Estimating the Per-Protocol Effect of Lithium on Suicidality in the VA CSP-590 Trial

Alejandro Szmulewicz

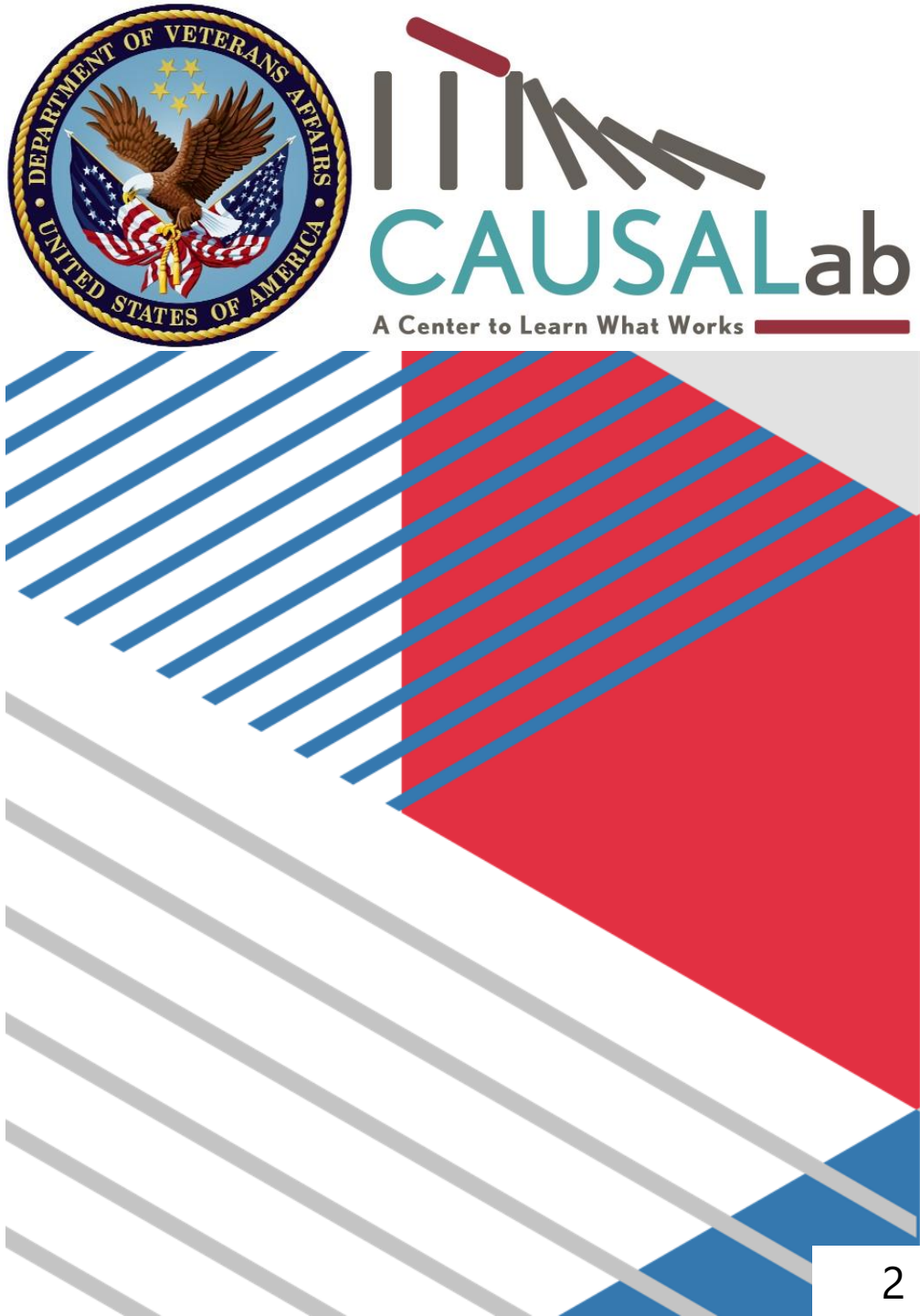
MD PhD

CAUSALab, Department of Epidemiology

Acknowledgements and Disclosures

- This research was supported by the U.S. Department of Veterans Affairs (VA) Office of Research and Development (ORD) Cooperative Studies Program (CSP) Epidemiology Center at the VA Boston Healthcare System through CSP #2032, by resources and the use of facilities at the VA Boston Healthcare System and VA Informatics and Computing Infrastructure (VINCI) (VA HSR RES 13-457).

- The VA CAUSAL Methods Core is a collaboration between the Massachusetts Veterans Epidemiology, Research, and Information Center (MAVERIC) Division of Population Health and Data Sciences and the CAUSALab at the Harvard T.H. Chan School of Public Health.



