

PREDICCIÓN DE LA RESPUESTA AL TRATAMIENTO EN PACIENTES CON CÁNCER DE MAMA TRIPLE NEGATIVO A PARTIR DE IMÁGENES DE BIOPSIA ALMACENADAS EN MÚLTIPLES CENTROS

Victor Dillard, Director de operaciones comerciales en Owkin.
Alberto Pardo, Jefe de sistemas en Fundación Jiménez Díaz.

CONTEXTO

La generalización de las evaluaciones médicas y las nuevas terapias eficaces han dado lugar a un aumento del 21% en la tasa de supervivencia a 5 años de los cánceres de mama no metastásicos desde 1975, hasta alcanzar el 91% actual¹. Sin embargo, sigue habiendo subtipos de cánceres de mama con grandes necesidades médicas por cubrir. El cáncer de mama triple negativo ("CMTN") es una forma agresiva de cáncer de mama que afecta aproximadamente al 10-15% de las pacientes con cáncer de mama². Se define por la ausencia de expresión del receptor 2 del factor de crecimiento epidérmico humano ("HER2"), del receptor de estrógeno ("RE") y del receptor de progesterona ("RP"). Como resultado, los tratamientos dirigidos a HER2 o la terapia endocrina son ineficaces contra el CMTN. Dado que las pacientes carecen hoy en día de terapias dirigidas, el CMTN suele asociarse a un pronóstico desalentador y a un alto riesgo de recaída³ y metástasis.

El tratamiento estándar actual para el CMTN no metastásico es la quimioterapia neoadyuvante ("QTN"), un régimen de tratamiento que suele durar 6 meses y cuyo objetivo es reducir el tamaño del tumor para permitir su extirpación quirúrgica.

La eficacia de la quimioterapia neoadyuvante para el CMTN no metastásico no se conoce bien y varía significativamente entre pacientes. Para aquellas pacientes que no responden a la quimioterapia neoadyuvante, esto significa 6 meses de tratamiento agresivo, con toda la carga física, mental y emocional que ello conlleva, y los efectos secundarios, incluida la cardiotoxicidad.

Predecir en el diagnóstico cómo responderán las pacientes podría influir significativamente en el curso del tratamiento. Por ejemplo, las pacientes que respondan bien podrían recibir una quimioterapia más ligera para reducir los riesgos de cardiotoxicidad. Aquellas con respuesta deficiente podrían empezar inmediatamente con agentes quimioterapéuticos alternativos o dirigirse a nuevos tratamientos en fase de ensayo clínico, como inmunoterapias, terapias celulares y otros tratamientos dirigidos.



El tratamiento estándar actual para el CMTN no metastásico es la quimioterapia neoadyuvante ("QTN")

¹ Cancer Stat Facts: **Female Breast Cancer, SEER Cancer Statistics Review (CSR) 1975-2018**. Bethesda, MD: National Cancer Institute.

² American Cancer Society. **Breast Cancer Facts & Figures 2017-2018**. Atlanta: American Cancer Society, Inc. 2017.

³ Portha, H. et al. **Non-metastatic triple-negative breast cancer in 2016: Definitions and management**. Gynecologie, obstetrique & fertilité 44, 492-504 (2016).

RETO

Aunque los clínicos puedan aproximar la respuesta del QTN inspeccionando manualmente las muestras de la biopsia y anotando manualmente las características celulares, este es un método heurístico informal y que depende directamente del profesional.

La tecnología de *Machine Learning* puede utilizarse para descubrir nuevos biomarcadores a partir de una recogida de muestras completa. Estos biomarcadores pueden utilizarse a su vez para clasificar los tumores en diferentes subgrupos y predecir su respuesta a tratamientos específicos, por ejemplo. Esto ya se ha utilizado con éxito en el diagnóstico, la clasificación y el pronóstico del cáncer de mama, por ejemplo^{4,5,6}. Incluso se ha demostrado que detecta de forma robusta características histológicas importantes para estimar la respuesta al QTN, como la Carga de Cáncer Residual y el porcentaje de linfocitos-T⁷ infiltrados en el tumor. Por lo tanto, este enfoque puede ser muy potente para dilucidar los factores que impulsan la heterogeneidad de la respuesta al QTN y mejorar nuestra comprensión sobre qué pacientes responderán al tratamiento.

Sin embargo, el *Machine Learning* es un método computacional que requiere muchos datos y, debido a su pobre prevalencia, los datos del CMTN son escasos. Además, no están disponibles en una única base de datos, sino que están distribuidos por los centros oncológicos, a menudo en volúmenes bajos inferiores a 200 pacientes.

Por lo tanto, el reto es acceder a suficientes datos de alta calidad para llevar a cabo un estudio sólido que permita dilucidar y potencialmente validar nuevos biomarcadores para la predicción de la respuesta al CMTN. Con esto, esperamos poder mejorar nuestra comprensión del CMTN y, en última instancia, la atención a los pacientes.



Machine Learning es un método computacional que requiere muchos datos

⁴ Couture, H.D., Williams, L.A., Geradts, J., Nyante, S.J., Butler, E.N., Marron, J.S., Perou, C.M., Troester, M.A. and Niethammer, M., 2018. **Image analysis with deep learning to predict breast cancer grade, ER status, histologic subtype, and intrinsic subtype.** NPJ breast cancer, 4(1), pp1-8.

⁵ Turkkı, R., Bychkov, D., Lundin, M., Isola, J., Nordling, S., Kovanen, P.E., Verrill, C., von Smitten, K., Joensuu, H., Lundin, J. and Linder, N., 2019. **Breast cancer outcome prediction with tumour tissue images and machine learning.** Breast cancer research and treatment, 177(1), pp41-52.

