

Sample size bias in the empirical assessment of the acute risks associated with Daylight Saving Time transitions*

José María Martín-Olalla

Universidad de Sevilla, Facultad de Física, Departamento de Física de la Materia Condensada, ES41012 Sevilla, Spain[†]

Jorge Mira

*Universidade de Santiago de Compostela, Facultade de Física,
Departamento de Física Aplicada and iMATUS, ES15782 Santiago de Compostela, Spain[‡]*

(Dated: September 5, 2022)

The assessment of the acute impact of Daylight Saving Time (DST) transitions is a question of great interest for an understanding of the benefits and inconveniences of a practice that is now under public scrutiny in Europe and America.

Here we report a thorough analysis of a record of twelve well-known research studies that reported increased risks associated with DST transitions in health issues —acute myocardial infarction, ischemic strokes— and in societal issues —accidents, traffic accidents and fatal motor vehicle accidents—.

We found that a 5% increase of the risks suffices to understand the reported increased risks associated with the spring transition. Reported values above this threshold are impacted by the sample size of the study.

In the case of the autumn transition, no increase of the risks is found.

Keywords: DST; summer time; latitude; sleep deprivation; spring transition; Europe; insolation; season; motor vehicle accidents

Sample size is the best-known study limitation of any research based on a sampled observation. Any finite sample size translates in a limited knowledge of the parent population that may have originated the observation. This limitation is essentially contained in the size of the confidence interval associated with the reported point estimate[1]. Researchers push hard to study large samples in the hope of a better determination of the effect. Nonetheless it is the case that the same phenomenon is analyzed by vividly different sample sized studies, which gives rise to a distribution of sampled results. Because every research study is analyzed on a solo basis, it is infrequent that the sample of reported point estimates is analyzed consistently in relation to the distribution of the sample sizes.

We bring here a comprehensive analysis of the research studies on the acute effects that Daylight Saving Time (DST) transitions may cause on public health and on societal issues. DST is the seasonal biannual changing of the clocks, a long-standing practice by which modern, extratropical societies —chiefly in America and Europe— adapt the phase of their daily rhythms to the early summer sunrises and to the late winter sunrises. The current criticism against the practice of DST focuses the disruption on human circadian rhythms it brings. The spring transition —when clocks are shifted forward—

charges with the burden of the proof because the advance in the phase of human social rhythms is accompanied by a sleep deprivation which, eventually, may give rise to the increase in the incidence of acute diseases and accidents.[2, 3]

DST transitions set a natural experiment in which every individual participates. Yet when it comes to a research study that assesses the correlation between transition dates and societal issues, things are further limited: it is only one health or one societal issue that is analyzed in a limited region —a country, some region, or some hospital— and during a limited period of time —one year, a few years, one or two decades. All these factors impact the sample size of the study that can vary by some orders of magnitude, whereas the point estimate of the effect ranges from a factor 0.8 to a factor 1.9 —where no effect is associated with 1. Eventually, and in view of these results, governments in Europe and America are considering discontinuing this practice in the fear that its impact on human health may be more severe than previously thought[4], where, understandably, the largest point estimates may have been overrated and the role of the sample size underrated.

We analyze a set of thirteen well-known research studies that associated the incidence of acute myocardial infarction, ischemic strokes or traffic accidents with DST transition dates. Many of them were retrieved from review reports[2, 5]. They are frequently cited in review literature to support the discontinuity of DST practice[6, 7]. They are also cited in ex-post impact assessment reports from the European Parliament[4] or from the German Parliament[8] and in technical reports[9]. We show that the parent Poissonian population that may have originated the distribution of incidence ratios in this set of



* olalla@us.es; <https://orcid.org/0000-0002-3750-9113>;

<https://ror.org/03yxnp24>

[‡] jorge.mira@usc.es; <https://orcid.org/0000-0002-6024-6294>;

<https://ror.org/030eybx10>

thirteen well-known research studies shows an increase of the risk around 5% in the spring transition. Any reported IR above this threshold is likely impacted by the sample size of the study.

METHODS

The point estimate of the acute impact of DST transitions on human health is primordially assessed by the incidence ratio $IR = O/N$: the ratio of the observed counts O in a societal issue after a transition —also known as the study group— to the expected counts N in the same societal issue —the control group. The window of time to compute acute effects is the week after the transition. Research studies report either the week IR or the day IR in the first seven days following a transition, or both.

