Received: 12 August 2021 | Revised: 25 September 2021 | Accepted: 21 October 2021

DOI: 10.1111/cts.13187

ARTICLE

Predicting completion of clinical trials in pregnant women: Cox proportional hazard and neural network models

Bomee Kim[1] | **Yun Ji Jang**[2] | **Hae Ram Cho**[2] | **So Yeon Kim**[2] | **Ji Eun Jeong**[2] | **Mi Kyoung Shim**[2] | **Myeong Gyu Kim**[3,4]

[1]Graduate School of Clinical Biohealth, Ewha Womans University, Seoul, Korea

[2]College of Pharmacy, CHA University, Pocheon, Korea

[3]College of Pharmacy, Ewha Womans University, Seoul, Korea

[4]Graduate School of Pharmaceutical Sciences, Ewha Womans University, Seoul, Korea

Correspondence
Myeong Gyu Kim, College of Pharmacy and Graduate School of Pharmaceutical Sciences, Ewha Womans University, 52 Ewhayeodae-gil, Seodaemun-gu, Seoul 03760, Korea.
Email: kimmg@ewha.ac.kr

Funding information
This work has supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT; No. NRF-2021R1C1C1013177)

Abstract

This study aimed to develop a model for predicting the completion of clinical trials involving pregnant women using the Cox proportional hazard model and neural network model (DeepSurv) and to compare the predictive performance of both methods. We collected data on 819 clinical trials performed on pregnant women and intervention studies using at least one drug as intervention from 2009 to 2018 from ClinicalTrials.gov. The Cox proportional hazard model and DeepSurv were used to develop models that predict clinical trial completion. The concordance index (C-index) was used to evaluate the predictive performance. The Cox proportional hazard model revealed that a sample size of $n \geq 329$ (hazard ratio [HR] = 0.53), very high human development index (HDI) country (HR = 0.28), abortion (HR = 3.30), labor (HR = 2.16), and iron deficiency anemia (HR = 2.29) were significantly related to the probability of clinical trial completion (all p value < 0.01). The C-index of the model development dataset and test dataset were 0.72 and 0.73, respectively. DeepSurv model consisted of one hidden layer with 16 nodes. DeepSurv showed the C-index comparable to the Cox proportional hazard model. The C-index of the training dataset and test dataset were 0.76 and 0.72, respectively. Further a nomogram that calculate a probability of clinical trial completion at 1 year, 3 years, and 5 years was developed. Both the Cox proportional hazard model and DeepSurv yielded sufficient predicting performance. We hope that this study will contribute to the execution of future clinical trials in pregnant women.

Study Highlights

WHAT IS THE CURRENT KNOWLEDGE ON THE TOPIC?

Machine learning algorithms have been applied to predict completion or termination of clinical trials. However, previous studies have not considered the time after clinical trials begin, an important factor of completion or termination.

WHAT QUESTION DID THIS STUDY ADDRESS?

Based on the characteristics of clinical trials in the stage of planning clinical trials, can we predict when clinical trials for pregnant women will be completed?

Can Cox proportional hazard and neural network models predict completion of clinical trials successfully?

WHAT DOES THIS STUDY ADD TO OUR KNOWLEDGE?

Sample size, study country, and target medical conditions (abortion, labor, and iron deficiency anemia) were significant predictor of completion of clinical trials on pregnant women. The Cox proportional hazard model and neural network model (DeepSurv) showed good performance (concordance index [C-index] >0.7) in predicting completion of clinical trial on pregnant women.

HOW MIGHT THIS CHANGE CLINICAL PHARMACOLOGY OR TRANSLATIONAL SCIENCE?

Using the developed model and nomogram, researchers can calculate the probability of completion of clinical trials before clinical trials for pregnant women are conducted. This allows researchers to anticipate the completion of clinical trials and help plan clinical trials.

INTRODUCTION

Pregnant women may require treatment for chronic diseases or acute conditions. Special medical conditions, such as preterm labor, pre-eclampsia, and gestational diabetes, also require treatment. Not treating pregnant women may be more dangerous than drug side effects.[1] More than 80% of pregnant women have used at least one drug during their pregnancy.[2] However, treatment for pregnant women is often difficult due to a lack of information about efficacy and safety based on clinical trials.

