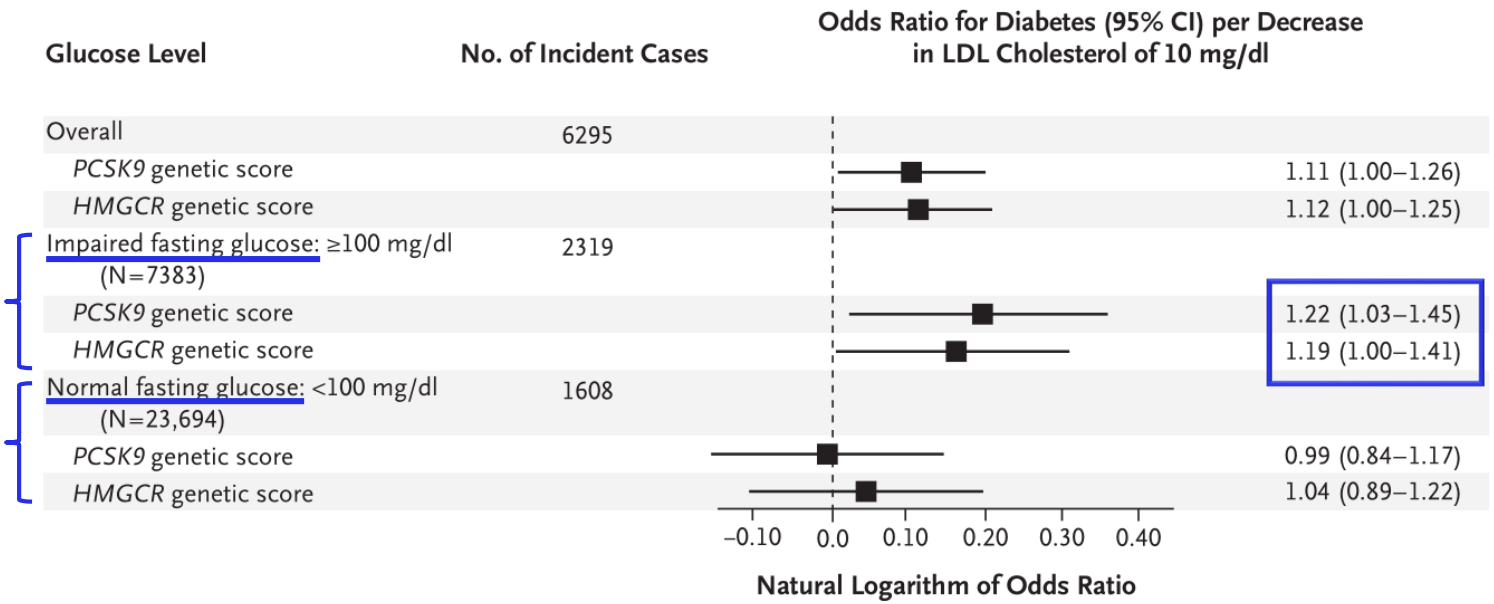


Figure 2. 2×2 Factorial Analysis of the Separate and Combined Effects of *PCSK9* and *HMGCR* Genetic Scores on the Risk of Cardiovascular Events and Diabetes.

Boxes represent point estimates of effect. Lines represent 95% CIs.

"The PCSK9 and HMGCR genetic scores had **additive effects** on LDL- C, the risk of cardiovascular events and diabetes!"

Effect of PCSK9 and HMGCR Scores on the Risk of Incident Diabetes



◆ What is Impaired Fasting Glucose (IFG) ?

- ✓ **A type** of prediabetes (100mg/dl – 126mg/dl)
- ✓ The blood sugar level during fasting is consistently higher than normal levels (< 100mg/dl)
- ✓ However, the level is not high enough to be diagnosed as diabetes mellitus (126mg/dl >)

Summary

- PCSK9 and HMGCR variants that mimic the effect of PCSK9 inhibitors and statins had independent and additive effects on the risk of both cardiovascular events and diabetes.

- PCSK9 and HMGCR variants were associated with approximately the same effect on the risk of cardiovascular disease (per unit decrease in the LDL cholesterol level)

- Treatment with a PCSK9 inhibitor should reduce the risk of cardiovascular events by approximately the same amount as treatment with a statin.

