International Journal of Clinical Trials
Lu Y et al. Int J Clin Trials. 2023 Aug;10(3):195-200
http://www.ijclinicaltrials.com
pISSN 2349-3240 | eISSN 2349-3259
DOI: https://dx.doi.org/10.18203/2349-3259.ijct20232191

Original Research Article

How big is a big hazard ratio in clinical trials?

Yuanyuan Lu[1], Weiliang Cen[1], Wei Wang[2], Yangxin Huang[1], Henian Chen[1]*

[1]Study Design and Data Analysis Center, College of Public Health, University of South Florida, USA
[2]Centre for Addiction and Mental Health (CAMH), Toronto, Canada

Received: 12 March 2023
Revised: 23 March 2023
Accepted: 13 July 2023

***Correspondence:**
Dr. Henian Chen,
E-mail: hchen1@usf.edu

ABSTRACT

Background: The hazard ratio (HR) has been widely used as an index of effect size in clinical trials for time-to-event data. The use of the Cox proportional hazards models and other hazard centered models is ubiquitous in clinical trials for time-to-event data. The relativity of effect sizes (small, medium, large) has been widely discussed and accepted when comparing magnitude of association for continuous and categorical data, but not yet for time-to-event outcomes.
Methods: We review published hazard ratios, investigate the relationships among HR, relative risk (RR), odds ratio (OR), and Cohen's d, and calculate the corresponding HRs for given event rate in control group (P_0) by adding standard normal deviation with 0.2 (small), 0.5 (medium) and 0.8 (large) to the event rate in the case group (P_1) based on equation $P_1 = 1 - (1 - P_0)^{HR}$.
Results: Our results indicate that HRs are from 1.68 to 1.16 when the event rate of control group moves from 1% to 90%, which are equivalent to Cohen's d = 0.2 (small). HRs are ranged between 3.43 and 1.43 when the event rate of control group moves from 1% to 90%, which are equivalent to Cohen's d = 0.5 (medium), HRs are valued between 6.52 and 1.73 when the event rate of control group moves from 1% to 90%, which are equivalent to Cohen's d = 0.8 (large).
Conclusions: This study provides general guidelines in interpreting the magnitudes of HRs for time-to-event data in clinical trials.

Keywords: Hazard ratio, Clinical trial, Effect size, Time-to-event data

INTRODUCTION

The use of the Cox proportional hazards models and other hazard ratio (HR) centered models is ubiquitous in clinical trials for time-to-event data. Survival analysis can estimate the regression coefficients as constant effects over the follow-up period. HR has been widely used as an index of effect size in clinical trials. Reviewers for manuscripts and grant proposals are interested in knowing how big is a big treatment effect of a medical intervention.[1,2] As we know, a statistically significant finding indicates only that the sample size was large enough to detect a non-random effect since p value is related to sample size. An indication of statistical significance, however, does not provide information about how big the finding is.[3] It is quite possible with a large sample to have a statistically significant finding from a weak treatment effect. There is no widely accepted recognition for large treatment effects in time-to-event or survival data analysis. For continuous outcome measures, Cohen suggested that d = 0.2, 0.5, and 0.8 are small, medium, and large effect sizes, respectively.[4] For clinical research, most of the outcome measures are binary or time-to-event.

For binary outcome, according to the calculations from Chen et al, odds ratio (OR)=1.68, 3.47, and 6.71 are equivalent to Cohen's d=0.2 (small), 0.5 (medium), and 0.8 (large), respectively, when disease rate is 1% in the non-exposed group.[5] It is well known that under certain

circumstances (low population rates of "cases" <5%), OR provides a good approximation to the relative risk (RR). However, it is also recognized that OR does not give a good approximation of RR when disease rates do not fall below 5%. OR has little meaning in biomedical research unless it can approximate RR.[5,6] We know that base rate (P_0: prevalence from the non-exposed group or control group) can have an influence on resultant effect size estimates. For example, for a RR (RR = P_1 / P_0) of 2, if P_0 (from the control group) is 5%, then P_1 (from the exposed group) should be 10%, and the difference ($P_1 - P_0$) is 5%; then we may treat this finding as small. However, if P_0=25%, then P_1 should be 50%, and the difference is 25%; then we may treat this finding as medium or large.