Clinical trials involving pregnant women have been limited due to ethical issues surrounding potential adverse effects (teratogenicity or genotoxicity).[3,4] Many pregnant women are reluctant to participate in clinical trials for this reason. A study reported that 95% of industry-sponsored clinical studies including female subjects excluded pregnant women.[5] Another study analyzed the difficulty of recruiting patients in an obstetric trial.[6] Only 22% participated in the study, with the rest excluded due to exclusion criteria (47%), patient refusal (21%), and obstetrician refusal (10%). For these reasons, the rationale for medication usage for pregnant women generally came from observational, retrospective, or epidemiological studies.[7]

There is increasing global agreement that pregnant women should be included in clinical studies to collect evidence about treatment options during pregnancy.[8,9] Pregnant women sometimes must use drugs without scientific evidence of the potential dangers to themselves and the fetus.[1] Clinical research can help to establish safe and effective treatment options and dosing regimens for pregnant individuals.[10] The 2002 Council for International Organizations of Medical Sciences (CIOMS) guidelines state that pregnant women should be presumed to be eligible for participation in biomedical research.[11,12] There have been many efforts to include pregnant women in clinical trials over the last 2 decades. In 2018, the US Food and Drug Administration (FDA) announced draft guidelines for scientific and ethical considerations for including pregnant women in clinical trials.[10]

Nevertheless, studies on pregnant women are still difficult to complete. Researchers and sponsors must predict whether a clinical trial will be completed during planning. Recently, two predictive models using machine learning algorithms have been developed to determine if clinical trials will be completed or terminated. Follet et al. used a random forest algorithm and found features associated with clinical trial termination (enrollment group, study phase, intervention assignment, primary purpose, and the appearance of some keywords ["'treat'," "'chemotherapy'," "'cancer'," "'patients'," and "'tumor"']).[13] However, the model's predictive performance was not excellent (sensitivity = 0.56; specificity = 0.71; accuracy = 0.71; precision = 0.07; and F1 score = 0.12).[13] Elkin et al. trained four types of classifiers (neural networks, random forest, XGBoost, and logistic regression) and found features related to clinical trial termination (i.e., eligibility words, study phase, industry sponsor, and cancer-related words).[14] The model's predictive performance of the model was satisfactory (balanced accuracy = 0.67 and area under the curve = 0.73).[14]

However, previous studies have not considered the time after clinical trials begin, an important factor in trial completion or termination. The time-to-event analysis allows researchers to predict the probability of clinical trial completion at a specific time after the clinical trial begins. The Cox proportional hazard model is a traditional method of time-to-event analysis, which allows the development of survival functions using multiple predictors.[15] However, this model is not suitable for nonlinear survival data, as it assumes linear proportional hazards.[16] Recently, DeepSurv, which incorporates neural networks

into time-to-event analysis, has improved nonlinear survival data over the Cox proportional hazard model.[16]

Both methods can be applied to predict the probability of clinical trial completion over time. This study aimed to develop a model for predicting the completion of clinical trials involving pregnant women using the Cox proportional hazard model and DeepSurv and compared the predictive performance of both methods.

METHODS

Data source

Clinical trial data were collected from ClinicalTrials.gov, a publicly available registry of clinical studies.[17] A trial record manager provides trial registration before enrolling the first subject and administers each trial record in the database. ClinicalTrials.gov is the largest clinical trial database and has been used in previous studies.[13,14] Search terms used in the "condition or disease" field of the website were "pregnant," "pregnancy," "maternal," "prenatal," and "gestational." Clinical studies that were initiated between January 1, 2009, and December 31, 2018, were collected. The search was conducted on May 10, 2021.

We used the following selection criteria for analysis: studies performed on pregnant women (during pregnancy to child-birth) and intervention studies using at least one drug as intervention. Observational studies or intervention studies on surgical or medical devices were excluded. The selection process was conducted by two researchers and any discrepancy was solved through discussion. Because this study involved analysis of pre-existing, non-human data, it was exempt from institutional review board approval.

Data preprocessing and feature engineering

In this time-to-event analysis, an event was defined as "completed" in the recruitment status field on the ClinicalTrials.gov website, which means the study ended normally. Other recruitment statuses were considered as censored data. Time was defined as the period from the study start date to the study completion date or the date of the last update posted (whichever comes first).

Features included quartile value of sample size, pregnancy stages, number of study countries, human development index (HDI) of the study country, study phase, sponsor, randomization, intervention assignment, participant blinding, primary purpose, placebo group, target medical condition, and number of eligibilities.

Preplanned sample size, not actual enrollment, was categorized into quartiles ($0 \le n < 80$, $80 \le n < 150$, $150 \le n < 329$, and $n \ge 329$). Information on the pregnancy stages required by clinical trials was collected from the eligibility criteria field. Pregnancy stage was defined as first trimester (0 to 13 6/7 weeks of gestational age), second trimester (14 0/7 to 27 6/7 weeks of gestational age), and third trimester (beyond 28 0/7 weeks of gestational age).[18] A plus or minus 1 week difference was allowed. Some clinical trials involved more than one pregnancy stage (i.e., first and second trimester), so each stage was designated as a binary feature.

