

Concerns in the Propensity Score–Matched Analysis

Authors

Yoshihide Inayama, MD, PhD^{1,2}; Toshiki Fukasawa², PhD; and Koji Yamanoi, MD, PhD¹

¹ Department of Gynecology and Obstetrics, Kyoto University Graduate School of Medicine,
Kyoto, Japan

² Department of Pharmacoepidemiology, Graduate School of Medicine and Public Health,
Kyoto University, Kyoto, Japan

Address all correspondence to:

Koji Yamanoi, M.D., Ph.D.

Department of Gynecology and Obstetrics, Kyoto University Graduate School of Medicine,
54 Shogoin Kawahara cho, Sakyo ku, Kyoto 606-8507, Japan

Tel: 81-75-751-3269

Fax: 81-75-761-3967

E-mail: kojiymni@kuhp.kyoto-u.ac.jp

Manuscript word count:

396 words

Concerns in the Propensity Score–Matched Analysis

To the Editor

We read with interest the propensity score matching (PSM) analysis by Suzuki et al., which reported poorer overall survival with hormonal therapy versus hysterectomy for early-stage endometrial cancer.¹ They reported “no difference” in outcomes for patients <40, but a higher risk for those aged 40–49.¹ While noteworthy, these results raise several methodological concerns.

First, the authors performed age-subgroup analyses after PSM; however, matching ensures balance in the overall matched cohort, and not necessarily within subgroups. Although PSM assumes exchangeability within matched pairs, this assumption may not hold across age subgroups without exact or stratified age matching.² Table 2 shows unequal subgroup sizes (<40: 861 vs. 850; ≥40: 178 vs. 189), suggesting that subgroup exchangeability is not ensured. Hence, the survival difference in the 40–49 age group could stem from covariate imbalance.² The authors should provide subgroup balance metrics (e.g., standardized mean differences) and, if possible, repeat subgroup analyses using approaches that ensure within-subgroup balance.

Second, the target estimand is insufficiently discussed.³ The primary hysterectomy group shrank from 13,870 to 1,039 under 1:1 PSM, excluding >90% patients, whereas the hormone-therapy cohort was largely preserved. Consequently, the study seemed to estimate the effect among those who received hormone therapy (i.e., average treatment effect on the treated [ATT]). Using an alternative treatment as a reference may be clinically reasonable, as hormone therapy is selected by patients strongly preferring uterine preservation, but the estimand should be specified to avoid misunderstanding. Moreover, the substantial drop in hysterectomy cases likely reflects 1:1 matching, but overlap (common support) between groups is not documented. The authors should show prematching and postmatching

propensity score histograms and specify the common support region.⁴ If the goal is ATT, standardized mortality ratio weighting (SMRW) can exploit more observations, improve efficiency, and clearly define the target population.³ Thus, discarding many hysterectomy cases to perform 1:1 matching offers limited methodological benefit. Additional SMRW-based analyses could be useful.

Third, the authors conclude “no difference” in outcomes for patients <40 years based on a hazard ratio 1.00 (95% confidence interval [CI], 0.50–2.00). Such a claim is justified only when the entire CI lies within a clinically unimportant range.⁵ This also highlights room for improvement in a method that excludes many hysterectomy patients, thereby reducing precision.

Evaluating survival after primary hormone therapy is clinically important; however, clear estimand specification and higher-precision analyses are needed to avoid misinterpretation.

Yoshihide Inayama, MD, PhD; Toshiki Fukasawa, PhD; and Koji Yamanoi, MD, PhD

Conflict of Interest Disclosures

Toshiki Fukasawa was previously employed by the Department of Digital Health and Epidemiology at Kyoto University with support from Eisai Co., Ltd. and Kyowa Kirin Co., Ltd.; he has received research funding from AstraZeneca K.K. and the Pfizer Health Research Foundation, consulting fees from Asahi Kasei Pharma Corporation, JMDC Inc., MSD K.K., and Real World Data Co., Ltd., and honoraria from EPS Corporation, DeSC Healthcare, Inc., and the Research Institute of Healthcare Data Science; and his spouse is employed by MSD K.K. The other authors declare that there are no conflicts of interest.

References

1. Suzuki Y, Huang Y, Xu X, et al. Survival After Fertility-Preserving Hormonal Therapy vs Hysterectomy for Early-Stage Endometrial Cancer. *JAMA Oncol.* Aug 28 2025;doi:10.1001/jamaoncol.2025.2761
2. Wang SV, He M, Jin Y, et al. A review of the performance of different methods for propensity score matched subgroup analyses and a summary of their application in peer-reviewed research studies. *Pharmacoepidemiol Drug Saf.* Dec 2017;26(12):1507-1512. doi:10.1002/pds.4328
3. Desai RJ, Franklin JM. Alternative approaches for confounding adjustment in observational studies using weighting based on the propensity score: a primer for practitioners. *BMJ.* Oct 23 2019;367:l5657. doi:10.1136/bmj.l5657
4. Zhu Y, Hubbard RA, Chubak J, Roy J, Mitra N. Core concepts in pharmacoepidemiology: Violations of the positivity assumption in the causal analysis of observational data: Consequences and statistical approaches. *Pharmacoepidemiol Drug Saf.* Nov 2021;30(11):1471-1485. doi:10.1002/pds.5338
5. Mansournia MA, Nazemipour M. Recommendations for accurate reporting in medical research statistics. *Lancet.* Feb 17 2024;403(10427):611-612. doi:10.1016/S0140-6736(24)00139-9