We included in our analysis research studies which, in addition to IR, reported the total counts in the control group N or this number could be deduced from the reading of the manuscript.

We found seven acute myocardial infarction (AMI) studies that met these requirements: six reported O in the first seven days following the spring and autumn transitions and deduced N from various estimates[10–15]; one reported the IR for the first seven days after a transition[16] and N could be deduced from the total number of AMI in their catalogue. We also considered one ischemic stroke study[17] reporting IR in the first seven days after a transition.

We also considered a series of four accident-related studies: accidental deaths[18], traffic accidents[19], road accidents[20] and fatal motor vehicle accidents[21]. The first two studies reported O and N ; the last two studies reported IR and N could be deduced from their catalogue.

Finally, we considered one study on trauma admissions in Austria, Germany and Switzerland[22] which reported O and N for the week after and prior to the transition, excluding the transition date. This study did not report IR.

In addition to day IR and week IR we also analyzed a set of stratified IR associated with patient characteristics and co-diagnoses —among others: gender, age group, cholesterol in the first week.[11, 12, 16]. We also considered the breakdown by accident type —accident, motorcycle accident, pedestrian, among others— in the trauma study[22].

Every of these studies were natural experiments with little design, in this way they are all comparable. The authors collected a dataset —such as the Swedish registry of AMIs[10], the National Highway Traffic Safety Administration registry of motor vehicle accidents[21] or the TraumaRegister DGU of the German Trauma Society— and crunched the numbers to get IR.

With the study group determined by the natural experiment, the control group —a contrafactual assessment of the events that would have been collected if the

clock changing had not occurred— is the key parameter that gives rise to the reported IR. Studies differ in the way the control group is counted. Older, seminal studies take N as the number of events just before the transition,[19] or an average of event just before and well after the transition.[10] Modern approaches consider sophisticated models that take into account seasonal and societal variations.[20, 21] We do not distinguish the ones from the others and assume every author choice of N and, therefore, IR is a solid point estimate.

In our analysis we identify the total counts in the control group with the sample size of the study. Eventually the number comes from the population size of the region analyzed, the number of events —strokes or accidents— that this population produces every year, and the number of years in the catalogue. Table 1 summarizes the main characteristics of every study.

Every study reported the 95% confidence interval (CI) associated with their point estimate of IR. Our goal in this meta-analysis is to test the distribution of reported IR and sample sizes N within the Poisson statistics. For this purpose we test parent populations with incidence ratio IR_t and retrieve exact Poissonian confidence intervals from the quantiles of the χ^2 distribution for $2N \cdot IR_t$ degrees of freedom (low bound) and $2 \cdot (N \cdot IR_t + 1)$ degrees of freedom (high bound)[23]. As IR_t is an adjustable parameter, we search for the value of IR_t whose 95%CI contains the 95% of the reported IR. In this way, the null hypothesis $IR = IR_t$ sustains. In simple words, we will contextualize the difference between $IR = 11\,000/10\,000 = 1.10$ and $IR = 110/100 = 1.10$ within the Poisson statistics.

We do not include in our meta-analysis research studies on road traffic accidents where IR was not reported[24] (Finland) or N was unavailable[25] (Florida, USA). Likewise we do not consider research studies where the window of time was longer than one week[26] (United Kingdom) since they do not account for acute effects. Likewise we do not exclude any study on AMIs, ischemic strokes and accidents that reported N and IR to the best of our knowledge. We understand that the sample of thirteen research studies here analyzed which produced 64 day IRs, 10 week IRs, and 85 stratified IRs is representative of the literature on the field.

RESULTS

Fig. 1 shows the reported IRs versus the sample size of the study for the spring transition (left) and the autumn transition (right). Generally speaking day IRs (intermediate ink) have smaller sample sizes than week IRs (darkest ink) and show a much larger variability.