The number of study countries was classified as either single or multicountry. Study country was classified as very high, high, medium, or low HDI countries.[19] The study phase included phase I (also including early phase I), phase II, phase III, and phase IV. There were clinical trials involving more than one phase (i.e., phases I and II), so each phase was designated as a binary feature. The sponsor was categorized as government, industry, or nonprofit.[20,21]

Randomization (randomized or nonrandomized), intervention assignment (single group, parallel, cross-over, sequential, or factorial assignment), participant blinding (participant blinded or non-blinded), and primary purpose (treatment, prevention, and others) were collected from the study design field. In this study, we also chose whether the clinical trials included a placebo group as a feature.

Target medical condition meant the disease, disorder, syndrome, illness, or planned surgery, which is why investigational drugs are used in a clinical trial. Target medical condition consisted of binary features (1 or 0) for 14 diseases, including noninfectious diseases (hypertension/pre-eclampsia, diabetes, iron deficiency anemia, and other noninfectious diseases), infectious diseases (malaria, human immunodeficiency virus, viral hepatitis, and other infectious diseases), pregnancy-specific conditions (abortion [induced termination of pregnancy], miscarriage [spontaneous loss of a fetus before the 20th week of pregnancy], preterm birth, labor, and other pregnancy-specific conditions), and others. In the case of multiple target medical conditions, for example, HIV and viral hepatitis, each feature was coded as 1.

The number of eligibilities was defined as the number of total criteria in the eligibility criteria field because some clinical trials did not clearly distinguish between inclusion and exclusion criteria. The number of eligibilities is a binary feature divided by its median value ($n \le 10$ and $n > 10$).

The dataset was randomly divided into a dataset for model development (80%) and an unseen dataset for testing (20%). Clinical trial completion rates of two datasets

were compared using Kaplan-Meier curves and the log rank test. The difference in feature distributions between the two datasets was evaluated by a chi-square test and *t*-test.

Cox proportional hazard model

The univariable Cox proportional hazard models were estimated to investigate the statistical significance of the association between each feature and the completion of the clinical trials. Next, multivariable Cox proportional hazard models were estimated. A stepwise selection procedure, based on likelihood ratio tests for nested models, was used to select a set of significant features.[22] A strict cutoff for significance, an alpha of 0.01, was used. R version 4.0.2 (R Foundation for Statistical Computing, Vienna, Austria) and a survival package was used for the Cox proportional hazard model. R code for Cox proportional hazard model is presented in the supplementary documents (Methods S1).

Neural network model for survival analysis

To develop a model that predicts clinical trial completion using DeepSurv, the dataset for model development (80% of the overall dataset) was further divided into an 80% training dataset (Table S1) and a 20% validation dataset (Table S2).

We used Python version 3.7 (Python Software Foundation, Delaware, United States) and pycox, a Python package for time-to-event analysis. The hyperparameters that yielded the largest concordance index (C-index) were identified by a grid search. DeepSurv was then trained with an adaptive moment estimation (Adam) optimizer on three NVIDIA Quadro RTX 8000 Graphical Processing Units (GPUs). Python code for DeepSurv is presented in the supplementary documents (Methods S2). DeepSurv does not need prior feature selection, so we trained a model including all features. To check for unnecessary features, we removed each feature from the final model and identified the change in C-index. Because the C-index can be unstable if the dataset is not large, we averaged the C-index after five final models were constructed per dataset.

Comparison of performance

The C-index was used to evaluate the performance of the Cox proportional hazard model and DeepSurv. The C-index indicates the proportion of samples that are correctly ranked when the samples are listed in the order of predicted survival time.[23] A value of 0.5 indicates that the model is no better at predicting an outcome than random chance, and a value of 1 means that the model perfectly predicts an outcome.

RESULTS

Dataset

A total of 6020 clinical trials were obtained from the database, 3602 of which included pregnant women. There were 2308 (64.1%) intervention studies, and 819 clinical trials (35.5% of intervention studies) which used at least one drug as an intervention were selected for the analysis.

Among the 819 clinical trials, 52.9% (n = 433) were completed during the median follow-up time of 22 months (0–144 months) and 106 (12.9%) were stopped early (terminated or withdrawn). Figure 1 shows the causes of early termination of the clinical trials. Poor enrollment was the biggest cause (39%), but other reasons were insufficient funding or a drug supply problem (16%) and principal investigator or organization changes (7%).

