Effectiveness and safety of using statin therapy for the primary prevention of cardiovascular diseases in older patients with chronic kidney disease who are hypercholesterolemic: a target trial emulation study

Dr. Edidiong Etete

Background

This paper by Xu et al.,2025 investigates the effectiveness and safety of statins for primary prevention of cardiovascular disease (CVD) in older adults with chronic kidney disease (CKD) and hypercholesterolemia,

Focusing particularly on patients aged 75–84 years and ≥85 years.

Using real-world electronic health records from Hong Kong and a target trial emulation design, the study includes over 700,000 person-trials spanning from 2008 to 2015.

DEFINITION

- Target trial emulation is a methodological framework used in observational studies to mimic the design and analysis of randomized controlled trials (RCTs).
- It aims to reduce biases inherent in observational data, such as immortal time bias and prevalent user bias, by aligning the study design with the principles of RCTs.

Hernán MA, Robins JM. Using big data to emulate a target trial when a randomized trial is not available. American journal of epidemiology. 2016;183(8):758-64.

Target trial

- In a study that uses the target trial emulation approach,
- Investigators specify the hypothetical RCT that would ideally be conducted to address a given study question and then specify the design elements of an observational study aligned with the elements of that “target” RCT[1].

Hubbard RA, Gatsonis CA, Hogan JW, Hunter DJ, Normand S, Troxel AB. Target trial emulation” for observational studies—potential and pitfalls. N Engl J Med. 2024;391(21):1975-7.

Ways to emulate target trials

Necessary design elements include

- Eligibility criteria,
- Participant selection,
- Treatment strategies,
- Treatment assignment,
- Start and end of follow-up,
- Outcome measure,
- Efficacy assessment, and
- Statistical analysis plan (SAP).

1. Hernán MA, Robins JM. Using big data to emulate a target trial when a randomized trial is not available. American journal of epidemiology. 2016;183(8):758-64. 2.Hernán MA. Methods of public health research—strengthening causal inference from observational data. New England Journal of Medicine. 2021;385(15):1345-8.

Target trial

- In this study, target trial emulation was employed using electronic health records from the Hong Kong Hospital Authority
- To investigate the long-term effectiveness and safety of statin therapy for the primary prevention of cardiovascular diseases in older patients with chronic kidney disease who are hypercholesterolemic.

Emulation details

- The emulation involved creating a sequence of 96 nested trials, where eligibility criteria were applied monthly from January 2008 to December 2015.
- Patients were categorized into age groups (60–74 years, 75–84 years, and ≥85 years), and the 60–74 years group served as a benchmark for validating the emulated trial.

Key features of Emulation

- This included:
- Aligning "time zero" with the point of treatment eligibility and assignment.
- Including only new statin users in the treatment group while excluding previous users.
- Using pooled logistic models to estimate hazard ratios for outcomes.
- Adjusting for time-varying confounders and competing risks in the per-protocol analysis.
- This approach allowed the study to assess the causal effects of statin therapy in real-world settings while addressing biases and confounding factors.

Defining target trial-hypothetical RCT features

- Population: Older adults (≥65) with CKD and LDL ≥2.6 mmol/L, no prior CVD
- Intervention: Start statins
- Comparator: No statins
- Follow-up: From baseline
- Outcomes: CVD events, death, safety

Eligibility criteria-target trial

- Have history of CKD and LDL ≥ 2.6
- mmol/L;
- Age ≥ 60 years old (categorized into three
- age groups to conduct separate analysis:
- 60-74, 75-74, ≥85 years old);
- Not use statin before baseline;
- Not use fibrate and other lipid regulating
- drugs on or before baseline.
- No history of CVD, liver dysfunction or
- myopathies on or before baseline
- No missing value of covariates at baseline

Supplementary Table 1. Specification and emulation of target trial

Protocol component	Target trial specification	Target trial emulation
Eligible criteria	Eligible for at least one pre-defined statin indication; Have history of CKD and LDL ≥ 2.6 mmol/L; Age ≥ 60 years old (categorized into three age groups to conduct separate analysis: 60-74, 75-74, ≥85 years old); Not use statin before baseline; Not use fibrate and other lipid regulating drugs on or before baseline. No history of CVD, liver dysfunction or myopathies on or before baseline No missing value of covariates at baseline	Same as for the target trials;
Treatment strategy	Initiate statin use vs. Not initiate statin use. Physician will decide whether to stop or start statin therapy when the contradiction (myopathies and liver dysfunction) occurred in the treatment group, or the indication occurred in the control group during the follow-up period.	Same as for the target trials 3-month gap was given for the ascertainment of statin discontinuation.
Treatment assignment	Subjects are randomly assigned to a treatment strategy at baseline and will be aware of the treatment strategy they are assigned to.	Participants were classified into different groups according to the prescription records at baseline.
Outcomes	Incidence of cardiovascular diseases (CVDs), all-cause mortality and major adverse events Three subtypes of CVDs • Myocardial infarction • Heart failure • Stroke Major adverse events • Myopathies • Liver dysfunction	Same as for the target trial
Follow-up	From baseline until the occurrence of death, the outcome of interest, or the administrative end of follow-up (31 December 2018)	Same as for the target trial
Statistical analysis	Intention-to-treat (ITT) analysis; Per-protocol analysis: censored the patients when they deviated treatment strategy. Subgroup analysis by sex, Charlson Comorbidity Index (CCI≤8/>8), and CKD stages at baseline	Same as for the target trials, where the ITT analysis and per-protocol analysis with conducted via sequence trial emulation. The statistical model was adjusted by the baseline covariates and the inverse probability weighting was applied to account for the selection bias introduced by the artificial censoring in the per-protocol analysis.