New Onset

- Like statins, **PCSK9 inhibitors may also increase the risk of new-onset diabetes.**

- Any potential increased risk of **new onset diabetes** during treatment with a PCSK9 inhibitor is likely to be **confined to persons with impaired fasting glucose levels(IFG).**

- The proportional **reduction in cardiovascular risk was much greater than the increased risk of diabetes.**

Any potential increased risk of diabetes. <<< The reduction in cardio-vascular risk with PCSK9 inhibitors

Practice

What is Two sample MR ?

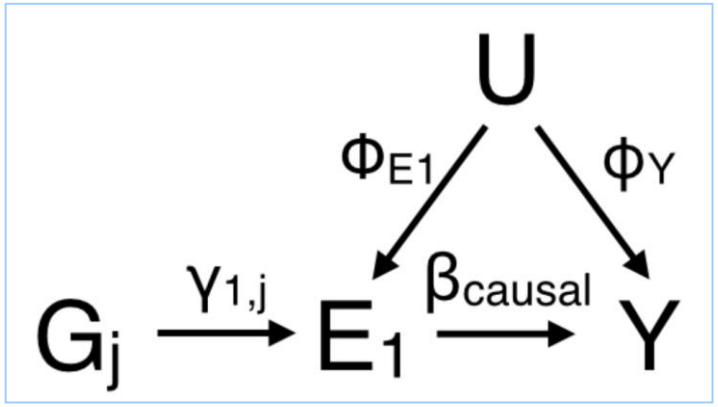
- Two sample Mendelian randomization (2SMR) is a method to estimate the causal effect of an exposure on an outcome using only summary statistics from GWAS
- Two-sample MR exploits the fact that it is not necessary to obtain the effect of the instrumental variable-risk factor association and instrumental variable-outcome association from the same sample of participants.

- The workflow for performing MR is as follows:

(1) Select instruments for the exposure (perform LD clumping if necessary)
(2) Extract the instruments from the MR Base GWAS database for the outcomes of interest
(3) Harmonise the effect sizes for the instruments on the exposures and the outcomes to be each for the same reference allele
(4) Perform MR analysis, sensitivity analyses, create plots, compile reports

What is Two sample MR ?

∗The causal effect of E1 on Y can be estimated by fitting the following regression models



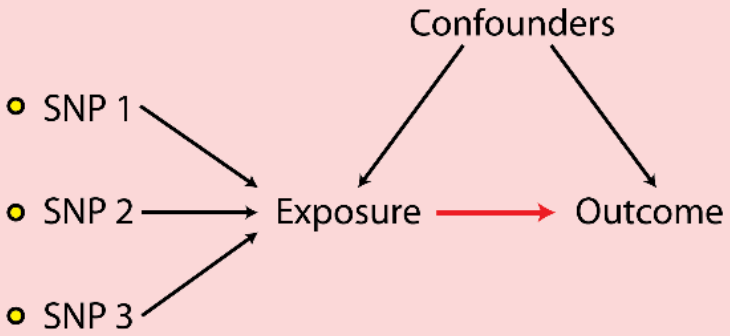
$$Y_i = \Gamma_0 + \Gamma_{1,j} G_{ij} + \varepsilon_{Y,i} \quad -(1)$$

$$E_{1,i} = \gamma_o + \gamma_{1,j} G_{ij} + \varepsilon_{E1,i} \quad -(2)$$

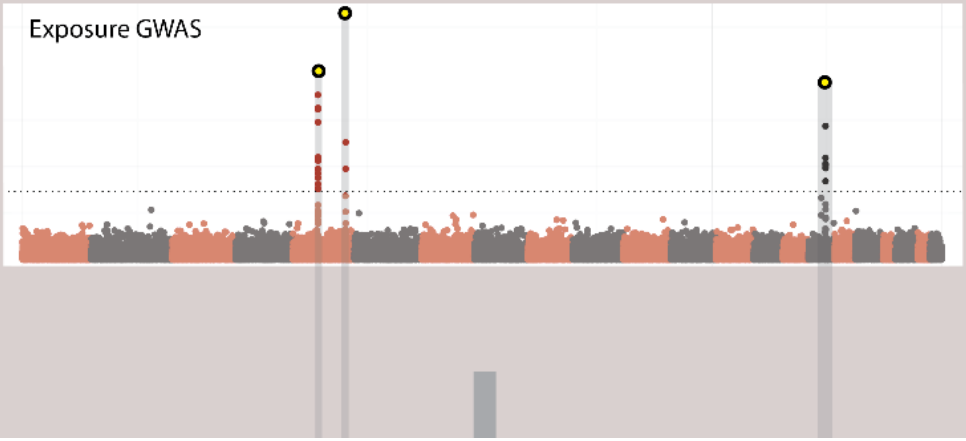
$$\frac{\Gamma_{1,j}}{\gamma_{1,j}} = \frac{\gamma_{1,j}\beta}{\gamma_{1,j}} = \beta, \quad \hat{\beta} = \frac{\hat{\Gamma}_{1,j}}{\hat{\gamma}_{1,j}}$$

