

Real-World Safety Evaluation through Electronic Medical Record - A Transformative Approach to Korean Post-Marketing Surveillance

Ha-Yeong Gil¹, Yeo Rae Moon², YongHyun Cho², Seung-Hee Jeong^{1†}

¹RWE Generation Team, Pfizer Inc., ²RWE Development Team, KakaoHealthcare Corp.

Purpose

Post-marketing surveillance (PMS) in Korea has been conducted for decades to generate safety data for regulatory purposes. However, traditional PMS based on Drug Use-Results Surveys (DURS) does not always adequately reflect real-world clinical practice, resulting in challenges in patient recruitment and substantial cost burdens. Electronic medical record-based data, combined with a hybrid approach of rule-based NLP and Large Language Models (LLMs), is emerging as an effective alternative for safety analysis. This study aims to evaluate whether retrospective safety assessments using EMR data can serve as a viable alternative to traditional PMS.

Conclusion

EMR-based safety analysis provided broader clinical information and identified adverse events in more patients with a shorter time compared to traditional PMS, demonstrating its viability as an alternative approach.

[†]Presenting Author:

Seung-Hee.Jeong@pfizer.com

Co-author:

Ha-Yeong.Gil@pfizer.com

Acknowledgment:

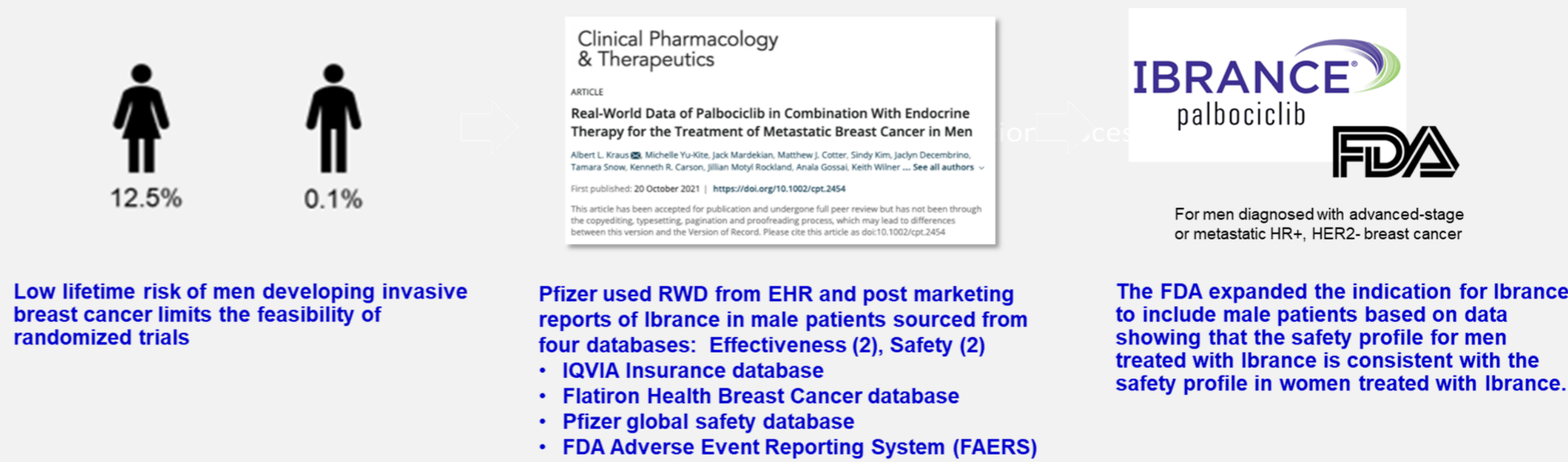
This study is sponsored by Pfizer Pharmaceuticals Korea Ltd. and we extend our sincere gratitude to all Pfizer colleagues who contributed to the IBRANCE database studies.



Background

- In Korea, post-marketing drug safety management mainly relies on traditional PMS through Drug Use-Results Surveys (DURS), which collect safety and efficacy data during routine clinical practice. However, these studies have major limitations: high cost and resource burden, lack of control groups, selection bias, and insufficient statistical validation — often resulting in descriptive case collection rather than robust evidence generation. Since the primary goal of post-marketing safety is to identify previously unknown adverse drug reactions, more proactive monitoring strategies are needed.
- Globally, the importance of Real-World Data (RWD) and Real-World Evidence (RWE) is increasing, with growing use in regulatory decision-making. Although frameworks now support real-world safety data collection, RWD utilization in Korean post-marketing studies remains rare, highlighting the need to explore and expand RWD-based approaches.