Methods/statistical analysis

- The study employed several statistical methods to analyze the data:
- **Pooled Logistic Regression**:
- They used pooled logistic regression as supposed to cox regression.

Method of pooled OR from logistic regression instead of using HR from survival analysis.

- In a paper that looked at the comparison time dependent Cox regression, pooled logistic regression and cross sectional pooling
- The compared the estimate across different methods
- The compared the slopes actually estimates, which is log OR from the pooled logistics and Log HR from the cox regression.
- The found that the are quite close, so they could use pooled logistics instead of the cox regression

Ngwa JS, Cabral HJ, Cheng DM, Pencina MJ, Gagnon DR, LaValley MP, et al. A comparison of time dependent Cox regression, pooled logistic regression and cross-sectional pooling with simulations and an application to the Framingham Heart Study. BMC medical research methodology. 2016;16:1-12.

Table 4 Comparison of longitudinal effect on survival ($N = 1000$)

Scenarios		CSP_UNADJUSTED					CSP_ADJUSTED & TDCM				
Event rate	γ	Estimate	SE	CP	Bias	MSE	Estimate	SE	CP	Bias	MSE
90 %	0.000	0.003	0.054	0.957	0.003	0.006	0.003	0.055	0.954	0.003	0.006
	0.500	0.498	0.055	0.954	−0.002	0.006	0.498	0.056	0.952	−0.002	0.006
	1.000	1.083	0.058	0.720	0.082	0.014	1.002	0.059	0.958	0.002	0.007
50 %	0.000	−0.001	0.075	0.953	−0.001	0.011	−0.002	0.076	0.944	−0.002	0.012
	0.500	0.499	0.070	0.947	−0.001	0.010	0.499	0.071	0.944	−0.001	0.010
	1.000	1.001	0.073	0.946	0.001	0.011	1.002	0.074	0.948	0.002	0.011
10 %	0.000	0.007	0.168	0.944	0.007	0.058	0.007	0.171	0.938	0.007	0.060
	0.500	0.501	0.158	0.947	0.001	0.051	0.501	0.161	0.946	0.001	0.053
	1.000	1.067	0.145	0.906	0.067	0.048	1.003	0.147	0.937	0.003	0.045
		PLR_UNADJUSTED					PLR_ADJUSTED				
Event rate	γ	Estimate	SE	CP	Bias	MSE	Estimate	SE	CP	Bias	MSE
90 %	0.000	0.003	0.066	0.957	0.003	0.008	0.003	0.067	0.957	0.003	0.009
	0.500	0.599	0.069	0.709	0.099	0.019	0.601	0.070	0.711	0.101	0.020
	1.000	1.342	0.080	0.005	0.342	0.130	1.255	0.082	0.111	0.255	0.078
50 %	0.000	−0.001	0.080	0.950	−0.001	0.013	−0.002	0.081	0.945	−0.002	0.013
	0.500	0.545	0.077	0.909	0.045	0.014	0.545	0.079	0.912	0.045	0.014
	1.000	1.109	0.084	0.746	0.109	0.026	1.109	0.086	0.744	0.109	0.027
10 %	0.000	0.007	0.170	0.945	0.007	0.059	0.007	0.173	0.937	0.007	0.061
	0.500	0.510	0.161	0.947	0.010	0.053	0.509	0.164	0.941	0.009	0.055
	1.000	1.091	0.150	0.888	0.091	0.055	1.027	0.152	0.926	0.027	0.049

Abbreviations: *SE* Standard Error, *CP* 95 % Coverage Probability, *MSE* Mean Square Error, *CSP_UN* Unadjusted Cross Sectional Pooling, *CSP_AD* Adjusted Cross Sectional Pooling; *PLR_UN* Unadjusted Pooled Logistic Regression, *PLR_AD* Adjusted Pooled Logistic Regression, *TDCM* Time Dependent Cox Regression Modeling

