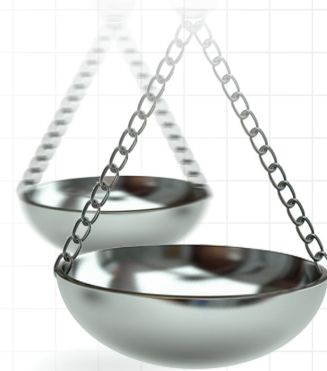




Directly comparing randomized and real-world controls in adjuvant breast cancer

Journal of **Comparative Effectiveness Research**

Brigham Walker^{*1} , Marley Boyd², Michelle Silver³, Junxin Shi⁴, Purva Bulsara⁴, Nino Balanchivadze⁵ & Nicholas J Robert⁴

¹Department of Health Policy and Management, Tulane University Celia Scott Weatherhead School of Public Health, LA 70112, USA

²TriNetX, MA 02140, USA

³Genesis Research, NJ 07030, USA

⁴Ontada, MA 02109, USA

⁵Virginia Oncology Associates, US Oncology Network, VA 23502, USA

*Author for correspondence: Tel.: +1 504 988 1312; bwalker6@tulane.edu

Aim: We sought to directly compare contemporaneous real-world (RW) and randomized controlled trial (RCT) control arms to assess the feasibility of utilizing RW controls in early breast cancer treatment.

Materials & methods: Women with human epidermal growth factor receptor 2-negative high-risk breast cancer receiving adjuvant docetaxel, doxorubicin and cyclophosphamide for six cycles between 29 May 2007 and 21 November 2013 were included, mirroring the control arm and study timeframe from our reference trial ‘Anthracyclines in Early Breast Cancer: The ABC Trials’. Structured data was extracted from the iKnowMed electronic medical record system, which included both RW and trial patients allowing for their direct comparison. Propensity scores were used for baseline characteristic balancing, and Kaplan–Meier analysis was conducted. **Results:** In the final sample of 306 patients (153 per arm drawn from 155 RCT patients and 703 RW patients in the study timeframe), the matching process yielded balanced cohorts along available covariates. Kaplan–Meier curves and disease-free survival distributions did not suggest significant differences between RW and trial arms (log rank $p = 0.98$). **Conclusion:** Despite limitations, including the use of structured data only, the study suggests that RW data could have served as a synthetic control in this setting.

Plain language summary

What is this article about? This article utilizes electronic medical records on patients who were in a clinical trial to evaluate whether similar patients who were not in the clinical trial but receiving the same treatment for early breast cancer had similar disease-free survival patterns. This is relevant to researchers because there is growing interest in understanding whether generating clinical evidence without a clinical trial control arm (and instead using a ‘real-world’ control arm) would be a reliable method for research.

What were the results? After selecting real-world patients who on balance were most similar to the clinical trial patients, we find that the patients had very similar disease-free survival patterns.

What do the results mean? These results suggest that, in this clinical setting, these patients would have likely served as reasonable substitutes for the clinical control arm of a clinical trial.

First draft submitted: 10 September 2025; Accepted for publication: 24 April 2026; Published online: 29 May 2026

Keywords: adjuvant breast cancer • electronic medical records • external controls • real-world controls • real-world data

Using synthetic control arms (SCAs), which are cohorts comprised of retrospective clinical trial data, or external control arms (ECAs), which are cohorts comprised of nonrandomized real-world (RW) patients, to evaluate cancer treatment efficacy has become increasingly prevalent. This is attributed to the complexity, cost, ethical and



enrollment challenges associated with randomized controlled trials (RCTs). In breast cancer, these methods have been used to verify single-arm trial results [1] and to measure racial disparities in outcomes [2]. Some studies have displayed overlapping Kaplan–Meier overall survival curves between SCAs and trial control cohorts suggesting similar patterns [3], while others have shown mixed trial replicability [4].

Multiple studies recommend careful specification [5,6], pre-specification [7] with matching [8,9] and safety data [6], and provide caution on the limitations of these methods in supplanting the RCT gold standard [10]. Using RW data further challenges analyses given potential missing data challenges [7,8] and differences in how the data are generated as compared with the trial setting [8,11].