There are relative effect sizes (small, medium, large) for binary outcome and continuous data, but not for time-to-event data. This paper investigates the relationships among HR, RR, OR, and Cohen's d, and provides general guidelines in interpreting the magnitudes of HRs for time-to-event data in clinical trials.

METHODS

Hazard ratio

Different statistical techniques have been developed to analyze survival data, including non-parametric, parametric and semi-parametric methods. Let T be a continuous measure of time to a predefined event (e.g., failure time) with density function$f(t)$, cumulative distribution function (cdf) $F(t) = P(T \le t)$, and survivor function $S(t) = P(T > t) = 1 - F(t)$. The hazard function for T is defined as the instantaneous risk per unit time and analytically it can be expressed as,

$$h(t) = \lim_{\Delta t \to 0} \frac{\Pr(t \le T < t + \Delta t \mid T \ge t)}{\Delta t} = \frac{f(t)}{S(t)} \quad (1)$$

The most commonly implemented statistical model for the hazard function is Cox's proportional hazards model, which defined hazard function as.[7]

$$h(t, z) = h_0(t)\exp(\boldsymbol{\beta}'\mathbf{z}) \quad (2)$$

where $h_0(t)$ is the unspecified baseline hazard function which could be any function of time as long as it is greater than 0. $\exp(\beta' z)$ is free of time, serving as the parametric part of the hazard function, where z is the covariate vector and β is the corresponding parameter vector. If we look a simplified example with only one predictor and let z = 0 designate no exposure and z=1 designate an exposed group. HR is the ratio of hazard functions in exposure group and non-exposure group. HR is widely used as a measure of association by comparing the hazard of certain event, e.g. death, between exposed group and non-exposed group at a given time. HR ranges from 0 to infinity, with 1 representing lack of association, greater than 1 indicating increased hazard in exposure group and smaller than 1 implying decreased hazard in exposure group. Cox proportional hazards model (Cox PH) assumes HR for any two individuals is constant over time and their survivorships are proportional. The HR for comparing exposure and non-exposure groups in Cox PH model is $exp(\beta)$, where β is the coefficient for group membership.

Hazard ratio, relative risk and odds ratio

Suppose $h_1(t)$ is the hazard function for exposure group, $h_0(t)$ represents the hazard function for non-exposure group, and $H_0(t)$ is the cumulative hazard function of non-exposure group. Under the proportional hazards assumption, the HR of exposure vs. the non-exposure is

$$\mathrm{HR} = \frac{h_1(t)}{h_0(t)} = e^{\beta}, t > 0 \quad (3)$$

The event rates for control group (P_0) and case group (P_1) at a given time t^* are

$$P_0(t^*) = 1 - S_0(t^*) = 1 - \exp\left\{-\int_0^{t^*} h_0(t)dt\right\} = 1 - \exp\{-H_0(t^*)\} \quad (4)$$

$$P_1(t^*) = 1 - S_1(t^*) = 1 - \exp\left\{-\int_0^{t^*} h_0(t)e^{\beta}dt\right\} = 1 - \exp\{-H_0(t^*)e^{\beta}\} = 1 - \exp\{-H_0(t^*) \times HR\} = 1 - (1 - P_0(t^*))^{HR} \quad (5)$$

The relative risk (RR) and odds ratio (OR) could be expressed by hazard ratio (HR) and event rate of control group (P_0) (time t^* is omitted for simplification of notation in the follow equations):

$$\mathrm{RR} = \frac{P_1}{P_0} = \frac{1 - (1 - P_0)^{HR}}{P_0} \quad (6)$$

$$\mathrm{OR} = \frac{\frac{P_1}{1-P_1}}{\frac{P_0}{1-P_0}} = \frac{\frac{1-(1-P_0)^{HR}}{(1-P_0)^{HR}}}{\frac{P_0}{1-P_0}} = \frac{[1 - (1 - P_0)^{HR}] \times (1 - P_0)}{P_0 \times (1 - P_0)^{HR}} \quad (7)$$