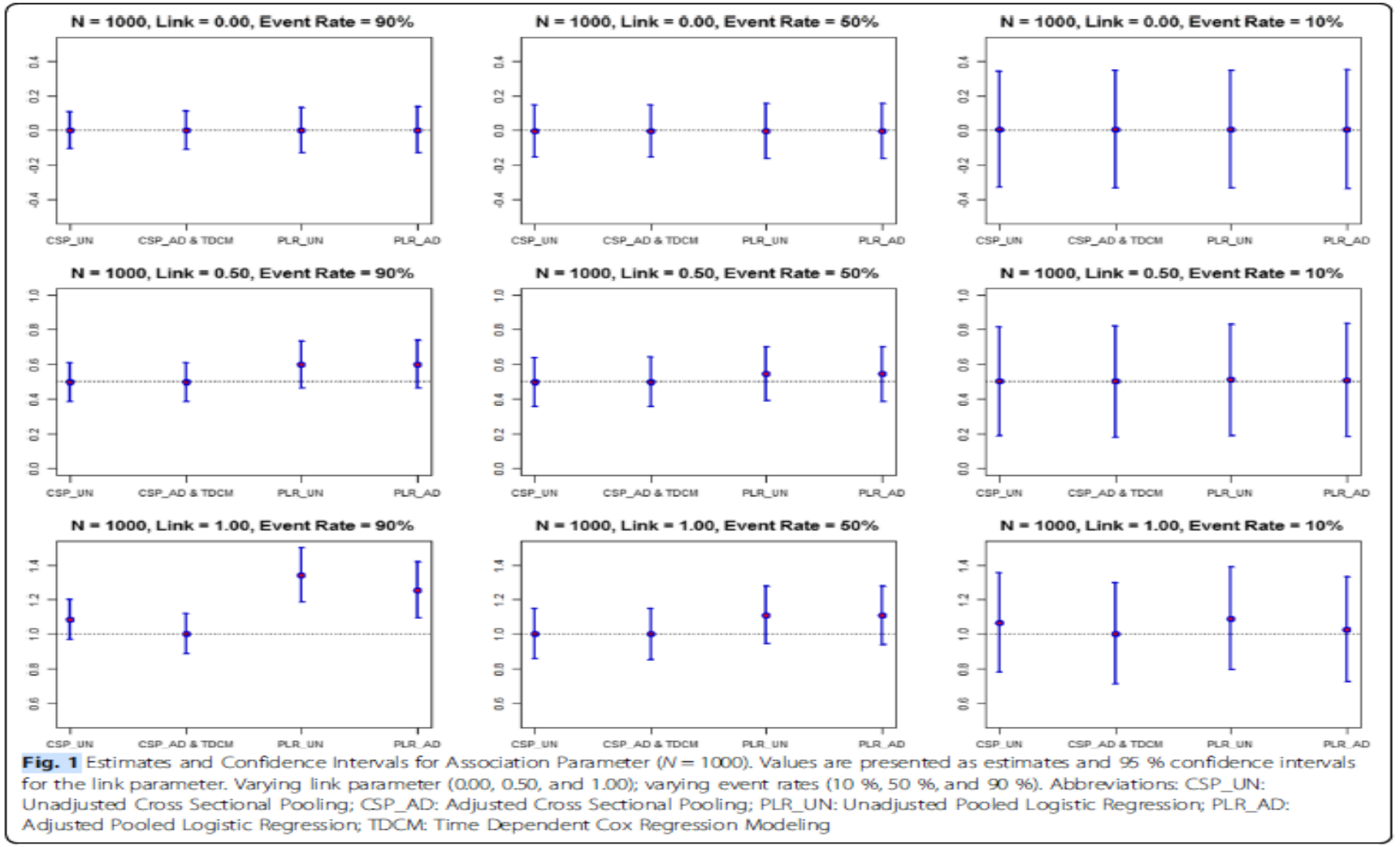


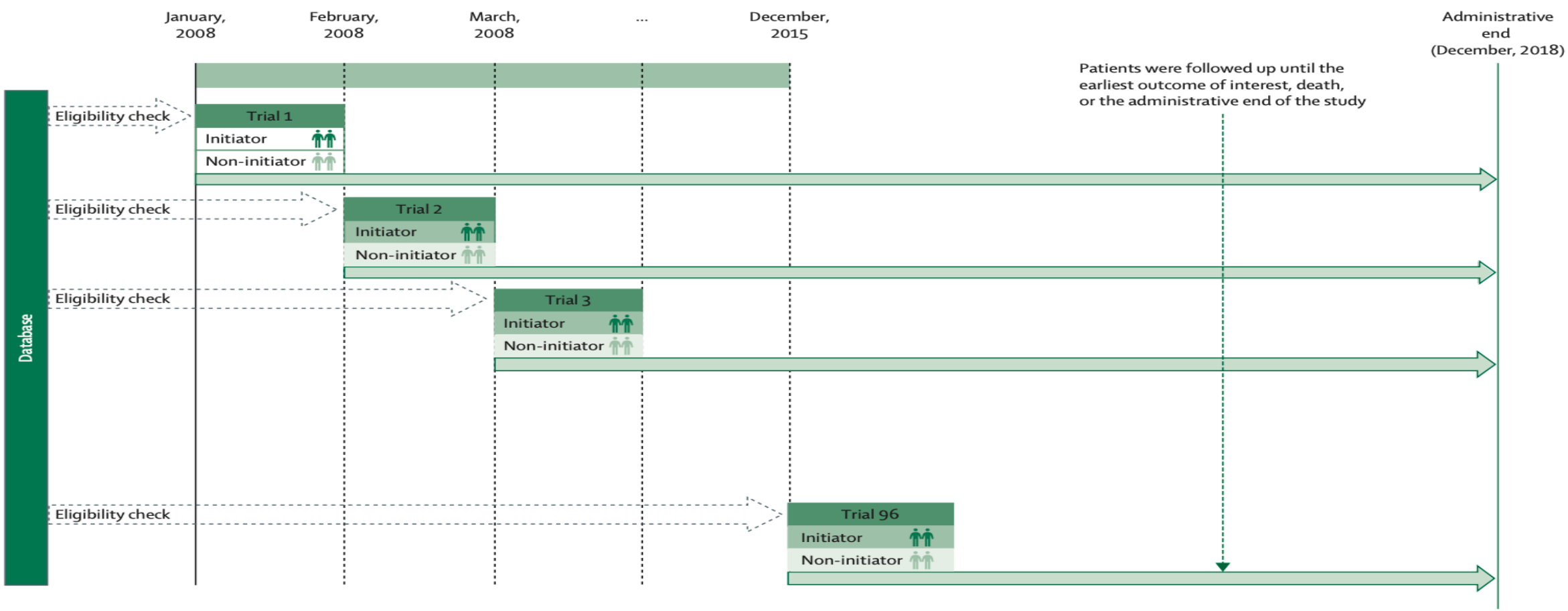
Fig. 1 Estimates and Confidence Intervals for Association Parameter ($N = 1000$). Values are presented as estimates and 95 % confidence intervals for the link parameter. Varying link parameter (0.00, 0.50, and 1.00); varying event rates (10 %, 50 %, and 90 %). Abbreviations: CSP_UN: Unadjusted Cross Sectional Pooling; CSP_AD: Adjusted Cross Sectional Pooling; PLR_UN: Unadjusted Pooled Logistic Regression; PLR_AD: Adjusted Pooled Logistic Regression; TDCM: Time Dependent Cox Regression Modeling