Here, we directly compared RW and RCT control cohorts in early breast cancer using data generated on a common electronic medical record (EMR) to assess the viability of contemporaneous RW controls in this setting. The reference RCT for this study was ‘Anthracyclines in Early Breast Cancer: The ABC Trials’ which compared a treatment arm of docetaxel and cyclophosphamide for six cycles (TC6) with a control arm of docetaxel, doxorubicin, and cyclophosphamide for six cycles (TAC6) in the adjuvant breast cancer setting [5].

In early-stage breast cancer, the superiority of four cycles of TC to four cycles of doxorubicin and cyclophosphamide (AC) had been documented [12]. However, the effectiveness of TC in comparison with a standard TAC6 regimen was unknown. In breast cancer patients with high-risk human epidermal growth factor receptor 2 (HER2) negative disease, the ABC Trials reported that TAC6 regimens displayed increased invasive disease-free survival compared with TC6 regimens over the study period.

Materials & methods

We included women with HER2-negative high-risk breast cancer who received TAC6 treatment in the adjuvant setting between 29 May 2007 (the earliest trial open date) and 21 November 2013 (the date of accrual closure) both within and outside the ABC trial. Patient data were extracted from the iKnowMed (iKM) EMR system, which is utilized exclusively in oncology practices affiliated with the US Oncology Network.

We selected the ABC trial as the reference RCT due to it being conducted within practices of the US Oncology Network, facilitating access to data from both RCT and RW patients within iKM. This enabled a direct comparison between patients enrolled in the ABC trial as controls and contemporary RW controls. The RW control arm was established by applying the inclusion and exclusion criteria from the ABC trials. For both arms, only structured data were utilized.

The primary end point of the study was disease-free-survival (DFS), defined as the time from initiation of adjuvant chemotherapy to recurrence. Recurrence was characterized as initiation of new chemotherapy or first line therapy after adjuvant treatment, a gap of 6 months or greater in endocrine therapy, noted metastases, or death.

To account for baseline differences between groups, we used 1:1 greedy nearest neighbor matching on propensity scores with a caliper of 0.1. This matching method ensured a balanced sample across key covariates, including age, BMI, stage, hormone receptor status, number of positive lymph nodes (obtained through staging data [N0, N1, N2 and N3]), smoking history and region and network of care. Time since diagnosis, patient race, Eastern Cooperative Oncology Group (ECOG) status and region and network of care were also considered but not selected for matching due to insufficient availability in the structured data. A standardized difference >0.2 was considered an imbalance between cohorts.

DFS was estimated using the Kaplan–Meier method with 95% CIs. At risk measures by cohort at each year interval were also generated. Statistical analyses were conducted using SAS version 9.4 [13].

Results

There were 155 RCT patients and 703 RW patients in the study timeframe. Propensity scores were calculated using a logistic regression model on age at diagnosis, BMI, stage at treatment initiation, smoking history, hormone receptor status and number of positive nodes.

Prior to matching, the RW sample was relatively younger (mean age of 51.6 vs 52.9 years), less likely to have stage I disease (3.1% vs 11.6%), more likely to be hormone receptor positive (72.8% vs 65.8%), more likely to have between 1 and 3 positive nodes (24.2% vs 14.8%), and less likely to have an undocumented node status (46.8% vs 62.6%) (Table 1 ‘Before matching’).

Table 1. Distributions and standardized differences (std diff) of covariates between cohorts.