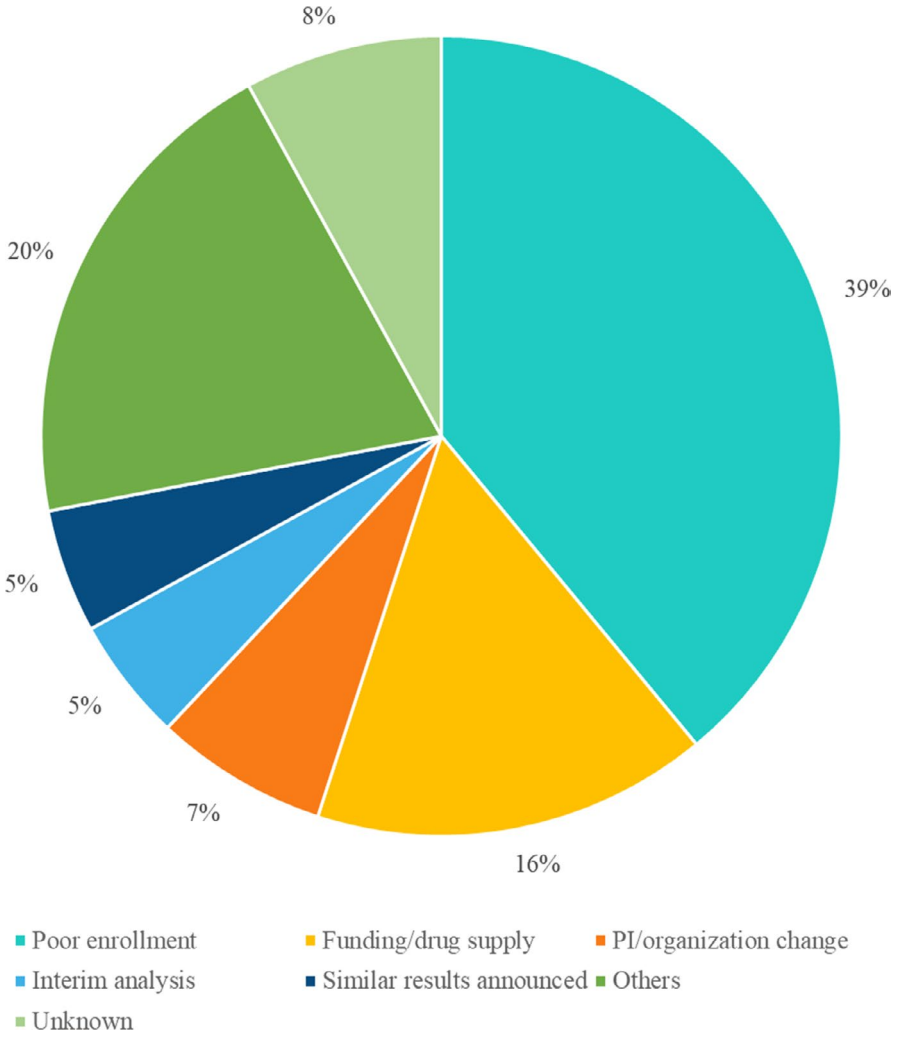
There was no difference in clinical trial completion rates between the dataset for model development and the test dataset (Figure 2; $p = 0.24$). There was also no significant difference in the distribution of clinical trial features contained in each dataset (Table S3).

Cox proportional hazard model

The results of the univariable Cox proportional hazard models are presented in the Table S4. In the univariable analysis, first trimester (hazard ratio [HR] = 0.74), second trimester (HR = 0.72), a sample size of $n \geq 329$ (HR = 0.63), and very high HDI country (HR = 0.39) were significantly related to a low likelihood of clinical trial completion. Abortion (HR = 2.21), labor (HR = 1.82), and iron deficiency anemia (HR = 2.32) were significantly related to a high likelihood of clinical trial completion.

Table 1 shows the results of the multivariable Cox proportional hazard model. A sample size of $n \geq 329$ (HR = 0.53) and very high HDI country (HR = 0.28) were significantly related to a low likelihood of clinical trial completion, while abortion (HR = 3.30), labor (HR = 2.16), and iron deficiency anemia (HR = 2.29) were significantly related to a high likelihood of clinical trial completion.

FIGURE 1 Causes of early clinical trial termination



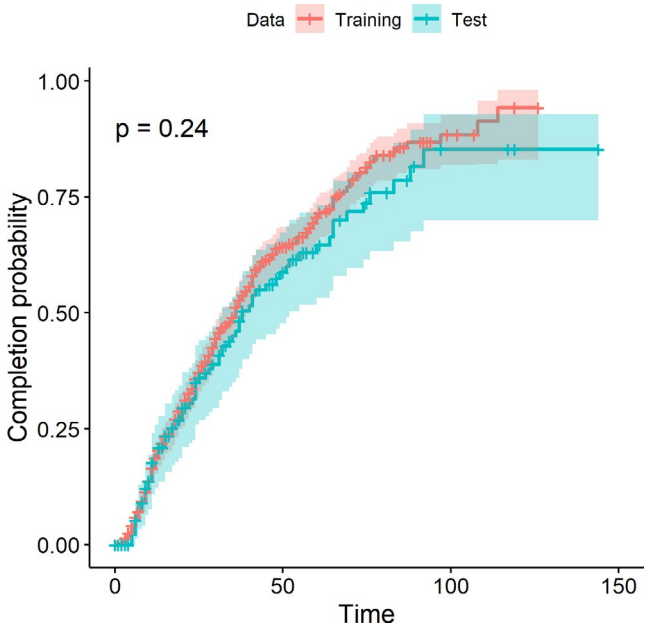


FIGURE 2 Distribution of clinical trial completion of training and test datasets. Colored areas represent 95% confidence interval