Hazard ratio and Cohen's d

For each event rate (P_0) in control group, we can have the relative standard normal deviation Z_0. By adding Z_0 with 0.2, 0.5 and 0.8, respectively, we get the standard normal deviation of P_1 for small, medium and large effects. Using equation (2), we can get HR

$$P_1 = 1 - (1 - P_0)^{HR} <=> \text{HR} = \frac{\log(1 - P_1)}{\log(1 - P_0)} \quad (8)$$

Where log(.) denotes natural logarithm.

RESULTS

The use of hazard ratio has increased rapidly over the previous two decades. We conducted a systematic literature review from PubMed database and found 76,713 publications that have used hazard ratio from January 1st, 1980 to December 31th, 2018. Figure 1 presents the number of publications that reported HR by year.

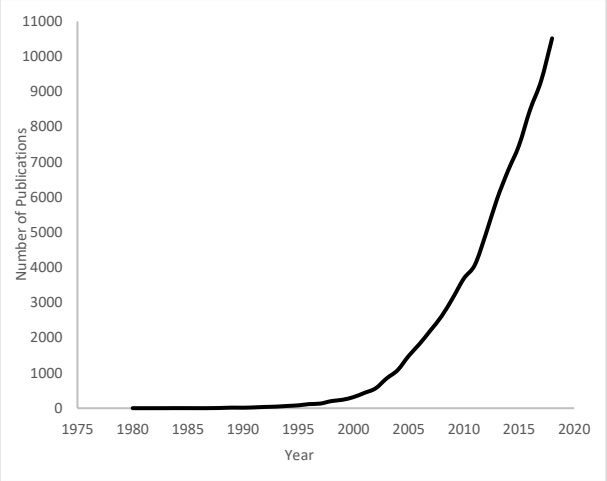

Figure 1: Number of Publications reporting hazard ratio (HR).

The number starts with 2 in year 1980 and increases up to 10,504 in year 2018, with dramatic increasing after year 2000.

Table 1: Magnitude of relative risk (RR) and odds ratio (OR) under various values of hazard ratio (HR) and event rate of control group (P_0)

		HR											
		1.00		1.50		2.00		3.00		4.00		5.00	
		RR	OR	RR	OR	RR	OR	RR	OR	RR	OR	RR	OR
P_0	0.01	1.00	1.00	1.50	1.50	1.99	2.01	2.97	3.03	3.94	4.06	4.90	5.10
	0.05	1.00	1.00	1.48	1.52	1.95	2.05	2.85	3.16	3.71	4.33	4.52	5.55
	0.10	1.00	1.00	1.46	1.54	1.90	2.11	2.71	3.35	3.44	4.72	4.10	6.24
	0.20	1.00	1.00	1.42	1.59	1.80	2.25	2.44	3.81	2.95	5.77	3.36	8.21
	0.30	1.00	1.00	1.38	1.65	1.70	2.43	2.19	4.47	2.53	7.38	2.77	11.55
	0.40	1.00	1.00	1.34	1.73	1.60	2.67	1.96	5.44	2.18	10.07	2.31	17.79
	0.50	1.00	1.00	1.29	1.83	1.50	3.00	1.75	7.00	1.88	15.00	1.94	31.00
	0.60	1.00	1.00	1.25	1.97	1.40	3.50	1.56	9.75	1.62	25.38	1.65	64.44
	0.70	1.00	1.00	1.19	2.18	1.30	4.33	1.39	15.44	1.42	52.48	1.43	175.94
	0.80	1.00	1.00	1.14	2.55	1.20	6.00	1.24	31.00	1.25	156.00	1.25	781.00
	0.90	1.00	1.00	1.08	3.40	1.10	11.00	1.11	111.00	1.11	1111.00	1.11	11111.00

Table 2: Cohen's d and the equivalent hazard ratio (HR).