Overview

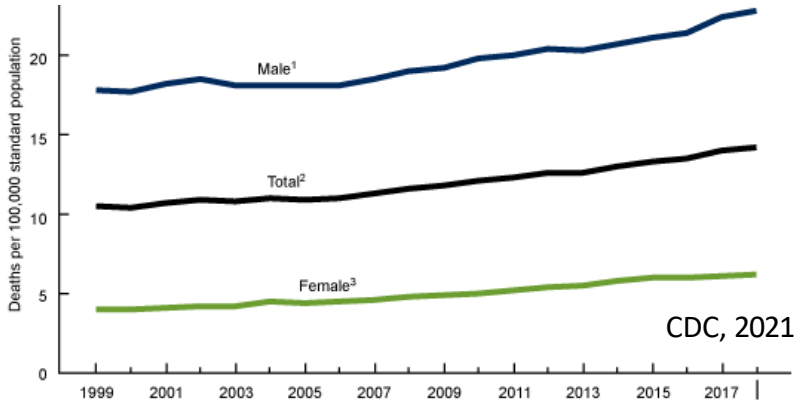
Background: presentation of clinical problem

Per-protocol analysis of CSP590: Re-analysis of a recent RCT adjusting for non-adherence

Moving forward: target trial emulation using VA observational databases

Suicide is a major public health concern

- Suicide in the US
  - ~50,000 deaths annually
  - 2nd cause of death among youth
  - 35% increase since year 2000



CDC, 2021

- Mood disorders entail a heightened risk of suicide
  - Lifetime risk in **bipolar disorder**: 30x general population
  - Lifetime risk in **major depression**: 10x general population

- Need/opportunity for targeted prevention efforts
  - Interventions to lower suicide risk among individuals with mood disorders

Treatment of mood disorders: Major depression, Bipolar disorder

- **Major depression disorder** is characterized by depressive episodes
  - First-line treatment are SSRI antidepressants
  - In case of failure, combination with antipsychotic, *lithium*, thyroid hormone, …
- **Bipolar disorder** is characterized by depressive episodes plus manic/hypomanic episodes
  - First line treatment are mood stabilizers: *lithium*, lamotrigine, valproic acid, carbamazepine
- There is a need to understand the comparative effectiveness of these treatments to prevent suicide

Lithium is a candidate preventative agent

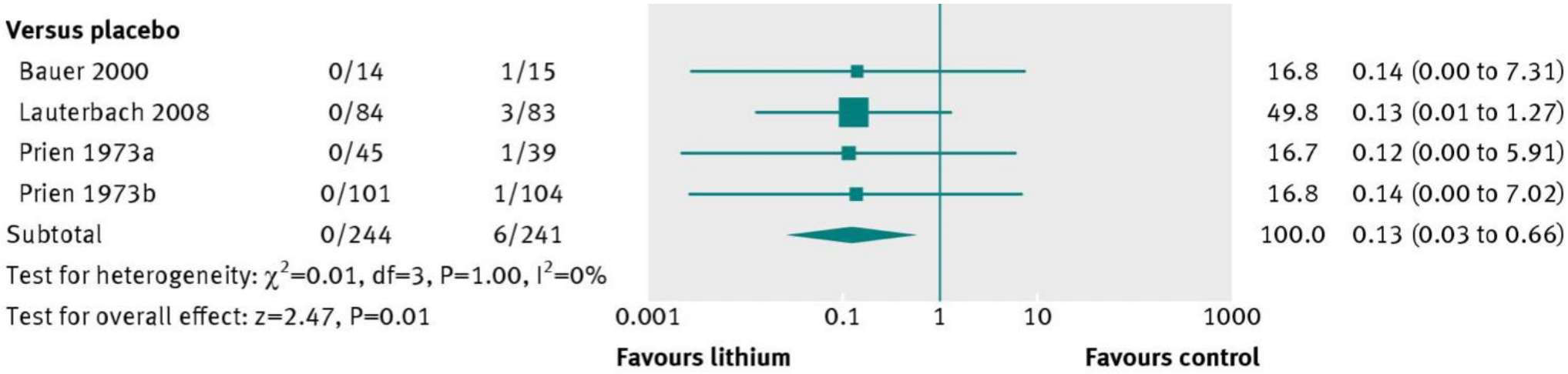
- Observation from early randomized trials (1970-1990) suggesting that patients in the lithium arm did not have suicides

- Systematic review of RCT *not specifically designed to look at suicide* comparing lithium initiation vs. no initiation:
  - **OR: 0.13 (95% CI: 0.02, 0.76)**
  - **Cipriani et al. (2013) BMJ**

- **Clinical guidelines**

- Lithium:
  - continue lithium in patients who needed lithium augmentation of antidepressants in acute treatment (B),
  - consider adding lithium to antidepressants in patients at high risk of relapse (B) or suicide (A),
  - do not routinely use lithium as monotherapy for relapse prevention but consider as a second-line alternative to antidepressants (B).