TABLE 1 The multivariable Cox proportional hazard model for clinical trial completion

Variable	HR	99% CI	*p* value
Sample size			
$0 \le n < 80$	Reference		
$80 \le n < 150$	0.59	0.39–0.87	<0.01
$150 \le n < 329$	0.56	0.38–0.84	<0.01
$n \ge 329$	0.53	0.36–0.77	<0.01
HDI of study country			
Low	Reference		
Medium	0.84	0.46–1.53	0.45
High	0.55	0.29–1.04	0.02
Very high	0.28	0.15–0.49	<0.01
Targeted medical conditions			
Abortion	3.30	1.92–5.69	<0.01
Labor	2.16	1.55–3.03	<0.01
Anemia	2.92	1.44–5.92	<0.01

Abbreviations: CI, confidence interval; HDI, human development index; HR, hazard ratio.

The C-index of model development dataset was 0.72 (standard error [SE] = 0.014) and that of the test dataset was 0.73 (SE = 0.027).

Neural network model for survival analysis

Final neural network model was one hidden layer with 16 nodes. Hyperparameter spaces and optimal values are shown in the Table S5. The C-index of training dataset was 0.76 (SE = 0.006) and that of test dataset was 0.72 (SE = 0.003). When DeepSurv was performed with only the selected features in the Cox proportional hazard model, the C-indices were 0.73 (SE = 0.003) and 0.71 (SE = 0.005), respectively (Table 2).

Three features decreased the C-index when added in the final model. Feature of phase III, malaria, and first trimester decreased the C-index by 0.004, 0.003, and 0.0005, respectively (Figure 3).

Nomogram

We used the Cox proportional hazard model to develop a nomogram for predicting clinical trial completion because prediction performance was comparable to DeepSurv. The probability of clinical trial completion at 1 year, 3 years, and 5 years can be obtained using the nomogram shown in Figure 4.

DISCUSSION

This study developed models for predicting the completion of clinical trials with pregnant women using both

C-index	Cox proportional hazard model (5 features)	DeepSurv (all features)	DeepSurv (5 features)
Training dataset	0.72 ± 0.014	0.76 ± 0.006	0.73 ± 0.003
Test dataset	0.73 ± 0.027	0.72 ± 0.003	0.71 ± 0.005

Note: Data are presented as mean ± standard error.

Abbreviation: C-index, concordance index.

TABLE 2 Predictive performance (C-index) of the Cox proportional hazard model and DeepSurv