- IBRANCE®**, also known as palbociclib, was first approved for the treatment of hormone receptor-positive, human epidermal growth factor receptor 2-negative (HR+/HER2-) metastatic breast cancer in postmenopausal women, in combination with letrozole.



- IBRANCE® PMS** was approved in Korea in 2016 with a regulatory requirement to collect prospective 3,000 cases within 6 years. However, patient enrollment was challenging because most patients were female breast cancer cases, and large hospitals prioritized global clinical trials or investigator-sponsored research over PMS. In 2020, the target sample size was reduced to 1,000 cases, but continuous enrollment remained difficult. As a result, study period was extended by 9 years. Therefore, we planned the IBRANCE DB protocol using EMR-based safety analysis as an alternative, and it was approved by Korea MFDS in July 2024.

Study objective

- This study aims to use a large real-world database to describe breast cancer patients treated with IBRANCE and assess its safety and effectiveness.**
- Primary Objective:** Evaluate safety, including adverse events during treatment, and identify related risk factors.
- Secondary Objectives:** Assess effectiveness through tumor response and compare treatment outcomes across patient groups based on predictors such as age and pathological characteristics.

Materials and Methods

Data Source This study used EMR-based database from tertiary multicenter hospitals in Korea, collected between August 29, 2015, and September 30, 2023.

This study included patients with HR(+)/HER2(-) advanced or metastatic breast cancer treated with palbociclib combined with an aromatase inhibitor or Fulvestrant.

Inclusion criteria

- Adult women (≥ 18 years of age) with at least one visit with a breast cancer ICD10 diagnosis code (C50.xx)
- Patients with locally advanced breast cancer or metastatic breast cancer with a staging classification of stage III, stage IV (distant metastases).
- Patients with a laboratory test HR(+) and HER2(-) before or up to 60 days after advanced/metastatic breast cancer diagnosis date.
- HR(+) is defined as any ER(+) or PR(+).
- HER2(-) is defined as any HER2-negative test and the absence of a positive test (immunohistochemistry (IHC) positive (3+), fluorescence in situ hybridization (FISH) positive /amplified)
- Patients who received at least one initial prescription of IBRANCE® following a diagnosis of advanced /metastatic breast cancer

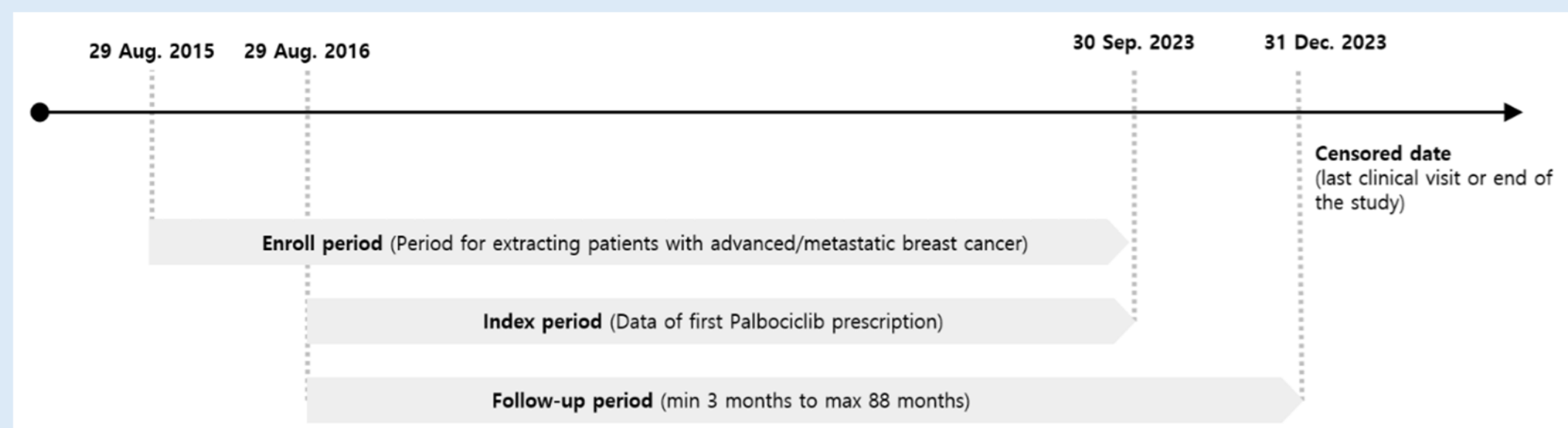
Exclusion criteria

- Breast cancer with HR(-) and HER2(+) on laboratory testing.
- More than a 90-day gap between breast cancer diagnosis date and next visit.
- Patients who treated as off-label