⁶ Naik, N., Madani, A., Esteva, A., Keskar, N.S., Press, M.F., Ruderman, D., Agus, D.B. and Socher, R., 2020. **Deep learning-enabled breast cancer hormonal receptor status determination from base-level H&E stains.** Nature communications, 11(1), pp.1-8.

⁷ Naylor, P., Boyd, J., Laé, M., Reyat, F. and Walter, T., 2019, April. **Predicting Residual Cancer Burden in a triple negative breast cancer cohort.** In 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019) (pp. 933-937). IEEE.

SOLUCIÓN

Dos de los principales centros oncológicos de Francia, el Centro Léon Bérard y el Instituto Curie, tienen sus propios datos de CMTN de alta calidad, listos para ser utilizados con *Machine Learning*. Entre los dos centros, se combinan datos de 529 pacientes de CMTN, una cantidad crítica para permitir un estudio sólido. Sin embargo, la extracción de los datos de estas instituciones es imposible debido a las preocupaciones en materia de privacidad de los pacientes y las regulaciones del LOPD.

En un proyecto puntero de colaboración (du Terrail et al. 2021⁸), los dos hospitales se han asociado con la empresa francesa de IA médica Owkin para abordar este complejo reto multidimensional, aplicando el *Machine Learning* Federado (Aprendizaje Federado) en el conjunto de datos dividido entre los dos hospitales.

El *Machine Learning* tradicional requiere que los datos estén centralizados en un único servidor donde se puedan ejecutar los algoritmos. El Aprendizaje Federado, sin embargo, es un nuevo paradigma de aprendizaje que permite realizar estudios multicéntricos sin agrupar los datos. Estos algoritmos de Aprendizaje Federado pueden accionarse sobre datos distribuidos no centralizados, compartiendo únicamente los parámetros del algoritmo y las ponderaciones del modelo entre instituciones, en lugar de sobre los propios datos (algo cada vez más difícil debido a la preocupación por la privacidad, la normativa federal, las limitaciones comerciales y el tamaño de los datos).

En esta colaboración, los algoritmos de *Machine Learning* de Owkin se enviaron detrás de los cortafuegos de cada hospital para entrenar los modelos locales individualmente en cada conjunto de datos⁹. A continuación, los parámetros del modelo se agregan sin extraer ningún dato confidencial o de identificación personal y el aprendizaje combinado de cada institución se integra en un modelo nuevo y actualizado. Este ciclo continúa hasta que se completan las tareas de entrenamiento y se utilizan todos los datos para entrenar y probar el modelo.

CONCLUSIÓN

En primer lugar, la colaboración demostró que cada modelo de *Machine Learning* entrenado con los conjuntos de datos de un hospital individualmente era capaz de predecir la respuesta al QTN con la misma precisión que los expertos clínicos, quienes dependen de anotaciones manuales que requieren de mucho tiempo¹⁰.

Este hallazgo es fundamental para generar confianza en el posible uso futuro de estos algoritmos informáticos en la práctica clínica habitual.

En segundo lugar, los investigadores descubrieron que el modelo colaborativo, entrenado mediante Aprendizaje Federado, mejora el rendimiento del modelo de *Machine Learning* y da lugar a predicciones de QTN que superan los métodos clínicos actuales en ambas instituciones¹¹. Aunque la mejora varía en ambos entornos, la fertilización cruzada de los aprendizajes de cada modelo local proporciona mejoras globales, lo que hace que este enfoque de *Machine Learning* sea más generalizable. Este crucial hallazgo confirma aún más la importancia de la colaboración interinstitucional y la forma en que los métodos tecnológicos pueden resolver los retos que ofrece el intercambio de datos.

En conclusión, la aplicación del Aprendizaje Federado al conjunto de muestras de biopsias de pacientes con CMTN ha proporcionado nuevos conocimientos sobre los impulsores de la respuesta al QTN, como la potencial importancia de la fibrosis y de las células tumorales discohesivas como biomarcadores, para seguir investigando. También ha producido un algoritmo más robusto y eficaz para ayudar a los clínicos expertos en su evaluación de las pacientes con CMTN en el momento del diagnóstico. Aunque todavía está en fase de investigación, este enfoque es prometedor para las pacientes de CMTN y los médicos que las tratan, y se espera que atraiga a más colaboradores que contribuyan con más datos de alta calidad a esta iniciativa de colaboración única.

⁸ du Terrail, J.O., Leopold, A., Joly, C., Beguier, C., Andreux, M., Maussion, C., Schmauch, B., Tramel, E.W., Bendjebbar, E., Zaslavskiy, M. and Wainrib, G., 2021. Collaborative Federated Learning behind Hospitals' Firewalls for Predicting Histological Response to Neoadjuvant Chemotherapy in Triple-Negative Breast Cancer. medRxiv.

⁹ Federated Learning - Owkin. 2022. Owkin, Inc.; [accessed 10/12/2021]. <https://owkin.com/federated-learning/>

¹⁰ du Terrail, J.O., Leopold, A., Joly, C., Beguier, C., Andreux, M., Maussion, C., Schmauch, B., Tramel, E.W., Bendjebbar, E., Zaslavskiy, M. and Wainrib, G., 2021. Collaborative Federated Learning behind Hospitals' Firewalls for Predicting Histological Response to Neoadjuvant Chemotherapy in Triple-Negative Breast Cancer. medRxiv.

¹¹ du Terrail, J.O., Leopold, A., Joly, C., Beguier, C., Andreux, M., Maussion, C., Schmauch, B., Tramel, E.W., Bendjebbar, E., Zaslavskiy, M. and Wainrib, G., 2021. Collaborative Federated Learning behind Hospitals' Firewalls for Predicting Histological Response to Neoadjuvant Chemotherapy in Triple-Negative Breast Cancer. medRxiv.