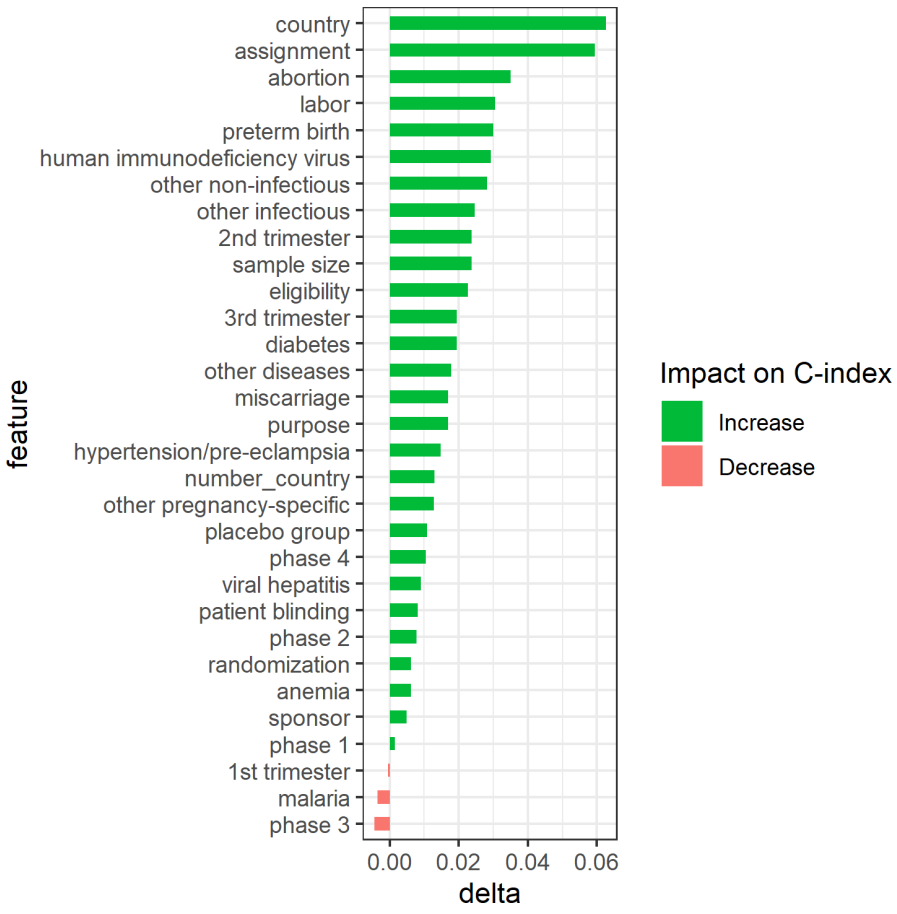


FIGURE 3 Effects of features on concordance index (C-index). It means the amount of change in the C-index when each feature is included in the model at the last time

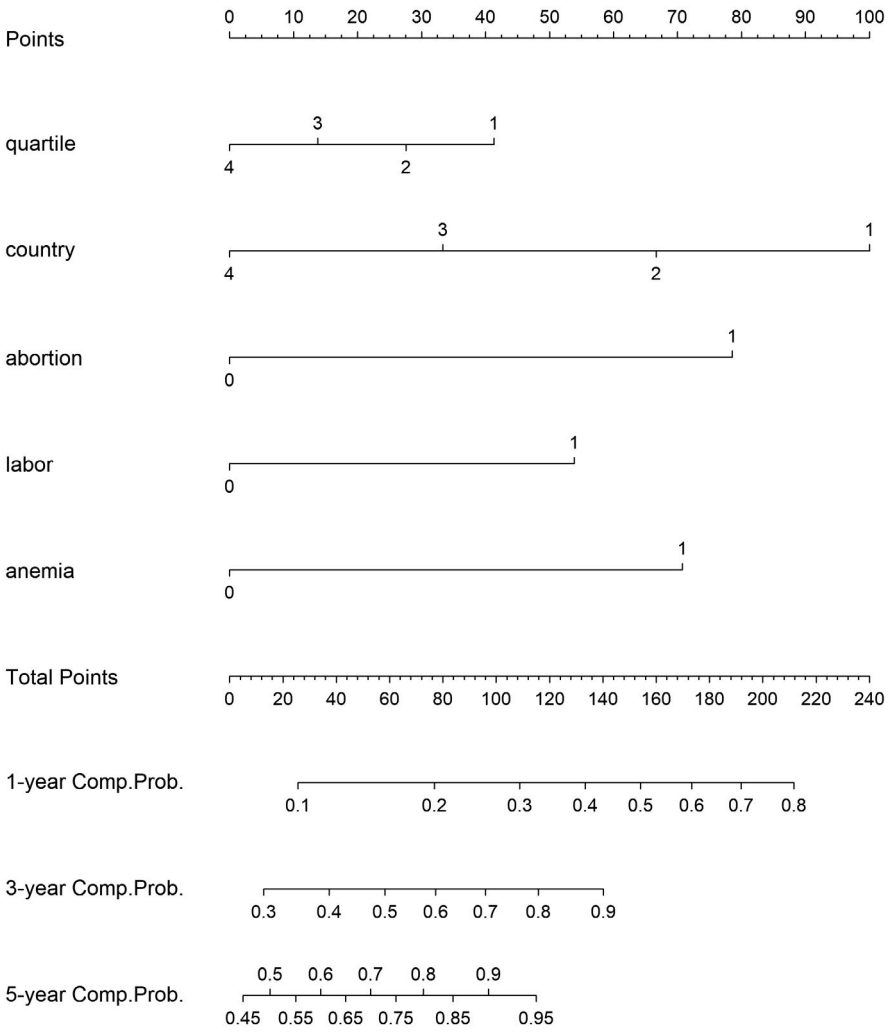


FIGURE 4 Nomogram for predicting clinical trial completion. Quartile: 1 ($0 \leq$ sample size [n] <80), 2 ($80 \leq n < 150$), 3 ($150 \leq n < 329$), and 4 ($n \geq 329$). Country: 1 (low human development index [HDI]), 2 (medium HDI), 3 (high HDI), and 4 (very high HDI). Abortion: 0 (no), and 1 (yes). Labor: 0 (no), and 1 (yes). Iron deficiency anemia: 0 (no), and 1 (yes)

the Cox proportional hazard and neural network models. DeepSurv showed a C-index comparable to the Cox proportional hazard model, although the value was high in the training dataset. Features affecting the completion of clinical studies include the study location, sample size, and some target medical conditions (abortion, labor, and iron deficiency anemia).