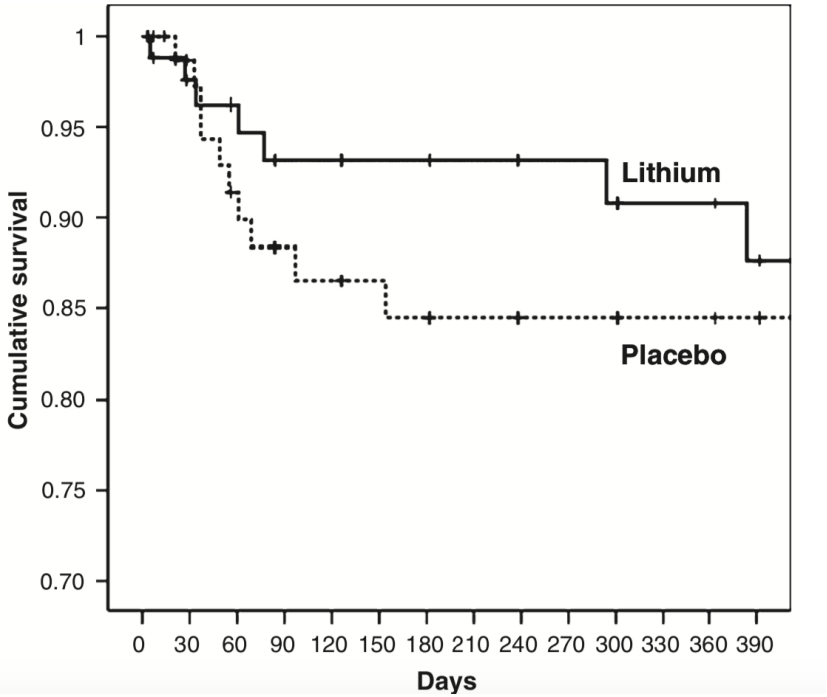
Cipriani et al. (2013)
Suicide mortality, lithium vs. placebo



- Two meta-analyses with different analytical approaches show unstable estimates
  - Nabi et al. 2022: 0.41 (0.03, 2.49)
  - Riblet et al. 2022: 0.30 (0.05, 1.15)

Pragmatic randomized trials: Lauterbach et al. (2008)

- 167 patients with bipolar disorder or major depression randomized to: (a) lithium or (b) placebo
- Primary outcome: "***suicidal acts***" (assessed by participant's report)



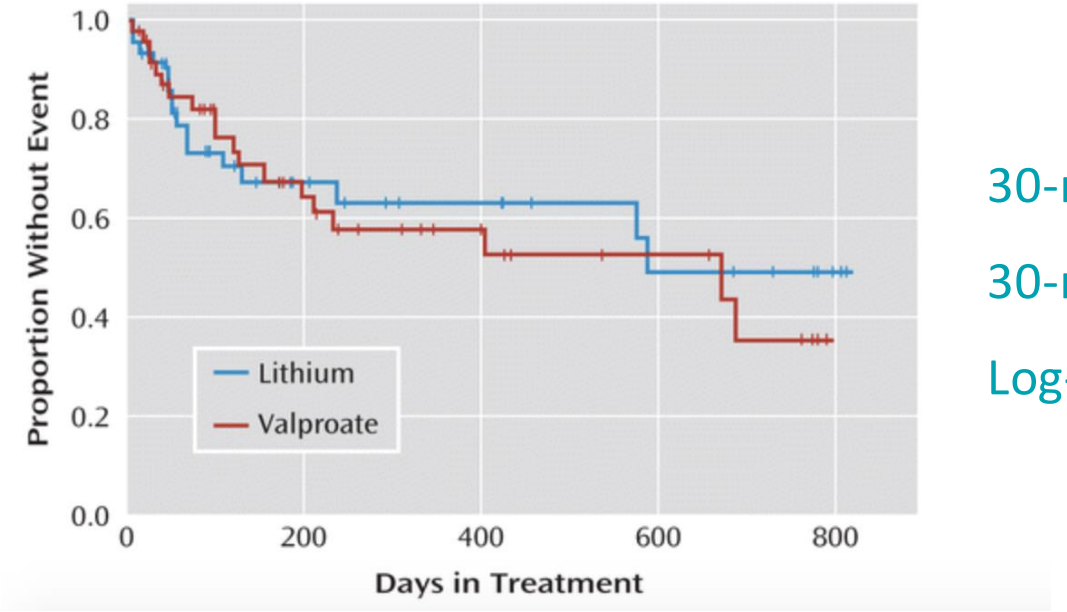
24-month risk in lithium: 8.3%

24-month risk in placebo: 12.0%

Hazard ratio: 0.52 (95% CI: 0.19, 1.44)

Pragmatic randomized trials
Oquendo et al. (2011)

- 98 patients with bipolar disorder randomized to: (a) lithium or (b) valproic acid
- Primary outcome: "***suicidal act***" (a composite of attempt, hospitalization or medication change in response to suicide)