	Before matching [†]			After matching		
	RCT	RW	Std diff.	RCT	RW	Std diff.
Mean age at diagnosis, years	52.9	51.6	(0.15)	52.8	53.2	(0.05)
Stage, %						
I	11.6	3.1	(0.33)	10.5	9.8	(0.02)
IA	5.2	5.0	(0.01)	5.2	5.2	(0.00)
IB	0.6	0.6	(0.01)	0.7	1.3	(0.07)
IIA	41.3	30.2	(0.23)	41.8	35.9	(0.12)
IIB	16.8	26.5	(0.24)	17.0	19.0	(0.05)
IIIA	21.9	24.3	(0.06)	22.2	26.1	(0.09)
IIIB	0.0	1.3	(0.16)	0.0	0.0	(0.00)
IIIC	2.6	9.1	(0.28)	2.6	2.6	(0.00)
Hormone receptor status, %						
Positive	65.8	72.8	(0.15)	66.0	68.6	(0.06)
Negative	33.5	27.2	(0.14)	33.3	31.4	(0.04)
Undocumented	0.6	0.0	(0.11)	0.7	0.0	(0.12)
Positive nodes, n, %						
0	14.2	16.2	(0.06)	14.4	11.1	(0.10)
1–3	14.8	24.2	(0.24)	15.0	11.1	(0.12)
4–9	6.5	8.8	(0.09)	6.5	5.2	(0.06)
10+	1.9	4.0	(0.12)	2.0	2.6	(0.04)
Undocumented	62.6	46.8	(0.32)	62.1	69.9	(0.17)

[†] A standardized difference >0.2 was considered an imbalance between cohorts.
RCT: Randomized controlled trial; RW: real-world.

The matching process subsequently resulted in cohorts of 153 patients each (i.e., two RCT patients excluded due to missing covariates and 550 RW patients excluded due to the availability of superior matches) that were balanced on each covariate with standardized differences <0.2 (Table 1). This indicates that the RW and RCT arms are comparable.

Mean follow-up was 50.0 months in the RCT arm and 50.1 months in the RW arm. The Kaplan–Meier curves for DFS distributions did not demonstrate a significant difference between groups (Figure 1). In the RCT arm, 15.0% of patients experienced an event, compared with 13.7% in the RW arm. The DFS in the RCT arm at 12 months was 97.4%, while in the RW sample it was 96.0%. Overall, the distribution of DFS was not significantly different between the two arms, with a log rank p-value of 0.98 (see Table 2).

Discussion

Our study contributes to the growing body of evidence evaluating the use of RW data in oncology research and in augmenting single arm trials. By demonstrating the feasibility of comparing RW and RCT control arms in early breast cancer, we highlight one potential use of RW data to augment or substitute traditional RCT controls, particularly in situations where conducting RCTs is challenging or not feasible.

Limitations

Our study has several limitations. Firstly, the sample size was small compared with the sample size reported in the ABC trials, yet our primary objective was to assess whether patients treated within the RCT had similar outcomes to those treated outside the RCT. Additionally, we relied solely on structured data, leading to the omission of important data points insufficiently captured (e.g., ECOG status, racial and ethnic identities, genomic measures, etc.) in this format. Certain important clinical and genomic characteristics were not considered in matching the cohort. We did not evaluate censorship directly; though, potential differences by cohort may lead to statistically significant differences in DFS (which we do not find). Moreover, indirect methods were utilized to determine disease free intervals and DFS was used as a proxy for invasive disease-free survival. Lastly, our results may not extend beyond settings with similar clinical characteristics.