The significantly lower probability of clinical trial completion conducted in very high HDI countries compared to those conducted in low HDI countries might be related to subjects' motivation to participate in the study. Most patients participate in clinical trials because they cannot afford expensive treatments. Financially stable patients may not want the hassle (i.e., visits, blood collection, and long questionnaires) of clinical trials.[8,24]

The estimated sample size is another important factor in predicting study completion. A large sample size makes it hard to meet target recruitment goals and leads to longer study periods. In this study, poor enrollment was the biggest reason why clinical trials were terminated early.

In the case of abortion or labor, participant recruitment would be easy because these are single, preplanned, and essential procedures. Iron deficiency anemia is a common medical condition in pregnancy, and subjects can participate easily as the intervention is merely iron supplementation.

Previous studies have shown a high risk of terminating clinical trials for cancer. In this study, clinical trials for pregnant women were analyzed, resulting in different results from previous studies. Cancer research in pregnant women is very rare ($n = 5$; 0.6%) and could not be designated as a feature. Instead, pregnancy-specific conditions could be designated as features in this study. The probability of clinical trial completion over time was predicted by conducting time-to-event analysis without dichotomizing the clinical trial status. These are the strengths of this study.

Cox proportional hazard model and DeepSurv both have a C-index greater than 0.7, which means they have sufficient discrimination ability.[25] Random survival forest has also been used for time-to-event analysis, but we did not include it in this study because previous research did not demonstrate better results than DeepSurv or the Cox proportional hazard model.[26,27] Contrary to expectations

for DeepSurv, its predictive performance was similar to the Cox proportional hazard model. The neural network model is not always superior. In a study predicting future fractures, the C-index of DeepSulv was 0.67 and that of the Cox proportional hazard model was 0.697.[26] The authors attributed the use of 45 features and unbalanced datasets with an 11% event rate.[26] Studies showing the superiority of DeepSurv used 5–14 features and datasets with 17–68% of event rates.[26] In our study, 52.9% of the clinical trials were completed, so the dataset had sufficient event rates. The use of 31 features can cause comparable performances for DeepSurv and Cox proportional hazard model. However, because both the Cox proportional hazard model and DeepSurv results were good, clinical trial completion is thought to be sufficiently explained by the linear model.

Several limitations exist due to the nature of registry-based research. First, this study did not analyze all clinical studies conducted on pregnant women but analyzed a subset registered at ClinicalTrials.gov. An investigator or sponsor may register their research in other registries, which are excluded from this analysis.[28] Second, because ClinicalTrials.gov is a US registry, it contains many studies conducted in North America, which is likely to overestimate the percentage of studies conducted in very high HDI countries.[20,28] In fact, 245 studies (29.9%) were conducted in the United States and Canada, which was higher than in other countries. If another registry was used, there may be differences in the proportion of study location and study characteristics, such as target medical conditions and trimesters. Last, the quality of the data entered in the database depends on study investigators or sponsors.[20,28] Some sections have not been filled out, and some information may not be correct.[28] However, this limitation is somewhat balanced by the fact that ClinicalTrials.gov is the biggest and most well-known clinical study registry.

Both the Cox proportional hazard and neural network models yielded sufficient predicting performance. Those derived as predictors in this study may not be alterable when planning clinical trials. It is good to have a small sample size for clinical trial completion, but we can only lower the sample size to the number in which sufficient statistical power is secured. If possible, clinical trials may be conducted in lower HDI countries. Moreover, predicting the completion of clinical trials can be applied to determine whether to proceed with clinical trials or to allocate resources in the planning stage. We hope that this study will contribute to the execution of future clinical trials in pregnant women.

CONFLICT OF INTEREST

The authors declared no competing interests for this work.

AUTHOR CONTRIBUTIONS

B.K. and M.G.K. wrote the manuscript. M.G.K. and M.K.S. designed the research. B.K. and M.G.K. performed the research. B.K., Y.G.J., H.R.C., S.Y.K., and J.E.J. analyzed the data.