In this scenario, the association estimates were similar among the different methods.

These FHS results are comparable to the simulation results with low event rate (10 %) and moderate association of the longitudinal measures to survival ($\gamma = 0.500$), as shown in Fig. 1.

Methods

- They used pooled logistic regression instead of Cox regression



Supplementary Table 4. Crude incidence rates of the outcome events

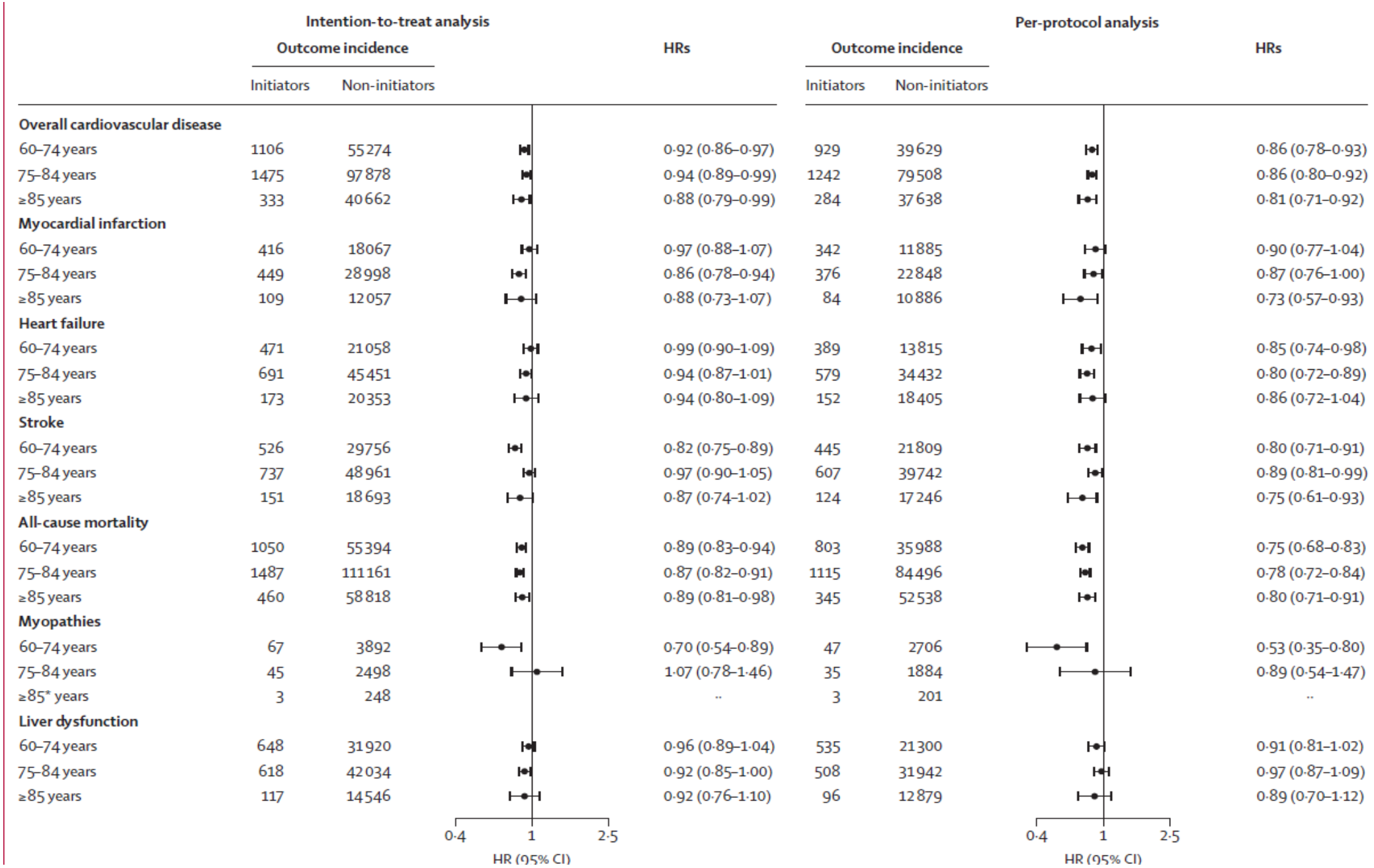
	ITT analysis				Per-protocol analysis			
	Treatment group		Control group		Treatment group		Control group	
	Number of event	Incidence rate (per 1,000 person-years)	Number of event	Incidence rate (per 1,000 person-years)	Number of event	Incidence rate (per 1,000 person-years)	Number of event	Incidence rate (per 1,000 person-years)
60-74 years old								
Overall CVD	1106	42.45 (40.02, 45.03)	55274	37.92 (37.60, 38.23)	929	41.51 (38.92, 44.26)	39629	37.81 (37.44, 38.19)
Myocardial infarction	416	14.81 (13.45, 16.30)	18067	11.50 (11.34, 11.67)	342	14.27 (12.84, 15.87)	11885	10.68 (10.49, 10.87)
Heart failure	471	16.87 (15.42, 18.47)	21058	13.49 (13.31, 13.68)	389	16.29 (14.74, 17.99)	13815	12.55 (12.34, 12.76)
Stroke	526	19.13 (17.56, 20.84)	29756	19.55 (19.33, 19.78)	445	18.90 (17.23, 20.74)	21809	20.22 (19.95, 20.49)
Death	1050	36.22 (34.09, 38.47)	55394	34.43 (34.14, 34.71)	803	32.50 (30.33, 34.83)	35988	31.94 (31.61, 32.27)
Myopathies	67	2.32 (1.82, 2.94)	3892	2.43 (2.35, 2.50)	47	1.91 (1.43, 2.54)	2706	2.41 (2.32, 2.50)
Liver dysfunction	648	23.40 (21.67, 25.27)	31920	20.72 (20.50, 20.95)	535	22.68 (20.84, 24.69)	21300	19.56 (19.30, 19.83)
75-84 years old								