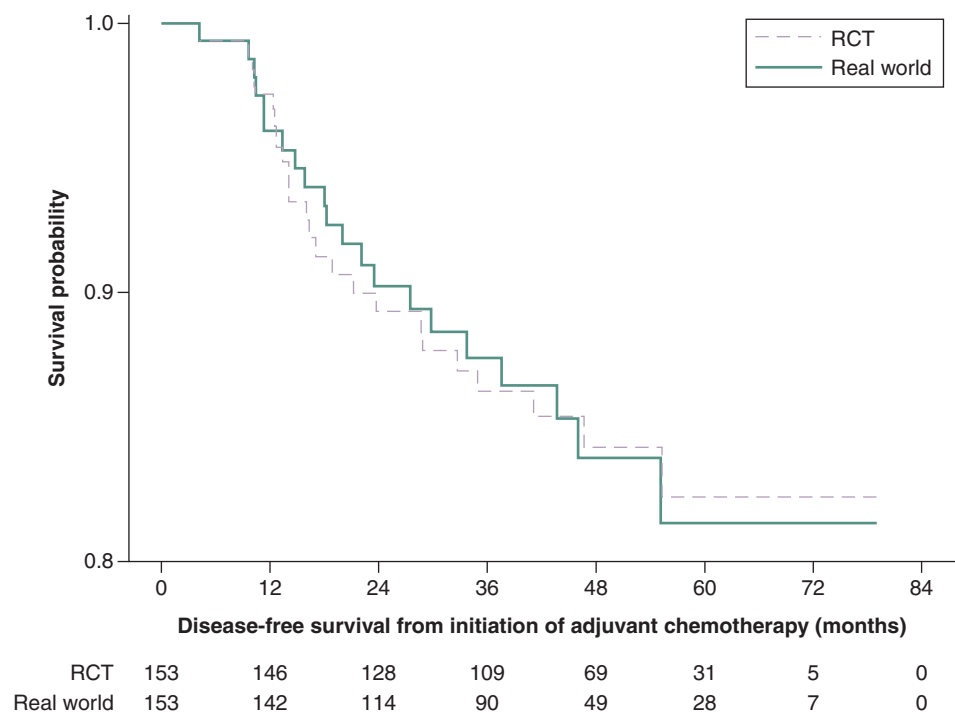


Figure 1. Kaplan–Meier curves of disease-free survival, real-world versus randomized controlled trial (in months). DFS was indexed on the initiation of adjuvant chemotherapy therapy. DFS: Disease-free-survival; RCT: Randomized controlled trial.

Table 2. Post-match Kaplan–Meier estimates of disease-free survival, by randomized controlled trial and real-world .			
	Treatment cohorts		Log rank p-value
	TAC6 RCT	TAC6 RW	
N	153	153	–
Events (%)	23 (15.0)	21 (13.7)	–
Mean follow time (SE), months	50.0 (1.1)	50.1 (1.1)	–
DFS, % (95% CI)			0.98
12 months	97.4 (93.1, 99.0)	96.0 (91.3, 98.2)	–
24 months	89.3 (83.1, 93.3)	90.2 (84.0, 94.1)	–
36 months	86.3 (79.6, 91.0)	87.6 (80.7, 92.1)	–
48 months	84.2 (76.9, 89.4)	83.9 (75.7, 89.5)	–
60 months	82.4 (74.1, 88.3)	81.5 (71.7, 88.1)	–
DFS: Disease free survival; RCT: Randomized controlled trial; RW: real-world; TAC6: docetaxel, doxorubicin, and cyclophosphamide for six cycles; TC6: docetaxel and cyclophosphamide for six cycles; SE: Standard Error.			

Conclusion

In this comparison, the matched contemporaneous RW control arm appears to replicate DFS from the available RCT arm. Our results cautiously suggest that, in this setting, RW data may have been considered as a potentially suitable contemporaneous ECA. However, the concerns raised in the literature relating to the merits of pre-specification, augmentation with additional data, and considering alternative matching techniques apply here and further research incorporating larger and more diverse datasets and rigorous validation methods is warranted to enhance the robustness and generalizability of these findings.

Summary points

- The use of synthetic (SCA) or external (ECA) control arms to evaluate cancer treatment efficacy has become increasingly prevalent, for reasons such as the complexity, cost, ethical and enrollment challenges associated with RCTs.
- However, using RW data in this context presents challenges.
- This study directly compared RW and RCT control cohorts in early breast cancer using data generated on a common EMR to assess the viability of contemporaneous RW controls in this setting. The ABC trial was used as the reference RCT.
- Women with HER2-negative high-risk breast cancer who received TAC6 treatment in the adjuvant setting between 29 May 2007 and 21 November 2013 both within and outside the ABC trial were included in the study.
- Disease-free survival (DFS), defined as the time from initiation of adjuvant chemotherapy to recurrence, was the primary end point, and was estimated using the Kaplan–Meier method with 95% CIs.
- Propensity score matching resulted in 153 patients in each cohort (RCT and RW).
- The Kaplan–Meier curves for DFS distributions did not demonstrate a significant difference between the RCT and RW groups (log rank p-value of 0.98).
- This study contributes to the growing body of evidence evaluating the use of RW data in oncology research and in augmenting single-arm trials. The findings cautiously suggest that, in this setting, RW data may have been considered as a potentially suitable contemporaneous ECA.
- However, concerns raised in the literature relating to the merits of pre-specification, augmentation with additional data and considering alternative matching techniques apply here and further research incorporating larger and more diverse datasets and rigorous validation methods is warranted to enhance the robustness and generalizability of these findings.