30-month risk in lithium: 12.6%

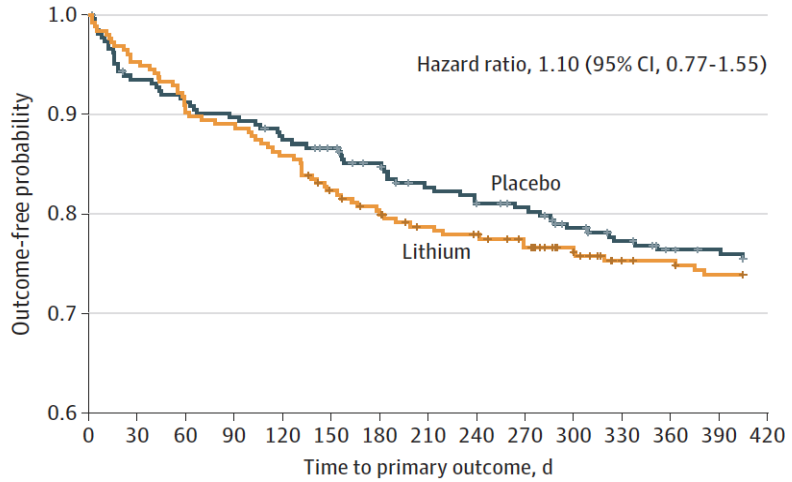
30-month risk in valproic acid: 16.3%

Log-rank test p-value > 0.05

Pragmatic randomized trials
CSP-590 (2022)

- 519 patients with bipolar disorder or major depression and a prior suicide attempt randomized to: (a) lithium or (b) placebo
- Primary outcome: suicidality

Figure 2. Time to Primary Outcome in the Lithium and Placebo Groups



No. at risk

Placebo	264	245	239	235	229	223	214	204	200	195	186	179	171	169	0
Lithium	255	243	232	227	219	207	200	192	189	181	173	163	160	156	0

12-month risk in lithium: 25.5%

12-month risk in placebo: 23.5%

Hazard ratio: 1.10 (95% CI: 0.77, 1.55)

Does lithium prevent suicide?

	Randomized trials	*Observational studies*
Main strength	Unbiased estimates at baseline	Large sample sizes
Outcome	*Suicidality*	Suicidality and suicide
Follow-up	1-2 years	> 5 years
Causal contrasts	Intention-to-treat	Observational analog to intention-to-treat
Results	Mostly compatible with no effect Pooled OR: 0.60 (0.32, 1.27)	Strong protective effects Pooled OR: 0.20 (0.16, 0.26)
Examples	Lauterbach et al. (2008); Oquendo et al. (2012); Girlanda et al. (2014); Katz et al. (2022), …	Goodwin et al (2003); Kessing et al. (2005); Hayes et al. (2016); Song et al. (2017); Smith et al. (2009, 2014, 2015, 2022)

Discrepancy between clinical intuition and trial's results

- Potential explanations:
  - Low adherence to lithium
  - Suicidality is a bad surrogate outcome for suicide

Overview

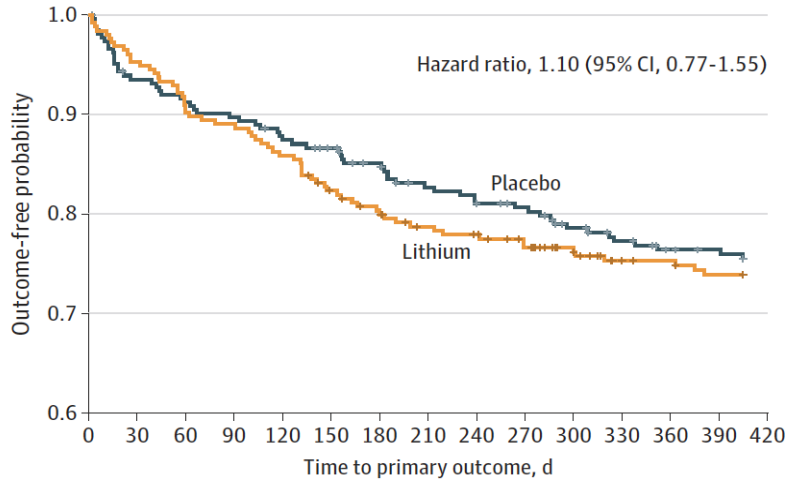
Background: overview of an ongoing debate

Per-protocol analysis of CSP590: re-analysis of a recent RCT adjusting for non-adherence

Pragmatic randomized trials
CSP-590 (2022)

- 519 patients with bipolar disorder or major depression and a prior suicide attempt randomized to: (a) lithium or (b) placebo
- Primary outcome: suicidality

Figure 2. Time to Primary Outcome in the Lithium and Placebo Groups



No. at risk

Placebo	264	245	239	235	229	223	214	204	200	195	186	179	171	169	0
Lithium	255	243	232	227	219	207	200	192	189	181	173	163	160	156	0

12-month risk in lithium: 25.5%

12-month risk in placebo: 23.5%

Hazard ratio: 1.10 (95% CI: 0.77, 1.55)

The CSP-590 trial was stopped for futility

- The ITT absolute risk reduction was 2.0% (favoring placebo) and hazard ratio 1.10 (0.77, 1.55)
  - The alternative hypothesis set by the Data Monitoring Committee was -9.7% and 0.34, respectively
  - On the basis of assumptions about future event rates, the trial was stopped (1,862 patients were anticipated)

Intention-to-treat effect

- The effect of being assigned to the treatment strategies, regardless of treatment actually received:
  - In the presence of significant deviations from protocol, if the treatment has an effect, will be **closer to the null** than the actual effect of treatment.