MRbase (https://mrcieu.github.io/TwoSampleMR/)

Objective: Infer the causal effect of the exposure on the outcome



1.

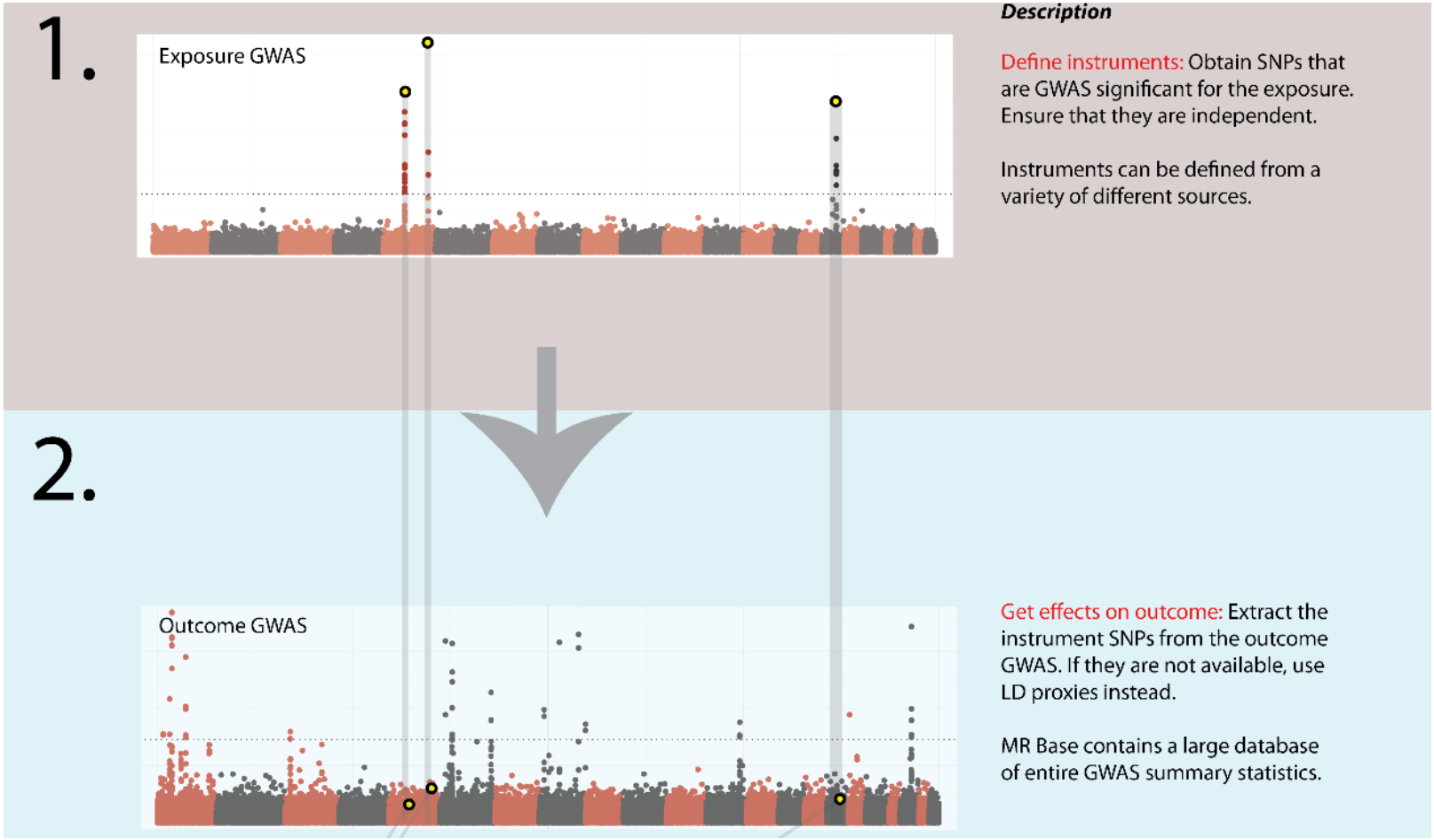


Description

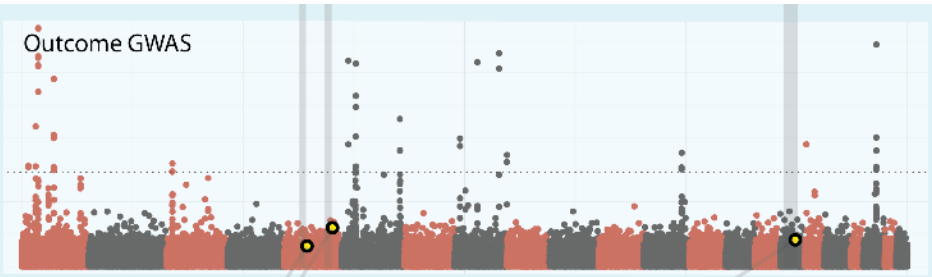
Define instruments: Obtain SNPs that are GWAS significant for the exposure. Ensure that they are independent.

Instruments can be defined from a variety of different sources.

MRbase (https://mrcieu.github.io/TwoSampleMR/)



MRbase (https://mrcieu.github.io/TwoSampleMR/)



Get effects on outcome: Extract the instrument SNPs from the outcome GWAS. If they are not available, use LD proxies instead.

MR Base contains a large database of entire GWAS summary statistics.

3.

	Exposure GWAS				Outcome GWAS			
SNP	*Effect*	*Effect allele*	*Other allele*	*Effect allele frequency*	*Effect*	*Effect allele*	*Other allele*	*Effect allele frequency*
rs123456	0.132	A	G	0.28	0.022	A	G	0.26
rs234567	-0.485	G	T	0.41	**0.056**	**T**	**G**	**0.61**
rs345678	0.203	G	C	0.11	**-0.046**	**G**	**C**	**0.88**

Harmonize effects

	Exposure GWAS				Outcome GWAS			