Fig. 1 left shows in broken lines the upper and lower bound of the 95%CI for $IR_t^{\text{spring}} = 1.00$, Table 2 summarizes the occurrences of reported IRs inside the 95%CI. In the first row of the table we see 59 (stratum, 75.6%), 47 (day, 73.4%) and 5 (week, 55.6%) reported IRs inside

Study	Region	Spring			Autumn		
		Years	N	IR [95 %CI]	Years	N	IR [95 %CI]
<i>Acute myocardial infarction</i>							
Ref. [10]	Sweden	15	10 251.0	1.05 [1.03, 1.07]	20	13 492.0	0.99 [0.97, 1.00]
Ref. [11]	Sweden	10	3115.5	1.04 [1.00, 1.08]	13	4095.5	0.99 [0.96, 1.03]
Ref. [12]	Two hospitals (Michigan, USA)	6	145.4	1.17 [1.00, 1.36]	6	158.3	0.99 [0.85, 1.16]
Ref. [13]	One hospital (Croatia)	6	45.9	1.15 [0.86, 1.51]	6	45.9	1.20 [0.90, 1.56]
Ref. [14]	Michigan (USA)	4	849.0	1.02 [0.95, 1.09]	4	658.0	0.98 [0.90, 1.06]
Ref. [15]	Finland	7	1257.2	1.01 [0.96, 1.07]	9	1633.2	0.99 [0.94, 1.04]
Ref. [16]‡	Ausburg (Germany)	26	491.9	1.08 [0.97, 1.25]	25	473.0	1.02 [0.93, 1.13]
<i>Ischemic strokes</i>							
Ref. [17]‡‡	Finland	1	210.7	1.14 [1.00, 1.30]	1	210.7	1.11 [0.98, 1.25]
<i>Accidents</i>							
Ref. [18]	USA	1	2750.0	1.07 [1.03, 1.10]	1	2938.5	0.98 [0.95, 1.02]
Ref. [19]†	Canada	1	2594.0	1.08 [1.04, 1.12]	1	4111.0	0.92 [0.89, 0.95]
Ref. [20]†	New Zealand	11	910.0	1.16 [1.02, 1.28]	12	1346.0	1.07 [0.94, 1.16]
Ref. [21]‡	USA	22	14 044.8	1.06 [1.03, 1.09]	22	14 044.8	1.00 [0.98, 1.03]
<i>Trauma admissions</i>							
Ref. [22]¶	Austria, Germany, and Switzerland	15	3456.0	1.08 [1.05, 1.12]	15	3755.0	0.94 [0.91, 0.97]

Table 1. The sample size N —the total number of events in the control week, the expected total number of events from a model or the average number of events in a week— and the incidence ratio IR associated with the week after the spring transition and the autumn transition. Values were obtained for the week after a transition, except when marked with †, in which case the largest IR reported in the first week after a transition is shown. Studies marked with ‡ did not provide N , we deduced them from an average of the record. Studies marked with ¶ did not report IR and their 95 %CI. Blue ink annotates p -values below the standard level of significance $\alpha = 0.05$. According to their N , the reported values can be located inside the 95 %CI for $IR_t = 1.05$ (spring), $IR_t = 1$ (autumn), see Figure 1 and Table 2.

the 95 %CI for $IR_t^{\text{spring}} = 1.00$. Therefore, the null hypothesis $IR_t^{\text{spring}} = 1.00$ does not sustain at the standard level of significance (5 %).

In contrast the null hypothesis $IR_t^{\text{spring}} = 1.05$ —whose 95 %CI is noted by a light shade background color— yields the largest occurrences inside the CI: 72 (92.3 %), 60 (93.8 %) and 9 (100.0 %) of the stratum, day and week observations (see second row in Table 2). Therefore the null sustains at the standard level of significance. A similar analysis for the autumn transition shows that the null hypothesis $IR_t^{\text{autumn}} = 1.00$ and $IR_t^{\text{autumn}} = 0.99$ sustains with similar scores.

As a result we understand that for the spring transition the incidence ratio is close to 1.05 or 5 % increase of the risks, while the autumn transition poses negligible small risks. Therefore, reported IRs above $IR_t^{\text{spring}} = 1.05$ —that often raises the alarm in decision makers, the public opinion, and researchers— are likely impacted by the limitation that the sample size always brings.

There is only one clear outlier in the catalogue of reported IR: motorcycle accident admissions hits $IR = 1.52$ in the trauma study[22]. The authors explain that motorcycle travels soars in Austria, Germany and Switzerland from March to April as many motorcyclists have summer license plates, which only operates from March to

Results inside the 95 %CI for IR_t

IR_t	Stratified	Day	Week
<i>spring</i>			
1.00	64/85 (75.3 %)	47/64 (73.4 %)	5/10 (50.0 %)
1.05	78/85 (91.8 %)	60/64 (93.8 %)	10/10 (100.0 %)
<i>autumn</i>			
1.00	70/85 (82.4 %)	56/64 (87.5 %)	9/10 (90.0 %)
0.99	71/85 (83.5 %)	54/64 (84.4 %)	9/10 (90.0 %)

Table 2. Occurrences of reported IRs inside the 95 % confidence interval of a Poissonian parent distribution for selected IR_t (see Fig. 1 for a graphical representation). First the number of occurrences relative to the total number of observations; then, in parentheses, the percentage score. The null hypothesis $IR = IR_t$ sustains at the standard level of significance for $IR_t^{\text{spring}} = 1.05$. For the autumn transition both $IR_t^{\text{autumn}} = 1.00$ and $IR_t^{\text{autumn}} = 0.99$ sustain similar scores.