	Cohen's d		
	0.2	0.5	0.8
P_0	HR	HR	HR
0.01	1.68	3.43	6.52
0.05	1.50	2.63	4.33
0.10	1.43	2.33	3.59
0.20	1.35	2.04	2.96
0.30	1.31	1.89	2.63
0.40	1.28	1.78	2.41
0.50	1.25	1.70	2.24
0.60	1.23	1.62	2.10
0.70	1.20	1.56	1.98
0.80	1.18	1.50	1.86
0.90	1.16	1.43	1.73

The numbers of survival trials using HRs as outcome measurement are 521, 671, 790, 863 and 702 for years 2014, 2015, 2016, 2017 and 2018, respectively. We reviewed publications (most are clinical trials) which reported significant HRs in the abstract from January 1st, 2017 to December 31th, 2018 (HR is converted to 1/HR, when HR<1) from New England Journal of Medicine (NEJM).

There are 226 significant HRs from NEJM. HRs reported in NEJM range from 1.02 to 33.33 with mean 2.52, and key percentiles as 1.39 (25%), 1.75 (50%), 2.40 (75%) and 4.39 (90%). Table 1 illustrates how the magnitudes of the RR and OR depend on both the value of the HR (ranged from 1 to 5) and the value of the event rate from the control group at baseline, P_0 (ranged from 0.01 to 0.90). A small P_0, such as 0.01, can be attributed to a short follow-up period or a small event rate at the end of the study. Results displays the RR and OR for increasing HR and increasing P_0.

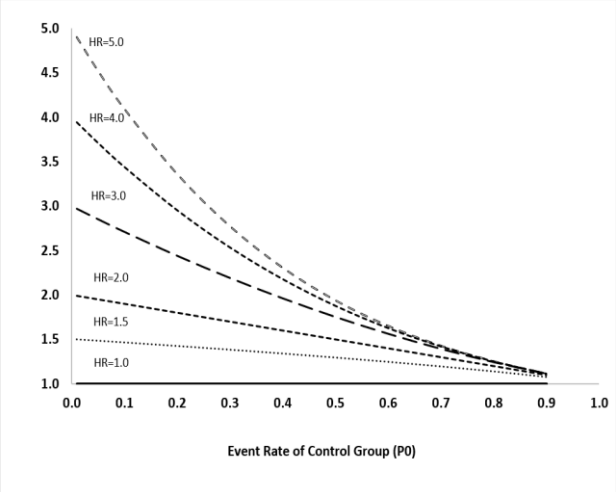

Figure 2: Magnitude of relative risk (RR) under various values of hazard ratio (HR) and event rate of control group (P_0).

There are three observations that should be noted from this table. First, for a constant HR, the value of the corresponding RR decreases and the OR increases as the event rate from the control group increases. Second, for a constant P_0, the HR is larger than the RR, but smaller than the OR. Lastly, the RR and OR are 1.0 when the HR is 1.0 regardless of the event rate from the control group (Table 1, Figures 2 & 3). The relationship between HR and RR is depicted in Figure 2. The curves show the decrease in RR as the base event rate P_0increases. Figure 3 depicts the relationship between HR and OR. The curves show the increase in OR as the event rate P_0 from the control group increases. (Table 2) displays the relationship between HR and Cohen's d as the event rate from the control group increases.

Cohen's d is said to be small, medium, and large, at 0.2, 0.5, and 0.8, respectively. HRs are from 1.68 to 1.16 when the event rate of control group moves from 1% to 90%, which are equivalent to Cohen's d=0.2 (small). HRs are ranged between 3.43 and 1.43 when the event rate of control group moves from 1% to 90%, which are equivalent to Cohen's d=0.5 (medium), HRs are valued between 6.52 and 1.73 when the event rate of control group moves from 1% to 90%, which are equivalent to Cohen's d=0.8 (large). Regardless of the magnitude of the effect size, HR decreases as P_0 increases. Figure 4 is the graphic presentation of (Table 2) and shows the relationship between HR and Cohen's d.