- Significant non-adherence rates:
  - 53% in Oquendo et al. (2012)
  - 60% in Lauterbach et al. (2008)
  - 83% in Katz et al. (2022)

Per-protocol effect

- The effect of receiving the treatment strategies as specified in the protocol

- Because the CSP-590 trial was stopped for futility, it would be ideal to know the per-protocol effect

  - to make a more informed decision of whether to stop the trial or not

  - to report a more meaningful clinical effect

What if patients had adhered to their assigned strategy throughout the trial?

- Initiate lithium at baseline and continue unless
  - ❑ serious side effects emerge (renal failure, neurotoxicity),
  - ❑ clinical concerns arise,
  - ❑ potentially dangerous drug-drug interactions (e.g., diuretics)

Per-protocol analysis: Naïve and non-naïve approaches

- Naïve per-protocol analysis:
  - **Step 1**: Censor individual's data once we have evidence of lack of adherence to the assigned strategy (if any).
  - **Step 2**: Conduct the analysis in the restricted per-protocol sample.

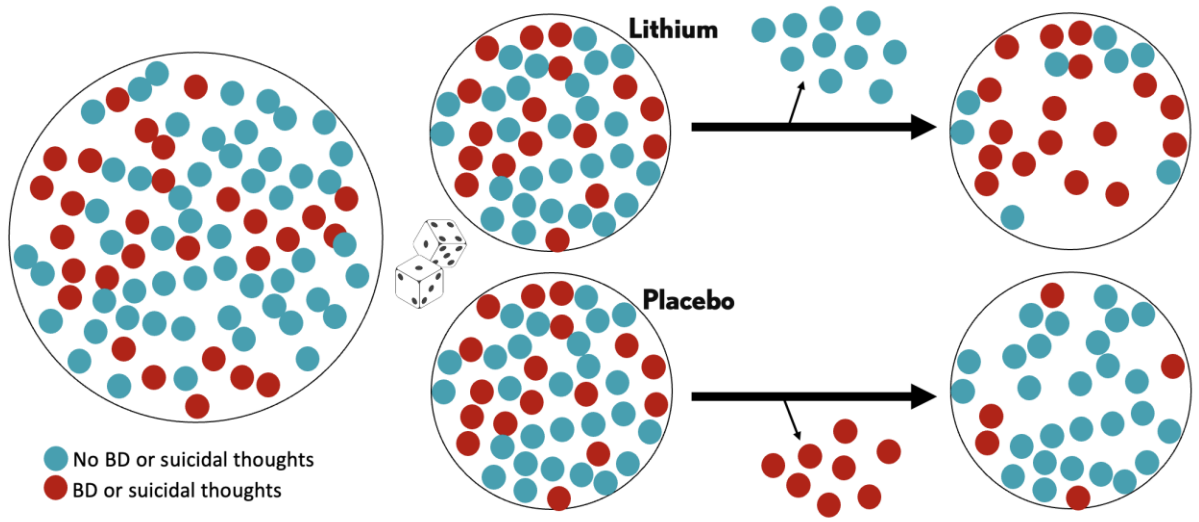
CSP-590 (2022)
Odds ratios of non-adherence

Lithium

- BD 0.77 (0.50, 1.09)
- Suicide thoughts 0.92 (0.42, 1.77)

Placebo

- **BD 1.31 (0.80, 2.16)**
- **Suicide thoughts 2.34 (1.20, 4.56)**



Per-protocol *analysis:* Naïve and non-naïve approaches

- Even in a randomized trial, treatment decisions after baseline (e.g., adherence) are not randomized:
  - Risk for confounding
  - If the post-baseline confounding is affected by prior treatment, adjustment using conventional methods (e.g., outcome regression or propensity score matching) will not work
  - Need to use g-methods (e.g., inverse probability weighting)

Non-naïve per-protocol analysis of CSP-590 Methods

- Naïve per-protocol analysis:
  - **Step 1**: Censor individual's data once we have evidence of lack of adherence to the assigned strategy (if any).
  - **Step 2**: Conduct the analysis in the restricted per-protocol sample.

- **Non-naïve** per-protocol analysis :
  - **Step 1**: Censor individual's data once we have evidence of lack of adherence to the assigned strategy (if any), *unless they stopped treatment for clinical reasons*
  - **Step 2**: Conduct the analysis in the restricted per-protocol sample, *after adjusting for confounding due to incomplete adherence*.