October.

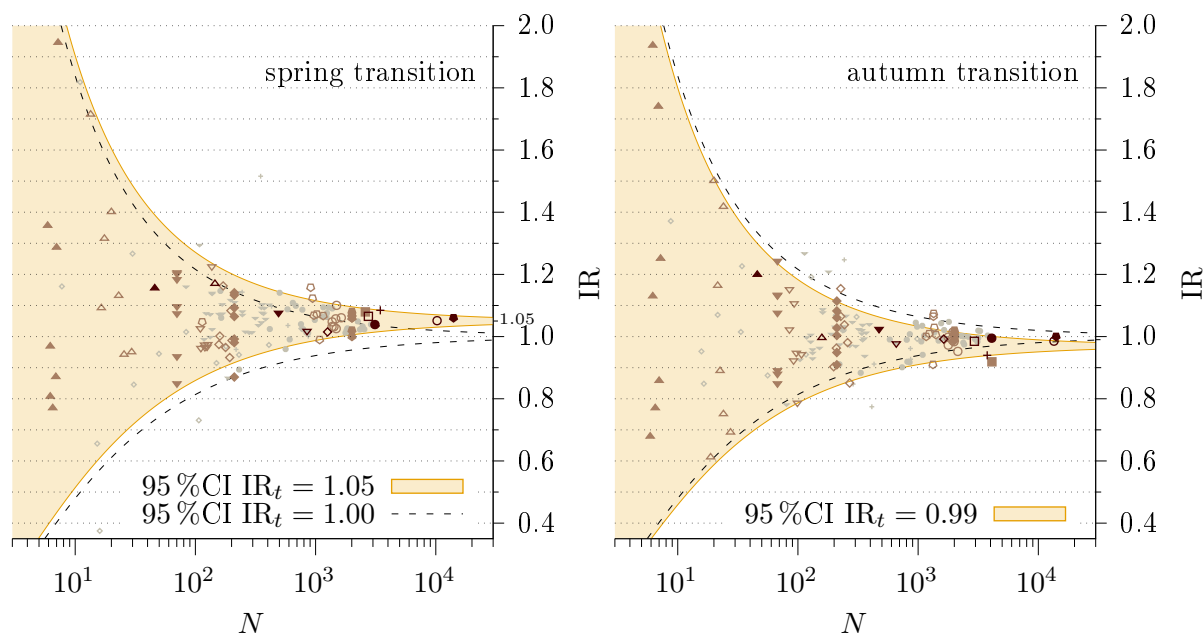


Figure 1. Scatter plot for stratified (lightest ink), day (intermediate ink) and week (darkest ink) IR associated with the spring transition (left panel) and the autumn transition (right panel). The lower and upper bounds of the 95 % confidence interval for a Poissonian parent distribution with $IR_t = 1.00$ are shown by broken lines. The 95 % confidence interval for a Poissonian parent distribution with $IR_t = 1.05$ (left) and $IR_t = 0.99$ (right) is shown in light shade background color. They roughly contains 95 % of the observations in either panel, see Table 2. Legend: open circles[10], solid circles[11], open up triangles[12], solid up triangles[13], open down triangles[14], open diamonds[15], solid down triangles[16], solid diamonds[17], open squares[18], solid squares[19], open pentagons[20], solid pentagons[21], crosses[22]. See Table 2 for a breakdown of occurrences inside the 95 % confidence intervals.