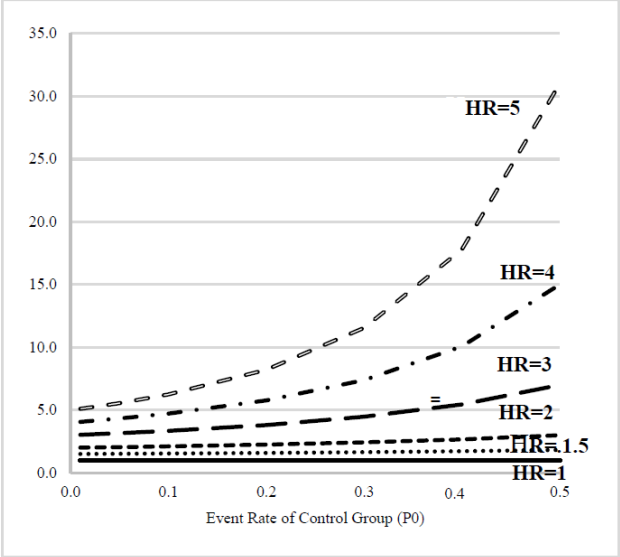

Figure 3: Magnitude of odds ratio (OR) under various values of hazard ratio (HR) and event rate of control group (P_0).

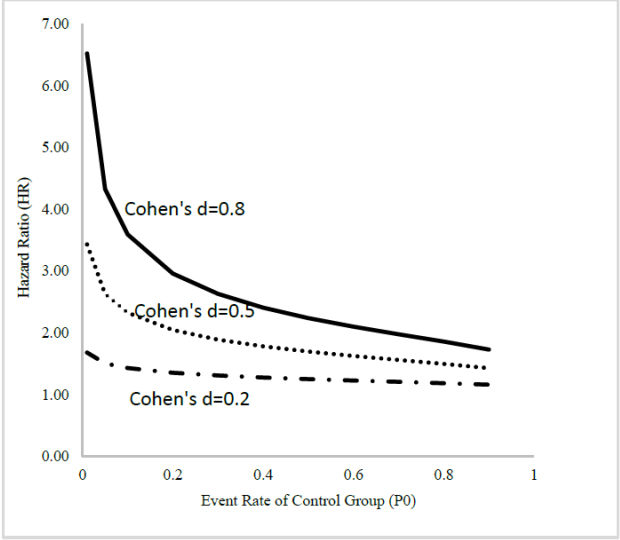

Figure 4: Cohen's d and equivalent hazard ratio (HR).

DISCUSSION

As is seen from (Table 1, Figure 2) HR exceeds RR, and is exceeded by OR. The divergence of them increases when increasing 1) the length of follow-up, 2) the average event rate at the end of the study, and 3) the magnitude of association (effect size). HR, RR, and OR numerically approximate one another when the follow-up period is short, the event rate is small, or the risk is close to 1. RR is a function of time, even when HR is independent of time. The HR should be described as a relative rate, not as a relative risk. The magnitude can be different between HR and RR.[8] Our calculations (Table 2) indicate that HRs are from 1.68 to 1.16 when the event rate of control group moves from 1% to 90%, which are equivalent to Cohen's d = 0.2 (small). HRs are ranged between 3.43 and 1.43 when the event rate of control group moves from 1% to 90%, which are equivalent to Cohen's d = 0.5 (medium), HRs are valued between 6.52 and 1.73 when the event rate of control group moves from 1% to 90%, which are equivalent to Cohen's d = 0.8 (large). Even though the hazard ratio might change over time, most studies only report a single HR at the end of the study. As a result, the conclusions from the study may critically depend on the duration of the follow-up [9]. It is possible to calculate an average HR if the HR is approximately constant. If the PH assumption is not violated, we can assume that the HR is constant over time. Otherwise, the partial likelihood estimator depends on underlying censoring time distribution when the PH assumption is violated. The average HR is under- or overestimated and we need to explore how the HR change over time and calculate the average HR under non-proportional hazards.[10-13]