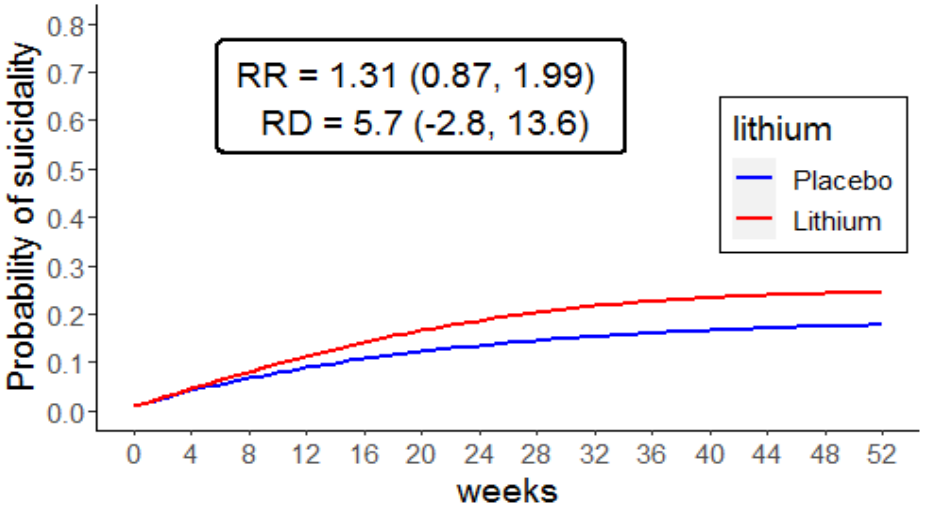
Non-naïve per-protocol analysis of CSP-590 Methods

- We adjusted for baseline
  - age, sex, race, diagnosis of PTSD, BD, personality disorder, substance abuse or dependence, prior suicide attempts

  and post-baseline
  - depression (PHQ-9), suicidal thoughts (C-SSRS), antipsychotic use, emergency room visits

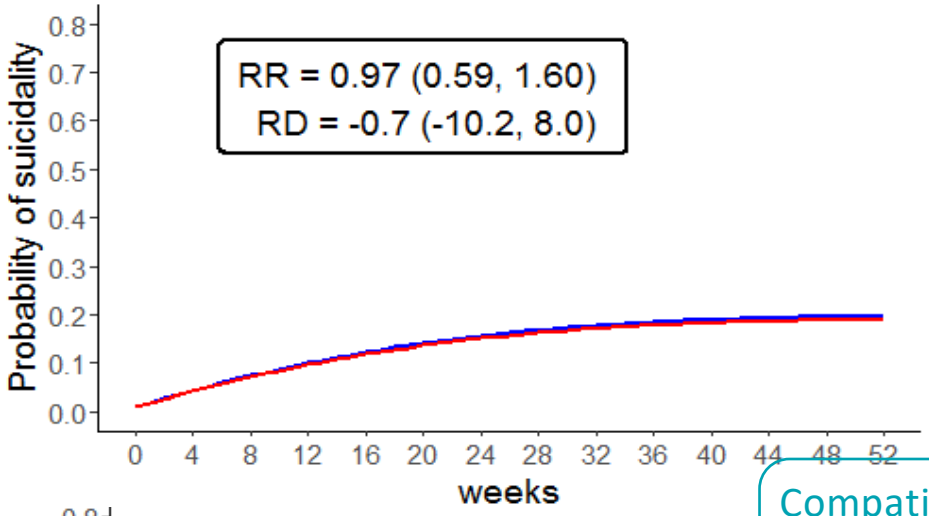
  prognostic factors associated with adherence via inverse probability weighting

Non-naïve per-protocol analysis of CSP-590 Results



mITT + censoring

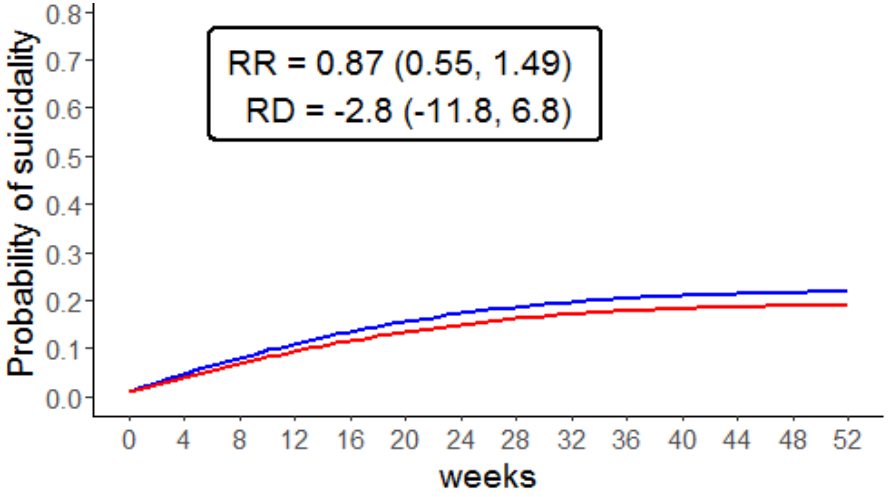
[88 events]