DISCUSSION

To understand the size of the spring effect we were able to estimate the relative sample standard deviation of the record in four studies[12, 14, 20, 21]. Figure 2 in [14] shows the residuals of the model on a daily basis. Visual perception suggests that the standard error of the model is at least some 20 % of the daily AMIs, the largest reported IR is in line with this number. Table 2b in [20] provides the standard error of the model as 0.069 in the \log_{10} space or $10^{0.069} = 1.17$ in the linear space: their largest reported IR is in line with this number. On the other hand, Figure 1 in [21] shows the sample standard deviation of the weekly observation in two periods of eleven years each. From this we infer a relative sample standard deviation $\sim 8\%$ in the week previous to the spring transition. Their reported IR is below this threshold. Finally, Table 1 in [12] contains N_i and O_i in the seven years of the record, from which we found a relative standard deviation of the sample $\sim 40\%$ for one week. In contrast [12] reported a week $IR = 1.17$ or 17 % increase in the week after the spring transition.

From these four studies — $N \in (150, 14000)$ — we understand that, while the spring transition shows a systematic increase of the risks, the size of the excursion is

below the sample relative standard deviation of O every year. Therefore the impact of the transitions is below the myriad of things that populates the variability of the tested quantity, which explains why the assessment of the impact of the clock transitions on health issues is elusive.

CONCLUDING REMARKS

We do not conclude from our analysis that DST transitions do not impact on public health. Instead, we bring attention to the fact that its impact might be as mild as previously thought. Decision makers and researchers should understand that the assessment of the IR is only the starting point of the balance of the risks associated with the practice. The risks of the practice should be balanced out against the risks of canceling a practice.

We bring the following recent example, US Senator Marco Rubio sponsored the Sunshine Protection Act of 2021 to “lock the clock”. Among other points, Rubio alleged that DST transitions caused 28 fatal motor vehicle accidents, a highlight from [21]. Much to the dismay of authors’ study and many others[6], Rubio vowed for locking the clock in the DST setting—that is, making DST the new perennial Standard Time— even though

the study also showed that the risks of fatal accidents doubled after DST spring transition was advanced three weeks in the US after the Energy Policy Act[27]. Note that the Sunshine Protection Act of 2021 would effectively advance DST “spring transition” by ten weeks and delay DST “autumn transition” by eight weeks to cancel out both.

The thing to emphasize is that having locked the clock in the permanent Standard Time during the 20th century would have caused also a greater number of fatal accidents. In other words, the advance of clocks in spring—and, therefore, the delay of sunrise times—enforced by DST also helped to prevent risks. Had sunrise time occurred in the 40° latitude at 4.30am in summer, greater shares of population would have found comfortable to advance their daily activity relative to present day scores.

This choice would have translated into early winter activity, which is more prone to poor light conditions and, eventually, to greater accidental risks. Current research studies can assess the impact of the practicing transitions, as they set a natural experiment to analyze. On the contrary, we can only speculate with the impact of not having practiced DST transitions by analyzing the impact of past actions in the present day.[28]

ACKNOWLEDGMENTS

The authors thank to Prof. Mercedes Conde Amboage from Universidade de Santiago de Compostela for fruitful discussion.

The authors declare no competing interest.

This project was not funded.