Limitations

To date there is no consensus as to what those values of HR may be. Cohen brought 0.2, 0.5 and 0.8 as small, medium and large effects, respectively, for continuous outcome, and warned that they are only "rules of thumb".[4] HR does not directly translate into differences in times to events and therefore can present difficulties in interpretation. HR is a unitless value (ratio) since we divide the hazard rates and the time variable cancels out. The units of time could be days, months or years and the HR would be the same. We should report the magnitude of benefit in time and the absolute difference in survival when we report a HR.[14-16]

CONCLUSION

This study provides general guidelines in interpreting the magnitudes of HRs for time-to-event data in clinical trials. As the Cox regression model becomes more popular, HR is increasingly utilized in clinical trials. HR offers a convenient way to summarize how the event rate among an exposed group of subjects differs from the event rate among a control group of subjects. It would be useful for values with corresponding qualitative descriptors that estimate the magnitudes for time-to-event data in prospective follow-up studies. Better guidelines are needed when we use HR as the index of effect size for time-to-event data in clinical trials.

Funding: No funding sources
Conflict of interest: None declared
Ethical approval: The study was approved by the Institutional Ethics Committee

REFERENCES

1. Guyatt GH, Oxman AD, Sultan S, Glasziou P, Akl EA, Alonso-Coello P, et al. GRADE guidelines: 9. Rating up the quality of evidence. J Clin Epidemiol. 2011;64(12):1311-6.
2. Pereira TV, Horwitz RI, Ioannidis JP. Empirical evaluation of very large treatment effects of medical interventions. JAMA. 2012;308(16):1676-84.
3. Wasserstein RL, Schirm AL, Lazar NA. Moving to a world beyond "p < 0.05". Am Stat. 2021.
4. Cohen J. Statistical Power Analysis for the Behavioral Sciences. 2nd ed. Hillsdale, NJ: Erlbaum; 1988.
5. Chen H, Cohen P, Chen S. How big is a big odds ratio? Interpreting the magnitudes of odds ratios in epidemiological studies. Commu Stat Simul Comput 2010;39:860-4.
6. Altman DG, Deeks JJ, Sackett DL. Odds ratios should be avoided when events are common. BMJ. 1998;317(7168):1318.
7. Prentice RL, Zhao S. Regression Models and Multivariate Life Tables. J Am Stat Assoc. 2021; 116(535):1330-45.
8. Sashegyi A, Ferry D. On the Interpretation of the Hazard Ratio and Communication of Survival Benefit. Oncologist. 2017;22(4):484-6.
9. Hernán MA. The hazards of hazard ratios. Epidemiology. 2010;21(1):13-5.
10. Schemper M. Analysis of survival data with nonproportional hazard functions. Control Clin Trials. 1981;2(2):149-59.
11. Xu R, O'Quigley J. Estimating average regression effect under non-proportional hazards. Biostatistics. 2000;1(4):423-39.
12. Wei G, Schaubel DE. Estimating cumulative treatment effects in the presence of nonproportional hazards. Biometrics. 2008;64(3):724-32.
13. Schemper M, Wakounig S, Heinze G. The estimation of average hazard ratios by weighted Cox regression. Stat Med. 2009;28(19):2473-89.
14. Blagoev KB, Wilkerson J, Fojo T. Hazard ratios in cancer clinical trials--a primer. Nat Rev Clin Oncol. 2012;9(3):178-83.
15. Uno H, Claggett B, Tian L, Inoue E, Gallo P, Miyata T, et al. Moving beyond the hazard ratio in quantifying the between-group difference in survival analysis. J Clin Oncol. 2014;32(22):2380-5.

16. Weir IR, Marshall GD, Schneider JI. Interpretation of time-to-event outcomes in randomized trials: an online randomized experiment. Ann Oncol. 2019;30:96-102.

Cite this article as: Lu Y, Cen W, Wang W, Huang Y, Chen H. How big is a big hazard ratio in clinical trials?. Int J Clin Trials 2023;10(3):195-200.