censoring at nonadherence

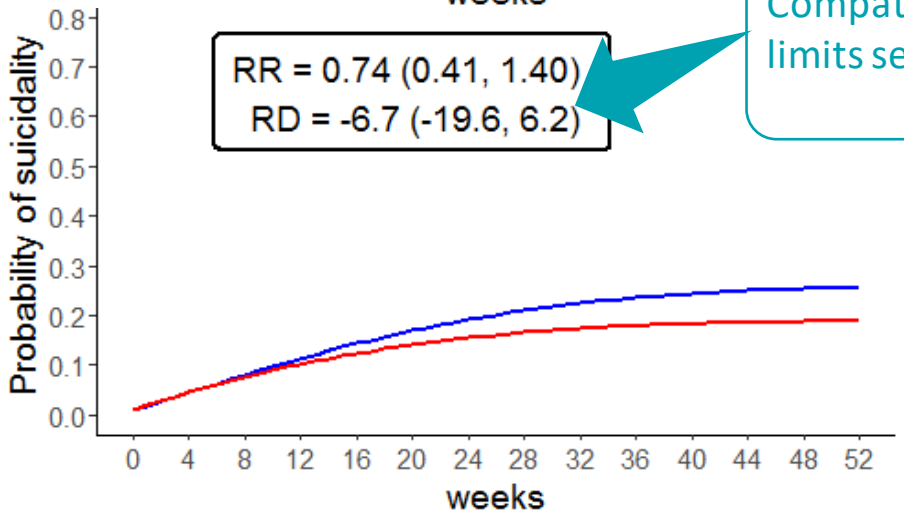
(unadjusted)

[49 events]



censoring at nonadherence

(adjusted for baseline variables)



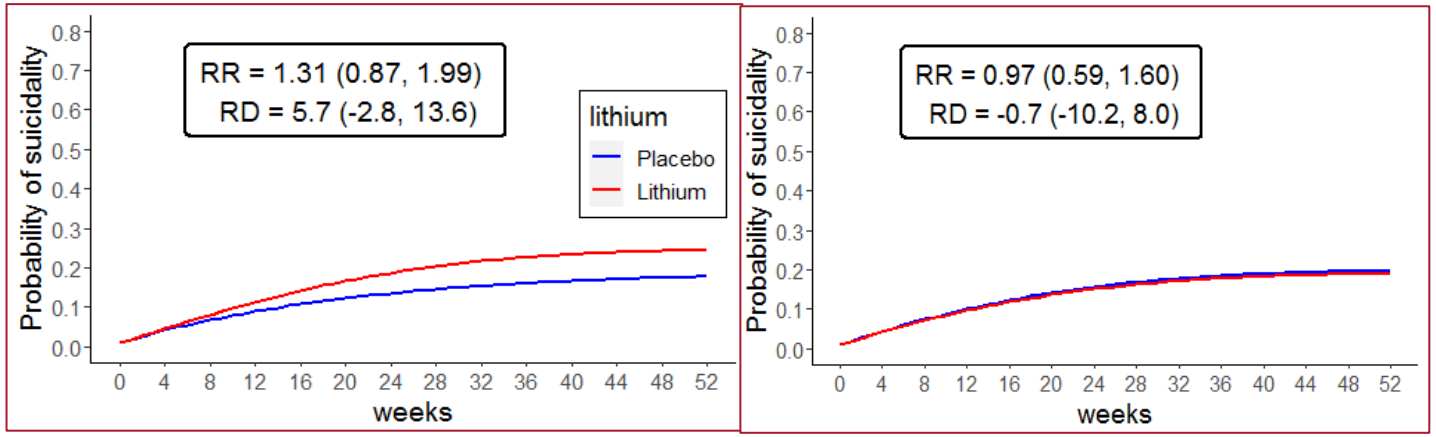
Compatible with the limits set by the DMC !

censoring at nonadherence

(adjusted for baseline and time-varying variables)