SNP	*Effect*	*Effect allele*	*Other allele*	*Effect allele frequency*	*Effect*	*Effect allele*	*Other allele*	*Effect allele frequency*
rs123456	0.132	A	G	0.28	0.022	A	G	0.26
rs234567	-0.485	G	T	0.41	**-0.056**	**G**	**T**	**0.39**
rs345678	0.203	G	C	0.11	**0.046**	**G**	**C**	**0.12**

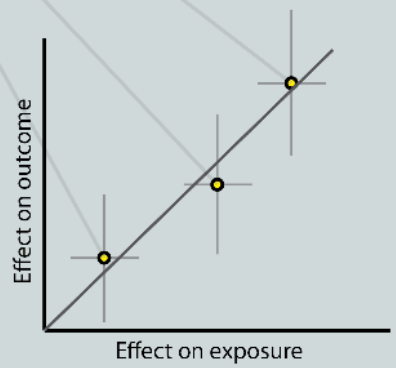
Harmonize effects: Ensure that the effect of the SNP on the exposure and the effect of the SNP on the outcome correspond to the same allele.

MRbase (https://mrcieu.github.io/TwoSampleMR/)

Harmonize effects

	Exposure GWAS				Outcome GWAS			
SNP	*Effect*	*Effect allele*	*Other allele*	*Effect allele frequency*	*Effect*	*Effect allele*	*Other allele*	*Effect allele frequency*
rs123456	0.132	A	G	0.28	0.022	A	G	0.26
rs234567	-0.485	G	T	0.41	**-0.056**	**G**	**T**	**0.39**
rs345678	0.203	G	C	0.11	**0.046**	**G**	**C**	**0.12**

4.



Perform analysis: Using the harmonized data, perform Mendelian randomization analyses and related sensitivity analyses.

The slope of the regression line corresponds to the causal effect of the exposure on the outcome

Practice