Exposure Patients with advanced/metastatic breast cancer who are HR(+)/HER2(-) and have received IBRANCE with combination therapy. The combination therapy is defined as the prescription of letrozole, anastrozole, fulvestrant within 28 days following IBRANCE.

• Palbociclib + Letrozole • Palbociclib + Anastrozole • Palbociclib + Fulvestrant

Index date & Follow-up The index date for all treatment-related events, including adverse events and tumor response assessments, is defined as the date of the first IBRANCE prescription.



- For the safety assessment, the observation period will be defined as from the index date to the last dose + 28 days. Adverse events include adverse events written in accordance with the National Cancer Institute Common Terminology Criteria for Adverse Events (CTCAE) Version 5.0.
- For the effectiveness assessment, the observation period will be defined as the index date to the date of death or censoring date, whichever occurs first. The censoring date is defined as the last clinical visit or the end of the study period. This includes the evaluation of overall survival (OS).

Data Management Data for IBRANCE DB was collected from EMR DB through clinical data warehouse (CDW). All data used in the study are standardized and de-identified during the extraction process through the CDW system.

Data Processing

- Structured data** Diagnosis codes (ICD-10), Prescription records, Laboratory test results, Demographic information, and Treatment history
- Unstructured data** Such as clinical notes and nursing records (i.e., free-text clinical charts)—is processed through an integrated approach combining rule-based logic and generative AI (LLMs) to reliably extract patient cohorts and detect adverse events with enhanced sensitivity.
- Safety - Adverse events** are defined using the National Cancer Institute's Common Terminology Criteria for Adverse Events (CTCAE). Definitions of adverse events based on the CTCAE were established using operational definitions derived from published literature and CTCAE laboratory criteria. When relevant literature was not available, appropriate criteria were developed through clinical expert review.

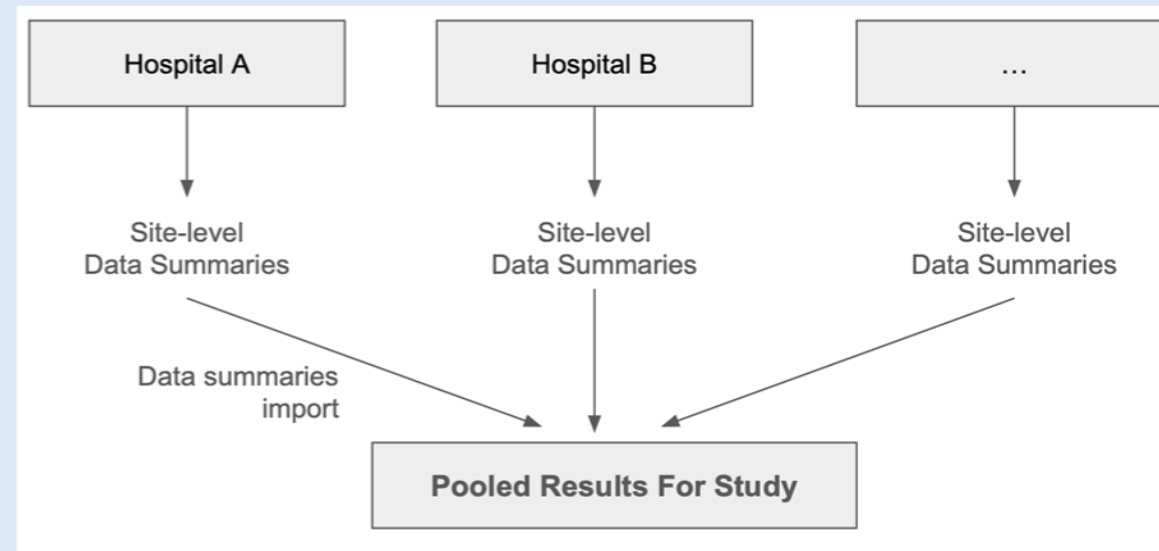
SOC	PT	Source	Operation definition
Infections and infectious	Upper respiratory tract infection	Diagnostic, record, medical note, clinical examination	If any of the following diagnosis codes not included and the following terms found from medical notes or clinical notes. 1) Diagnosis code: No diagnosis received with ICD-10 J00-J06, J39, A36.8, or A36.9 codes within 14 days from the index date. 2) Medical records: terms related to upper respiratory tract infections ('URL', 'nasopharyngitis', 'pharyngitis', 'sinusitis', 'laryngitis', 'tonsillitis'). 3) Clinical examination: confirmed sinusitis on PMS routine examination
Infections and infectious	Infections	Diagnostic records, Medication prescriptions	An event is defined as occurring if either of the two definitions are met. 1) If any of the following diagnosis codes included and the following medication prescribed. Diagnosis codes: ICD-10 A00-99, B00-99, medications: systemic antibiotics, antivirals, antifungals. 2) If any of the following diagnosis codes included (ICDM-9-CM) 001.xx-018.xx, 030.xx-041.xx, 045.xx-049.xx, 050.xx-057.xx, 110.xx-118.xx, 070.xx-079.xx, 130.xx-136.xx, 460.xx-466.xx, 480.xx-488.xx, 038.0x, 038.19, 038.80, 038.9x, 995.91, 995.92