Overall CVD	1475	70.77 (67.25, 74.47)	97878	65.79 (65.38, 66.20)	1242	69.98 (66.20, 73.98)	79508	65.60 (65.15, 66.06)
Myocardial infarction	449	18.98 (17.30, 20.81)	28998	17.36 (17.16, 17.56)	376	18.81 (17.00, 20.81)	22848	17.10 (16.88, 17.33)
Heart failure	691	30.15 (27.98, 32.48)	45451	28.01 (27.75, 28.27)	579	29.77 (27.44, 32.30)	34432	26.62 (26.34, 26.90)
Stroke	737	32.89 (30.60, 35.36)	48961	30.75 (30.48, 31.02)	607	31.99 (29.54, 34.64)	39742	31.02 (30.71, 31.32)
Death	1487	61.04 (58.02, 64.23)	111161	64.74 (64.36, 65.12)	1115	54.25 (51.16, 57.53)	84496	62.21 (61.79, 62.63)
Myopathies	45	1.85 (1.38, 2.48)	2498	1.46 (1.40, 1.51)	35	1.71 (1.22, 2.38)	1884	1.39 (1.33, 1.45)
Liver dysfunction	618	26.50 (24.49, 28.67)	42034	25.57 (25.32, 25.81)	508	25.86 (23.71, 28.21)	31942	24.38 (24.12, 24.65)
≥85 years old								
Overall CVD	333	92.52 (83.10, 103.01)	40662	94.34 (93.42, 95.26)	284	93.30 (83.05, 104.81)	37638	94.38 (93.43, 95.33)
Myocardial infarction	109	26.79 (22.20, 32.32)	12057	24.59 (24.16, 25.03)	84	24.56 (19.83, 30.42)	10886	24.26 (23.80, 24.72)
Heart failure	173	44.43 (38.28, 51.57)	20353	43.41 (42.82, 44.01)	152	46.46 (39.63, 54.47)	18405	43.01 (42.40, 43.64)
Stroke	151	38.55 (32.86, 45.21)	18693	39.87 (39.30, 40.44)	124	37.56 (31.50, 44.79)	17246	40.07 (39.48, 40.67)
Death	460	110.06 (100.45, 120.60)	58818	116.79 (115.85, 117.74)	345	98.53 (88.66, 109.49)	52538	114.79 (113.82, 115.78)
Myopathies	3	0.72 (0.23, 2.23)	248	0.49 (0.44, 0.56)	3	0.86 (0.28, 2.66)	201	0.44 (0.38, 0.50)
Liver dysfunction	117	28.97 (24.17, 34.72)	14546	29.94 (29.46, 30.43)	96	28.37 (23.22, 34.65)	12879	29.10 (28.61, 29.61)