Non-naïve per-protocol analysis of CSP-590 Discussion

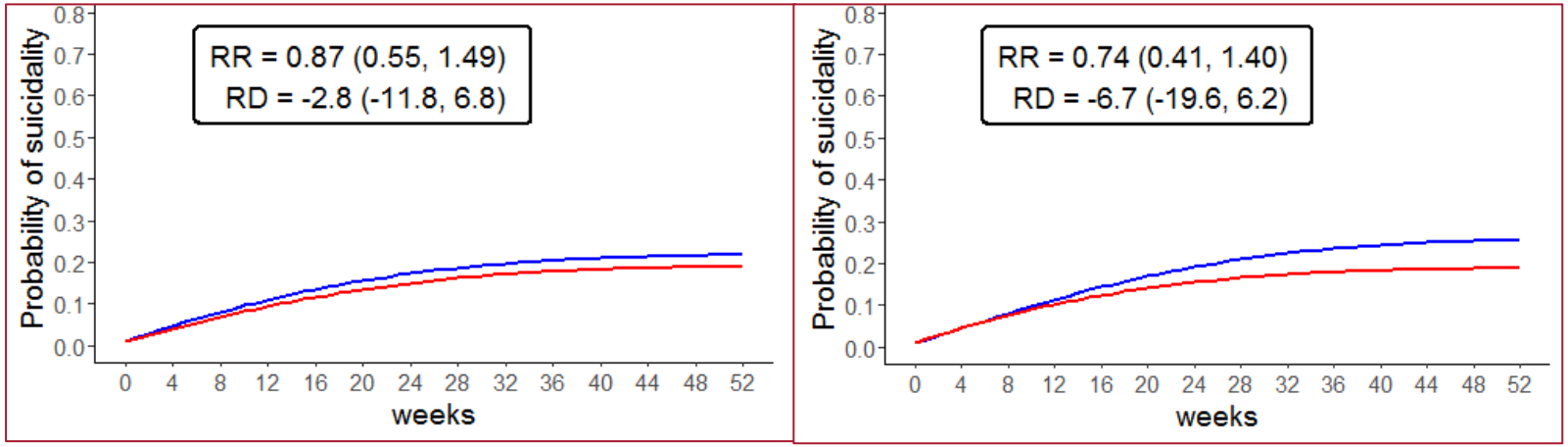
- The effect of censoring at nonadherence
  - The risk of suicidality was higher in individuals who deviated from protocol in the **lithium group** than among those who deviated from protocol in the **placebo group**



Main analysis (12-month risk)	
Censoring at non-adherence (12-month risk)	↓

Non-naïve per-protocol analysis of CSP-590 Discussion

- The effect of adjusting for baseline and time-varying confounding
  - The per-protocol restriction **overloads the lithium arm** with patients with baseline markers of severity (e.g., bipolar disorder).



Non-naïve per-protocol analysis of CSP-590 Conclusions

- The estimated per-protocol effect of lithium on suicidality is consistent with a **protective** effect:
  - but confidence intervals are wide
  - further research with larger sample sizes are required

Overview

Background: overview of an ongoing debate

Per-protocol analysis of CSP590: re-analysis of a recent RCT adjusting for non-adherence

Moving forward: target trial emulation using VA observational databases

Discrepancy between clinical intuition and RCT results

- Potential explanations:
  - Low adherence to lithium in most RCT
  - Suicidality is a bad surrogate outcome for suicide

Alternatively, can the trial findings be negative because of the use of suicidality and not suicide as the outcome?

	Suicide	*Suicidality*
Nomenclature	Completed suicide, suicide death	Suicidal thoughts and behaviors
Epidemiology	More frequent in men and adults	More frequent in female and youths
Frequency	Infrequent event	30 times more common
Methods	Higher lethality (firearm)	Lower lethality (overdose, cutting)
Genetics	Higher heritability ($h^2_{SNP} = 0.30$)	Lower heritability ($h^2_{SNP} = \sim 0.03$)
Environment	Gun accessibility, antisocial personality	Social isolation, interpersonal and work-related events
Predictors	Non-affective and affective psychosis	Social isolation, anxiety

Edwards et al. 2021, Am J Psych

Target trial emulation using VA observational databases

- RCT can provide unbiased estimates but are unable to study suicide
- Observational studies can study rare outcomes, but are susceptible to confounding by indication (suicide ideation, clinical status)
  - Highly protective effects in studies before 2013 (e.g., HR: 0.44, Kessing et al. (2005))
  - Harmful effects in studies after 2013
    - Smith et al. 2014: 1.22 (0.82, 1.81)
    - Smith et al. 2022: 1.50 (1.05, 2.15)
- Moving forward, the optimal path seems to be to combine strengths from randomized trials and observational studies

The randomized trial as a benchmark

- We can use the estimates from CSP590 as a benchmark:
  - and compare that result with the emulation of a target trial with the same protocol as CSP590.

- This procedure ensures that the same causal question is asked in the same population and same healthcare setting

- Then, we can more confidently, extend to
  - longer follow-up periods
  - suicide as the outcome

How to improve observational analyses?

A combination of

- High-quality data
- Sound methodology
- Subject-matter knowledge

Lithium for the prevention of suicide: A target trial emulation using VA databases

- High-quality data
  - **VA Office of Mental Health and Suicide Prevention** dataset can be used to supplement and validate suicide risk factors data
  - Veterans Health Administration directive outlines policy and guidance for the proper use of **Patient Record Flags** to identify patients that are at high risk for suicide.
- Sound methodology
  - emulating a hypothetical target trial (similar to CSP590) by comparing initiators vs. non-initiators of lithium.
- Subject-matter knowledge
  - guiding the choice of confounders
  - understanding singularities of the databases

Thank you! Questions?