- Effectiveness** was evaluated in patients who received IBRANCE in combination with an aromatase inhibitor as first-line hormone therapy, and who were prescribed IBRANCE for at least 4 cycles.
- Data QC** To ensure the validity and reliability of the data, quality checks were conducted in accordance with standard quality control processes.

Statistical Analysis

- Safety analysis** focused on evaluating the safety profile of patients who received at least one dose of IBRANCE combined with letrozole, anastrozole, or fulvestrant. Incidence of adverse events will be calculated as the proportion of patients experiencing at least one event, with 95% CIs.
- Effectiveness** was assessed in patients receiving ≥ 4 cycles of IBRANCE with an aromatase inhibitor as first-line therapy and documented response; reported with 95% CIs (Wilson method), analyzed using Kaplan–Meier curves with 95% CIs.

- Overall Survival (OS): Time from index date to death or censoring
- Real-World Tumor Response (rwTR): Clinician-assessed response (CR, PR, SD, PD) based on radiologic and clinical data; best overall response (BOR) summarized
- Objective Response Rate (ORR): Proportion of patients achieving CR or PR



Workflow for Generating Analysis-Ready AE Data

Unstructured Data

... Patient reports diarrhea since yesterday, with 4 or more bowel movements per day.
Stools { ...
Mild abd. "Since starting Ibrance, patient complains of abdominal pain, described as dull discomfort.
Will pr. Abd. pain not worsening, managed conservatively.
Bowel movements regular.
Will continue to monitor.

Rule-based NLP

- Keyword matching
- Contextual interpretation
- CTCAE mapping

* Applied to unstructured text

Structured Data

Laboratory test Table			
Patient id	Reported datetime	Value	Value unit
1			
2			
3			
4			

Diagnosis Table			
Patient id	Diagnosis date	ICD-10	Diagnosis name
1	XXXX.XX.XX	K58.0	XXXXXX
2	XXXX.XX.XX	R41	XXXXXX
3	XXXX.XX.XX	N72	XXXXXX
4	XXXX.XX.XX	G96.0	XXXXXX

Adverse event Definition

- CTCAE-based AE definition
- Event date determination
- Record de-duplication

Analysis-ready AE table

Patient ID	Group	SOC	PT
1	Ibrance+AI	Gastrointestinal disorders	Diarrhoea
2	Ibrance+Fulvestrant	Blood and lymphatic system disorders	Anaemia
3	Ibrance+AI	Skin and subcutaneous tissue disorders	Dry skin

Limitations

- Retrospective EMR data may miss mild or non-specific AEs compared with structured PMS reporting
- Site-level differences in coding and documentation reduce consistency versus protocol-based PMS
- Causality assessment is limited in retrospective analyses compared with physician-judged PMS
- PMS-specific fields (e.g., detailed discontinuation reasons) are not consistently captured in EMRs

Future Work

- Develop an EMR–PMS hybrid model to enable automated, PMS-aligned data capture
- Apply advanced NLP for more accurate AE detection in free-text notes
- Promote national coding standardization to support scalable EMR-based PMS
- Include comparator cohorts to strengthen interpretation of safety outcomes