ln 1000 = 42.45
100= ?
100/1000 x 42.45
=4.245

Methods/Statistical Analysis

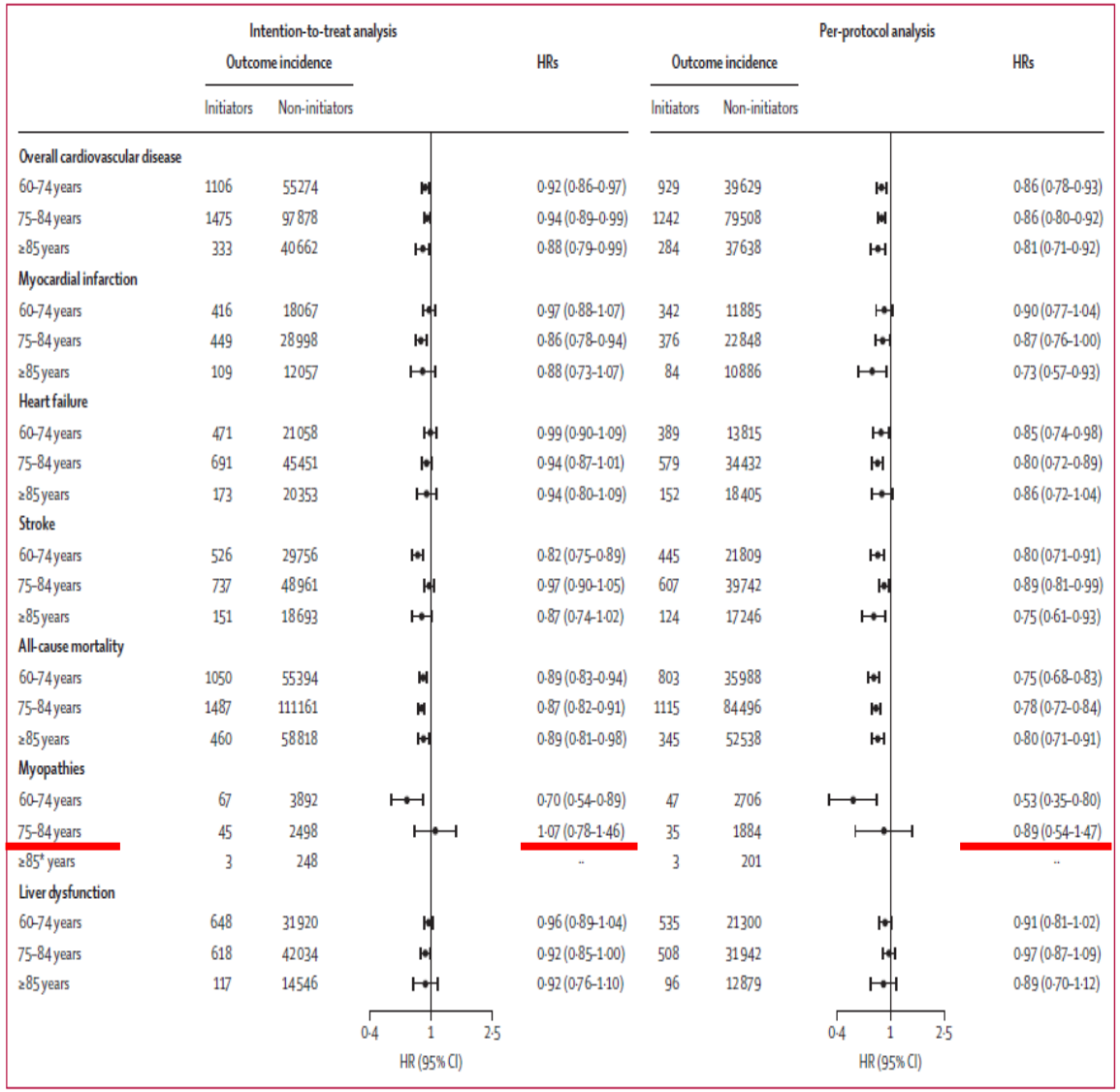
- The main analysis used Inverse probability weighting in the per protocol(PP) analysis but not for the intention to treat analysis(ITT).
- In the per protocol section the following was stated:

1. To adjust for selection bias resulting from the artificial censoring process, each person-trial was weighted at each timepoint by the inverse probability of receiving their assigned treatment strategy, conditional on baseline and time-varying covariates.
2. A pooled logistic model was fitted to predict the probability of receiving statin therapy at each timepoint.
3. The time-varying covariates used in this model included lifestyle behaviors, clinical parameters, comorbidities, drug use, and service use, as detailed above.