Financial disclosure

The author(s) received no financial and/or material support for this research or the creation of this work.

Competing interest disclosure

All authors at some point were employed by or affiliated with the organization that provided data for this manuscript. The authors have no other competing interests or relevant affiliations with any organization or entity with the subject matter or materials discussed in the manuscript apart from those disclosed.

Writing assistance disclosure

No funded writing assistance was utilized in the production of this manuscript.

Data transparency statement

The authors certify that this manuscript reports the original results of a real-world evidence study. A prespecified study protocol was developed. The underlying data and analytic code used in the analysis are not publicly available. Reporting checklists were not provided for this manuscript.

Open access

This work is licensed under the Attribution-NonCommercial-NoDerivatives 4.0 Unported License. To view a copy of this license, visit <https://creativecommons.org/licenses/by-nc-nd/4.0/>

References

1. Amiri-Kordestani L, Xie D, Tolane SM *et al.* A Food and Drug Administration analysis of survival outcomes comparing the Adjuvant Paclitaxel and Trastuzumab trial with an external control from historical clinical trials. *Ann. Oncol.* 31(12), 1704–1708 (2020).
2. Watson ML, Holcomb K, Newman L *et al.* Evaluating progression-free survival in black and white women with triple-negative breast cancer in pooled clinical trials from a synthetic control database (SCD) and real-world electronic medical records (EMR). *J. Clin. Oncol.* 38(Suppl. 15), e13102 (2020).
3. Davi R, Yin X, Stewart M. Exploring the validity of a synthetic control arm (SCA) for augmentation or replacement of a randomized control in difficult-to-study indications: a case study in relapsed or refractory multiple myeloma (R/R MM). *J. Clin. Oncol.* 38(Suppl. 15), e20521 (2020).
4. Kanapuru B, Gong Y, Mishra-Kalyani PS *et al.* Project Switch: lenalidomide and dexamethasone (Len-Dex) as a potential synthetic control arm (SCA) in relapsed or refractory multiple myeloma (rrMM). *J. Clin. Oncol.* 37(Suppl. 15), 8047 (2019).
5. Mishra-Kalyani PS, Amiri Kordestani L, Rivera DR *et al.* External control arms in oncology: current use and future directions. *Ann. Oncol.* 33(4), 376–383 (2022).

6. Hermans SJF, van der Maas NG, van Norden Y *et al.* Externally controlled studies using real-world data in patients with hematological cancers: a systematic review. *JAMA Oncol.* 10(10), 1426–1436 (2024).
7. Siu DHW, Lin FPY, Cho D *et al.* Framework for the use of external controls to evaluate treatment outcomes in precision oncology trials. *JCO Precis. Oncol.* 8, e2300317 (2024).
8. Schmidli H, Häring DA, Thomas M, Cassidy A, Weber S, Bretz F. Beyond randomized clinical trials: use of external controls. *Clin. Pharmacol. Ther.* 107(4), 806–816 (2020).
9. Walker B, Ray HE, Shao P, D'Ambrosio C, White C, Walker MS. Comparing prospectively assigned trial and real-world lung cancer patients. *J. Comp. Eff. Res.* 13(7), e230176 (2024).
10. Nuño MM, Pugh SL, Ji L, Piao J, Dignam JJ, Steingrimsson JA. On the use of external controls in clinical trials. *J. Natl Cancer Inst. Monogr.* 2025(68), 30–34 (2025).
11. Walker B, Boyd M, Aguilar K *et al.* Comparisons of real-world time-to-event end points in oncology research. *JCO Clin. Cancer Inform.* 5, 45–46 (2021).
12. Jones SE, Savin MA, Holmes FA *et al.* Phase III trial comparing doxorubicin plus cyclophosphamide with docetaxel plus cyclophosphamide as adjuvant therapy for operable breast cancer. *J. Clin. Oncol.* 24(34), 5381–5387 (2006).
13. SAS version 9.4. SAS (Statistical Analysis System) Institute Incorporated, NC, USA. <https://support.sas.com/software/94/>