-
- [1] V. Amrhein, S. Greenland, and B. McShane, Scientists rise up against statistical significance, *Nature* **567**, 305 (2019).
 - [2] Y. Harrison, The impact of daylight saving time on sleep and related behaviours, *Sleep Medicine Reviews* **17**, 285 (2013).
 - [3] M. M. e Cruz, M. Miyazawa, R. Manfredini, D. Cardinali, J. A. Madrid, R. Reiter, J. F. Araujo, R. Agostinho, and D. Acuña-Castroviejo, Impact of daylight saving time on circadian timing system: An expert statement, *European Journal of Internal Medicine* **60**, 1 (2019).
 - [4] E. Parliament, D.-G. for Parliamentary Research Services, and I. Anglmayer, *EU summer-time arrangements under Directive 2000/84/EC Study* (European Parliament, 2017).
 - [5] R. Manfredini, F. Fabbian, A. D. Giorgi, B. Zucchi, R. Cappadona, F. Signani, N. Katsiki, and D. P. Mikhailidis, Daylight saving time and myocardial infarction: should we be worried? a review of the evidence., *European Review for Medical and Pharmacological Sciences* **22**, 750 (2018).
 - [6] T. Roenneberg, A. Wirz-Justice, D. J. Skene, S. Ancoli-Israel, K. P. Wright, D.-J. Dijk, P. Zee, M. R. Gorman, E. C. Winnebeck, and E. B. Klerman, Why should we abolish daylight saving time?, *Journal of Biological Rhythms* **34**, 227 (2019).
 - [7] N. F. Watson, Time to show leadership on the daylight saving time debate., *Journal of clinical sleep medicine : JCSM : official publication of the American Academy of Sleep Medicine* **15**, 815 (2019).
 - [8] C. Caviezel, C. Revermann, and S. Rabaa, Bilanz der sommerzeit : Endbericht zum ta-projekt, *Arbeitsbericht Nr 165*, 10.5445/IR/1000101425 (2016).
 - [9] J. Kearney, S. Chirico, and A. Jarvis, The application of summertime in europe (2014).
 - [10] I. Janszky and R. Ljung, Shifts to and from daylight saving time and incidence of myocardial infarction, *New England Journal of Medicine* **359**, 1966 (2009).
 - [11] I. Janszky, S. Ahnve, R. Ljung, K. J. Mukamal, S. Gautam, L. Wallentin, and U. Stenstrand, Daylight saving time shifts and incidence of acute myocardial infarction - swedish register of information and knowledge about swedish heart intensive care admissions (riks-hia), *Sleep Medicine* **13**, 237 (2012).
 - [12] M. R. Jiddou, M. Pica, J. Boura, L. Qu, and B. A. Franklin, Incidence of myocardial infarction with shifts to and from daylight savings time, *The American Journal of Cardiology* **111**, 631 (2013).
 - [13] V. Čulić, Daylight saving time transitions and acute myocardial infarction, *Chronobiology International* **30**, 662 (2013).
 - [14] A. Sandhu, M. Seth, and H. S. Gurm, Daylight savings time and myocardial infarction, *Open Heart* **1**, e000019 (2014).
 - [15] J. O. Sipilä, P. Rautava, and V. Kytö, Association of daylight saving time transitions with incidence and in-hospital mortality of myocardial infarction in finland, *Annals of Medicine* **48**, 10 (2015).
 - [16] I. Kirchberger, K. Wolf, M. Heier, B. Kuch, W. V. Scheidt, A. Peters, and C. Meisinger, Are daylight saving time transitions associated with changes in myocardial infarction incidence? results from the german monica/kora myocardial infarction registry, *BMC Public Health* **15**, 1 (2015).
 - [17] J. O. Sipilä, J. O. Ruuskanen, P. Rautava, and V. Kytö, Changes in ischemic stroke occurrence following daylight saving time transitions, *Sleep Medicine* **27-28**, 20 (2016).
 - [18] S. Coren, Accidental death and the shift to daylight savings time, *Perceptual and Motor Skills* **85**, 921 (1996).
 - [19] S. Coren, Daylight savings time and traffic accidents, *New England Journal of Medicine* **334**, 924 (1996).
 - [20] D. Robb and T. Barnes, Accident rates and the impact of daylight saving time transitions, *Accident Analysis and Prevention* **111**, 193 (2018).
 - [21] J. Fritz, T. VoPham, K. P. Wright, and C. Vetter, A chronobiological evaluation of the acute effects of daylight saving time on traffic accident risk, *Current Biology* **30**, 729 (2020).
 - [22] A. Nohl, C. Seelmann, R. Roenick, T. Ohmann, R. Lefering, B. Brune, V. Weichert, and M. Dudda, Impact of dst (daylight saving time) on major trauma: A european cohort study, *International Journal of Environmental Re-*

- search and Public Health 2021, Vol. 18, Page 13322 **18**, 13322 (2021).
- [23] J. Sun, Y. Ono, and Y. Takeuchi, A simple method for calculating the exact confidence interval of the standardized mortality ratio with an sas function, *Journal of Occupational Health* **38**, 196 (1996).
- [24] T. Lahti, E. Nysten, J. Haukka, P. Sulander, and T. Partonen, Daylight saving time transitions and road traffic accidents, *Journal of Environmental and Public Health* **2010**, 10.1155/2010/657167 (2010).
- [25] J. E. Molina, A. Kitali, and P. Alluri, Relationship between daylight saving time and traffic crashes in florida;, *Transportation Research Record* , 036119812211083 (2022).
- [26] R. Singh, R. Sood, and D. J. Graham, Road traffic casualties in great britain at daylight savings time transitions: a causal regression discontinuity design analysis, *BMJ Open* **12**, e054678 (2022).
- [27] J. M. Martín-Olalla, Traffic accident increase attributed to daylight saving time doubled after energy policy act, *Current Biology* **30**, R298 (2020).
- [28] J. M. Martín-Olalla, A chronobiological evaluation of the risks of canceling daylight saving time, *Chronobiology International* **39**, 1 (2022).