The results are similar, so they are reversed. ITT we can not rule out selection bias because the did just regression adjusting

Main Analysis



Sensitivity Analysis

Supplementary Table 10a. Sensitivity analysis of adopting propensity score matching

	ITT analysis		Per-protocol analysis	
	Hazard ratio	95% CI	Hazard ratio	95% CI
60-74 years old				
Overall CVD	0.90	(0.83, 0.98)	0.88	(0.76, 1.03)
Myocardial infarction	0.95	(0.83, 1.08)	0.90	(0.68, 1.18)
Heart failure	0.91	(0.80, 1.03)	0.81	(0.63, 1.04)
Stroke	0.88	(0.78, 0.99)	0.88	(0.70, 1.10)
Death	0.86	(0.79, 0.93)	0.80	(0.67, 0.95)
Myopathies	0.89	(0.63, 1.23)	0.78	(0.40, 1.50)
Liver dysfunction	0.89	(0.81, 0.99)	0.82	(0.66, 1.01)
75-84 years old				
Overall CVD	0.93	(0.87, 1.00)	0.85	(0.75, 0.97)
Myocardial infarction	0.85	(0.75, 0.97)	0.99	(0.78, 1.27)
Heart failure	0.95	(0.85, 1.05)	0.79	(0.64, 0.97)
Stroke	0.96	(0.87, 1.06)	0.85	(0.71, 1.03)
Death	0.85	(0.80, 0.92)	0.79	(0.68, 0.91)
Myopathies	1.24	(0.80, 1.93)	0.41	(0.17, 0.99)
Liver dysfunction	0.90	(0.81, 1.00)	1.15	(0.92, 1.43)
≥85 years old				
Overall CVD	0.81	(0.69, 0.93)	0.68	(0.54, 0.85)
Myocardial infarction	0.87	(0.67, 1.13)	0.70	(0.47, 1.06)
Heart failure	0.94	(0.76, 1.16)	0.81	(0.57, 1.14)
Stroke	0.72	(0.58, 0.89)	0.51	(0.36, 0.71)
Death	0.86	(0.76, 0.97)	0.83	(0.66, 1.04)
Myopathies*	/	/	/	/
Liver dysfunction	0.84	(0.65, 1.08)	0.65	(0.42, 1.00)

* Estimates cannot be obtained due to small number of events

The results are similar, so they are reversed. ITT we can not rule out selection bias because the did just regression adjusting

Results

- The results are similar to some extent but in myopathies its so different with a HR of 1.07 (0.78-1.46) in the ITT and HR of 0.89 (0.54-1.47) in the PP
- In the main analysis ITT, we cannot rule out selection bias because they did just regression adjustment they did not even use PSM or IPW
- Whereas in the main analysis per protocol they did inverse probability weighting
- However, in the sensitivity analysis they used propensity score matching and weighting again, but the result came out as quite different.

Methods/statistical analysis

- **Intention-to-Treat Analysis**:
  - Compared the assigned treatment strategy (statin initiation at baseline) with non-initiation.
  - Included adjustments for baseline covariates and follow-up periods.
- **Per-Protocol Analysis**:
  - Compared continuous statin users with those who never used statins during the follow-up period.
  - Artificial censoring was applied when patients deviated from their assigned strategy.
  - Adjusted for selection bias resulted from censoring using inverse probability weighting for treatment strategy and competing risks (e.g., death).
  - Missing values for time-varying covariates were handled using the last observation carried forward (LOCF) method.

Methods/statistical analysis

Main analysis

- They used regression adjustment in the ITT
- Only one equation for estimating treatment effect not the 2 equations that they usually use like propensity score first then use the score to either match or balance with inverse probability weighting.
- They did not use the first equation, but moved to the outcome model
- But adjusted for several covariate.
- Rather used inverse probability weight in the per protocol analysis and adjusting for censoring bias

Methods used to analyse

Sensitivity Analyses:

- Eight was conducted to test the robustness of results by varying assumptions, such as

1. shortening the gap for statin discontinuation,
2. truncating weights,
3. using propensity score matching, and

4 & 5. excluding participants with early outcomes.

Sensitivity Analyses

6. Removing the exclusion criteria requiring at least one follow-up visit during the follow-up period, and instead censoring patients 2 years after their last visits to the local health system

7. Competing risk adjustment in the intention-to-treat analysis was conducted.

8. Multiple imputation was used to handle missing baseline data.

Sensitivity analysis

- If we have to predict the causal effect, we have to show that it is not bias by unobserved or uncontrolled confounders, so we calculate and measure confounder.
- This is done by using E value to measure the confounder.

VanderWeele TJ, Ding P. Sensitivity analysis in observational research: introducing the E-value. Annals of internal medicine. 2017;167(4):268-74.