REFERENCES

1. Lyerly AD, Little MO, Faden R. The second wave: toward responsible inclusion of pregnant women in research. *Int J Fem Approaches Bioeth*. 2008;1:5-22.
2. Lupattelli A, Spigset O, Twigg MJ, et al. Medication use in pregnancy: a cross-sectional, multinational web-based study. *BMJ Open*. 2014;4:e004365.
3. Rubin R. Addressing barriers to inclusion of pregnant women in clinical trials. *JAMA*. 2018;320:742-744.
4. Turner MA, Kenny L, Alfirevic Z. Challenges in designing clinical trials to test new drugs in the pregnant woman and fetus. *Clin Perinatol*. 2019;46:399-416.
5. Shields KE, Lyerly AD. Exclusion of pregnant women from industry-sponsored clinical trials. *Obstet Gynecol*. 2013;122:1077-1081.
6. Madan A, Tracy S, Reid R, Henry A. Recruitment difficulties in obstetric trials: a case study and review. *Aust N Z J Obstet Gynaecol*. 2014;54:546-552.
7. Unger A, Jagsch R, Jones H, et al. Randomized controlled trials in pregnancy: scientific and ethical aspects. Exposure to different opioid medications during pregnancy in an intra-individual comparison. *Addiction*. 2011;106:1355-1362.
8. Ballantyne A, Pullon S, Macdonald L, Barthow C, Wickens K, Crane J. The experiences of pregnant women in an interventional clinical trial: Research In Pregnancy Ethics (RIPE) study. *Bioethics*. 2017;31:476-483.
9. Haas DM. From no to yes: the history and ethics of including pregnant women in clinical trials. *Clin Invest*. 2011;1:1349-1351.
10. US Food and Drug Administration. Pregnant Women: Scientific and Ethical Considerations for Inclusion in Clinical Trials Guidance for Industry. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/pregnant-women-scientific-and-ethical-considerations-inclusion-clinical-trials. Accessed September 25, 2021.
11. Macklin R. Enrolling pregnant women in biomedical research. *Lancet*. 2010;375:632-633.
12. Council for International Organizations of Medical Sciences. International ethical guidelines for biomedical research involving human subjects. *Bull Med Ethics*. 2002;182:17-23.
13. Follett L, Geletta S, Laugerman M. Quantifying risk associated with clinical trial termination: a text mining approach. *Inf Process Manag*. 2019;56:516-525.
14. Elkin ME, Zhu X. Predictive modeling of clinical trial terminations using feature engineering and embedding learning. *Sci Rep*. 2021;11:3446.
15. Bradburn MJ, Clark TG, Love SB, Altman DG. Survival analysis part II: multivariate data analysis–an introduction to concepts and methods. *Br J Cancer*. 2003;89:431-436.
16. Katzman JL, Shaham U, Cloninger A, Bates J, Jiang T, Kluger Y. DeepSurv: personalized treatment recommender system using a Cox proportional hazards deep neural network. *BMC Med Res Methodol*. 2018;18:24.

17. Gillen JE, Tse T, Ide NC, McCray AT. Design, implementation and management of a web-based data entry system for ClinicalTrials.gov. *Stud Health Technol Inform*. 2014;107:1466-1470.
18. Morgan JA, Cooper DB. Pregnancy dating. *StatPearls*. https://www.ncbi.nlm.nih.gov/books/NBK442018/. Accessed September 25, 2021.
19. Conceição P. *Human development report 2019*. United Nations Development Programme; 2019.
20. Stockmann C, Sherwin CMT, Koren G, et al. Characteristics and publication patterns of obstetric studies registered in ClinicalTrials.gov. *J Clin Pharmacol*. 2014;54:432-437.
21. Bourgeois FT, Murthy S, Pinto C, Olson KL, Ioannidis JP, Mandl KD. Pediatric versus adult drug trials for conditions with high pediatric disease burden. *Pediatrics*. 2012;130:285-292.
22. Lewis F, Butler A, Gilbert L. A unified approach to model selection using the likelihood ratio test. *Methods Ecol Evol*. 2011;2:155-162.
23. Harrell FE Jr, Califf RM, Pryor DB, Lee KL, Rosati RA. Evaluating the yield of medical tests. *JAMA*. 1982;247:2543-2546.
24. Rodger MA, Makropoulos D, Walker M, Keely E, Karovitch A, Wells PS. Participation of pregnant women in clinical trials: will they participate and why? *Am J Perinatol*. 2003;20:69-76.
25. Shi X, Xu L, Ma B, Wang S. Development and validation of a nomogram to predict the prognosis of patients with gastric cardia cancer. *Sci Rep*. 2020;10:14143.
26. de Vries BCS, Hegeman JH, Nijmeijer W, Geerdink J, Seifert C, Groothuis-Oudshoorn CGM. Comparing three machine learning approaches to design a risk assessment tool for future fractures: predicting a subsequent major osteoporotic fracture in fracture patients with osteopenia and osteoporosis. *Osteoporos Int*. 2021;32:437-449.
27. Byun S-S, Heo TS, Choi JM, et al. Deep learning based prediction of prognosis in nonmetastatic clear cell renal cell carcinoma. *Sci Rep*. 2021;11:1242.
28. Zippel C, Ronski SC, Bohnet-Joschko S, Giesel FL, Kopka K. Current status of PSMA-radiotracers for prostate cancer: data analysis of prospective trials listed on ClinicalTrials.gov. *Pharmaceuticals*. 2020;13(1):12.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Kim B, Jang YJ, Cho HR, et al. Predicting completion of clinical trials in pregnant women: Cox proportional hazard and neural network models. *Clin Transl Sci*. 2022;15:691–699. doi:10.1111/cts.13187