Challenges and issues in bivariate endpoints in study designs of translational medicine

To the Editor:

We read with great interest the recent study by Chuah *et al.*[1] aiming to find immune predictors of treatment response and mechanisms of response versus immune-related adverse events (irAE) in hepatocellular carcinoma (HCC) patients treated with anti-programmed cell death 1 (PD-1). The authors noticed the recent successful combinations of immunotherapies achieved increased response rates compared to monotherapy, but with an accompanying increase in irAEs. The authors were inspired by this and conducted an innovative study to investigate how to increase response rates and reduce irAEs simultaneously. In this study, the peripheral blood samples of patients with HCC who had received anti-PD-1 monotherapy were collected for cytometry by time of flight and single-cell RNA sequencing to find the immune components related to response and toxicity, respectively. Further, on the basis of cell-cell interaction analysis, *in vivo* mouse models were established to verify that the immunotherapy targets can improve response rates without increasing toxicity. Defining optimal treatment combinations which balance activity and toxicity is an important issue in oncology clinical trials.[2,3] This study provides novel insights into the preclinical study of drug combinations. However, there are a number of concerns with the study design which should be pointed out.

In this study, when investigating the enriched immune components and cell-cell interactions for response and toxicity, the two binary endpoints of efficacy and safety were analyzed separately. The interaction or the bivariate nature of the two endpoints were ignored. Bivariate endpoints are not uncommon in clinical studies. For example, in early drug toxicity trials, there were statistical models that considered drug activity and toxicity at the same time.[4] But in basic and translational research, experiments are often based on a binary endpoint to explore a mechanism. The study follows the traditional design, showing the differences of targets between response and non-response and between toxicity and non-toxicity independently. This separate analysis, no matter from the perspective of bioinformatics or basic research, cannot accurately capture the combined active mechanism of response and toxicity. We believe that the study population should be reorganized into four groups in order to evaluate efficacy and safety simultaneously. As Fig. 1 shows, patients from each study cohort could be divided into four quadrants according to their two binary outcomes (*i.e.* response and toxicity) as a previous study on a bivariate endpoint trial design presented.[5] Patients from the upper right quadrant I achieved the most desirable result from anti-PD-1 treatment with response and non-toxicity, while the lower left quadrant reflected a negative result with no response but toxicity. In quadrant II and IV, patients faced equivocal benefit from the treatment. In terms of the purpose of this study, the authors could find suitable targets for combined immunotherapy by analyzing the difference between patients in quadrant I and other quadrants. Through this classification, we can not only compare the targets between response and no response and between toxicity and no toxicity groups, but also analyze the characteristics of patients in quadrant I, so as to find specific targets for both response and non-toxicity. Therefore, it is necessary to make such a distinction among the population first through this approach. However, one challenge of this classification is the limited sample size for each group. Although the number of each quadrant is relatively small, all four types of patients exist as expected.

In addition, only patients receiving monotherapy were included to evaluate the immune targets for combined therapies. To highlight the results of the combined anti-PD-1 and anti-TNFR2 treatment, a control group receiving combined immunotherapies which have been proven to be positive in randomized clinical trials should be added to the *in vivo* mouse models. Other issues include the limited sample size of each group for *in vivo* mouse models. In addition, the reporting of how to find the two targets of TNFR1 and TNFR2 from a cluster of ligand-receptor pairs is unclear.

Besides, we re-analyzed the characteristic data of the two cohorts from Table S1 of the paper (see our Table S1) and found that the age of the Korea cohort using for validation was 10 years younger than that of the Singapore cohort. Some other variables, such as BCLC stage, viral status, extrahepatic spread, were significantly different between the two cohorts as well.

Finally, we noticed that as shown in Fig. 5A, there were only two time points of measurements after treatment, reflecting the changes of cell proportions. The authors should provide the rationale for using the word 'trajectory' in the title and conclusion of the study.

In summary, the findings of the study need to be confirmed by further well-designed large-sample animal experiments before they can be applied to clinical translational studies. Preclinical design is very important for drug development and good study design is necessary to reduce potential bias in the process of mechanism research.

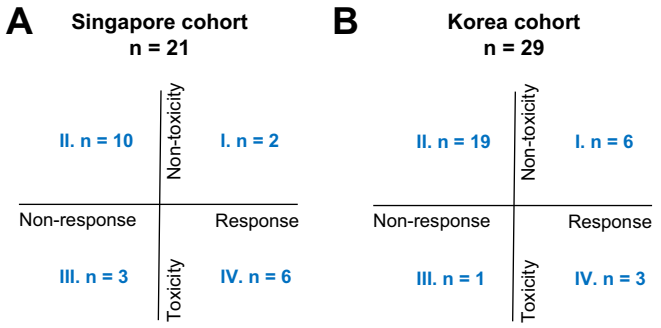

Fig. 1. Bivariate endpoints analysis maps. Four quadrants were presented based on the two binary endpoints of response and toxicity.

Received 27 April 2022; received in revised form 16 May 2022; accepted 31 May 2022; available online 13 June 2022
https://doi.org/10.1016/j.jhep.2022.05.035

Financial support
The authors received no financial support to produce this manuscript.