E value/interpretations

- The main effect size in the paper is HR 0.94 [95%CI 0.89 – 0.99]
- So, in Stata we used this command evalue HR 0.94, lcl(0.89) ucl(0.99) fig
- evalue or 0.94, lcl(0.89) ucl(0.99) fig
- E-value (point estimate): 1.324
- E-value (CI): 1.111

- The observed OR of 0.94 could be explained away by an unmeasured confounder that associated with both the treatment and the outcome by a OR of more than 1.3-fold each, above and beyond the measured confounders, but weaker confounding could not do so. It is good since its close to 1

Stata/SE 18.0

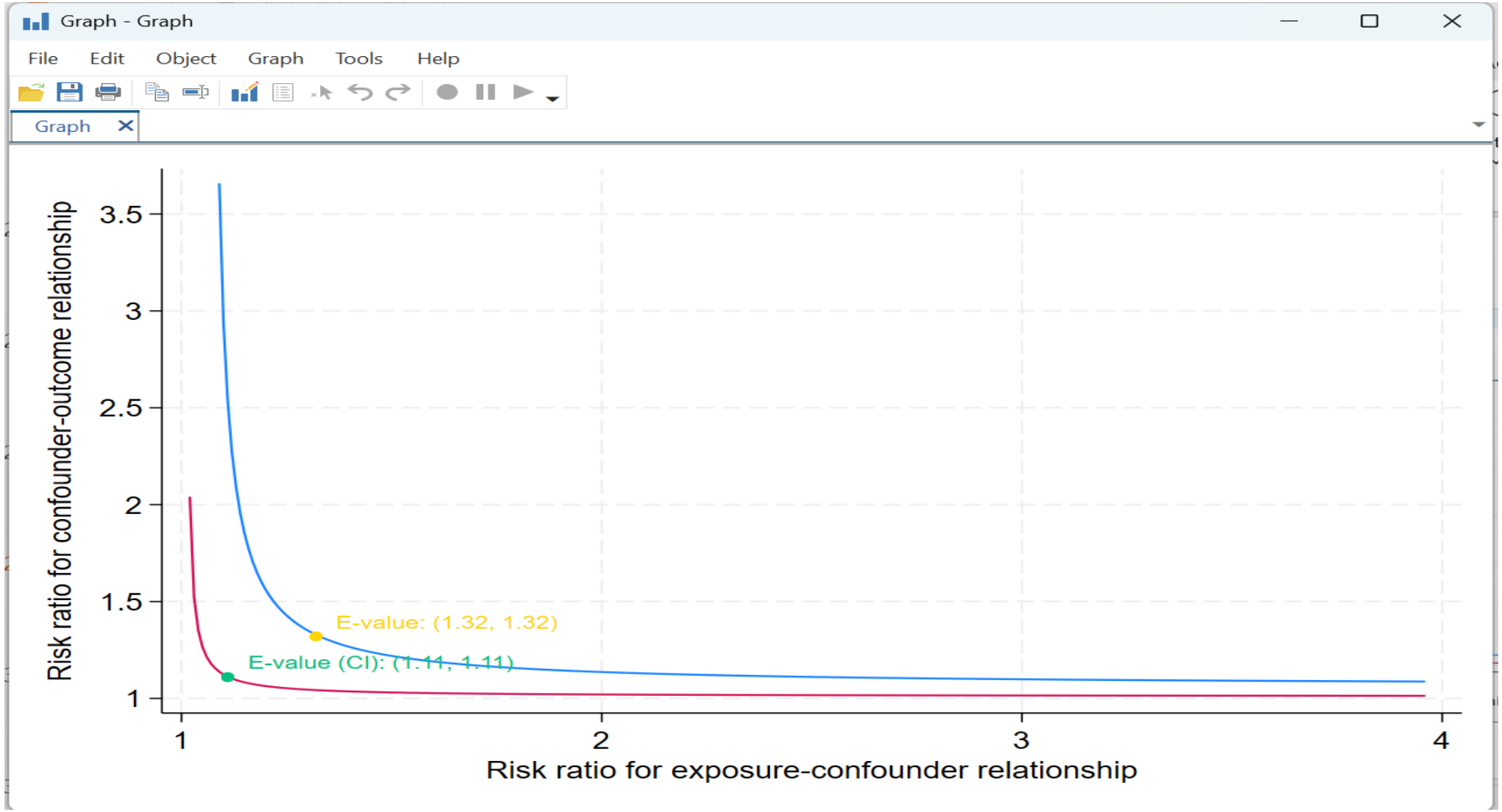
File Edit Data Graphics Statistics User Window Help

H... Filter comman # Comma... 1 evalue o...

```
  /__    /   ____/   /   ____/      18.0
 ___/   /   /___/   /   /___/       SE—Standard Edition

 Statistics and Data Science       Copyright 1985-2023 StataCorp LLC
                                   StataCorp
                                   4905 Lakeway Drive
                                   College Station, Texas 77845 USA
                                   800-STATA-PC        https://www.stata.com
                                   979-696-4600        stata@stata.com

Stata license: 15-user network perpetual
Serial number: 401806310883
  Licensed to: Clinical Epidemiology and Biostatistics
               Ramathibodi hospital

Notes:
      1. Unicode is supported; see help unicode_advice.
      2. Maximum number of variables is set to 5,000 but can be increased;
          see help set_maxvar.
      3. New update available; type -update all-

. evalue or 0.94, lcl(0.89) ucl(0.99) fig
   E-value (point estimate): 1.324
   E-value (CI): 1.111

.
```

Command

E value graph



Conclusion

- This paper results from the main analysis are bit different from the propensity score matching in the sensitivity analysis.
- For the effectiveness, it's quite straightforward that it helps but for the adverse event especially the myopathies in the 75-84 age group statin may fare worse than no statin
- But the result are controversial, however, we can't deny completely the possibility of being worse.

THANK YOU