Conflict of interest
The author declares no conflicts of interest that pertain to this work.

Please refer to the accompanying ICMJE disclosure form for further details.

Supplementary data
Supplementary data to this article can be found online at https://doi.org/10.1016/j.jhep.2022.05.035

References
[1] Chuah S, Lee J, Song Y, Kim HD, Wasser M, Kaya NA, et al. Uncoupling immune trajectories of response and adverse events from anti-PD-1 immunotherapy in hepatocellular carcinoma. J Hepatol 2022 Sep;77(3):683–694.
[2] Ahern E, Solomon BJ, Hui R, Pavlakis N, O'Byrne K, Hughes BGM. Neo-adjuvant immunotherapy for non-small cell lung cancer: right drugs, right patient, right time? J ImmunoTherapy Cancer 2021;9(6):e002248.
[3] Agostinetto E, Losurdo A, Nader-Marta G, Santoro A, Punie K, Barroso R, et al. Progress and pitfalls in the use of immunotherapy for patients with triple negative breast cancer. Expert Opin Investig Drugs 2022;9:1–25.
[4] Colin P, Delattre M, Minini P, Micallef S. An escalation for bivariate binary endpoints controlling the risk of overtoxicity (EBE-CRO): managing efficacy and toxicity in early oncology clinical trials. J Biopharm Stat 2017;27(6):1054–1072.
[5] Kittelson JM, Spyropoulos AC, Halperin JL, Kessler CM, Schulman S, Steg G, et al. Balancing risk and benefit in venous thromboembolism trials: concept for a bivariate endpoint trial design and analytic approach. J Thromb Haemost 2013;11(8):1443–1448.

Qian Zhou*
Department of Medical Statistics, Clinical Trials Unit, The First Affiliated Hospital of Sun Yat-sen University, 58 Zhong Shan Er Road, Guangzhou, China, 510080
*Corresponding author. Address: Department of Medical Statistics, Clinical Trials Unit, The First Affiliated Hospital of Sun Yat-sen University, 58 Zhong Shan Er Road, Guangzhou, China, 510080. Tel./fax: +86-20-87755766-8649,.
E-mail address: zhouq49@mail.sysu.edu.cn

Reply to: "Challenges and issues in bivariate endpoints in study designs of translational medicine"

To the Editor:

We thank Zhou[1] for their interest and comments on our work.[2] Zhou commented that the two binary endpoints of efficacy and safety should be analyzed separately or reorganized into four groups. Indeed, the relationship between response and toxicity to immunotherapy is widely recognised, which we also acknowledged by citing the clinical trials of CheckMate040 and IMBRAVE150, where immune related adverse events (irAEs) increased in tandem with greater objective response rates (ORRs) in combination immunotherapy.[3,4] There were also other analyses that correlated irAEs with clinical benefits of immune checkpoint blockade (ICB), both in HCC and other cancer types.[5,6] With this in mind, we did explore the interaction between response and toxicity to treatment. Firstly, our multivariate analyses showed that the incidence of irAEs was a significant factor relating to progression-free survival, as would be expected (Table S8).[2] We also split our cohorts into the four different groups to verify that our biomarkers were not confounded by the interaction between response and toxicity (Figs S2C and S3B).[2] Due to the small sample size in some of the groupings, we have elected instead to conduct our main analysis on response and toxicity separately, while still acknowledging their potential co-dependence.

The author further commented on combination therapy, where we have only included patients receiving monotherapy and the limited sample size of each group in our *in vivo* mouse studies. We would like to highlight that the objective of our study was to analyse patients undergoing anti-PD1 monotherapy in order to discover potential combination therapies that would improve response rates while not increasing the incidence of irAEs. As such, we do not yet have an adequate patient cohort to act as a control for 'combined immunotherapies which have been proven to be positive in randomized clinical trials'. In fact, we are planning to conduct the first-in human trials using combination immunotherapy of anti-PD-1+anti-TNFR2, where we will be able to assess the patient data in the near future. Rather, in this current study, we validated the response and toxicity of the combination immunotherapy in our pre-clinical HCC mouse model, based upon our discovery in the HCC patient cohort treated with anti-PD-1 ICB. Hence, our study focused on discovery and pre-clinical validation, not clinical validation. The deeper mechanistic insights into the involvement of T cells and antigen-presentation cells in clinical response to ICB is another highlight of our study. Despite the small sample sizes, the strong data we observed provided sufficient proof-of-concept for an anti-PD1+anti-TNFR2 combination immunotherapy, which could encourage larger scale animal experiments and clinical trials in HCC.

The differences between the Singapore and Korea cohorts were already explained in the description of our patient cohorts. The Singapore patient cohort was recruited from real-world clinical patients with advanced HCC undergoing anti-PD1 immunotherapy. As a result, their clinical characteristics were very heterogenous. To ensure that the immune targets we discovered were not affected by this heterogeneity, we collaborated and validated our targets with the group from Korea, who had a more controlled cohort of patients with HCC from their clinical trial (NCT03695952). The analyses

Received 27 July 2022; accepted 14 August 2022; available online 24 August 2022
https://doi.org/10.1016/j.jhep.2022.08.017