

Description of Additional Supplementary Files

File name: Supplementary Data 1

File description: Clinical Characteristics of All Cases and Hormone Receptor-Positive Samples Across Datasets. For each test cohort and variable, we present the number and percentage of (1) all slides, (2) hormone receptor positive (HR+) slides according to the pathology reports, and (3) those identified as HR+ by our system. The system identified hormone receptor-positive cases in each of the categories, except for the hormone receptor-negative and ER-negative categories. Although there is a tendency for lower grades in hormone receptor-positive patients, high-grade tumors were also accurately identified as hormone receptor-positive. A smaller proportion of ERBB2-positive cases were found in the identified HR+ group compared to the HR+ group in pathology reports, indicating a lower likelihood of the system identifying hormone receptor-positiveness within ERBB2-positive tumors than within ERBB2-negative tumors. Counts for some clinical variables do not sum to 100% because their missing values are not listed. ER-low-positive cases are defined as $(1\% \leq \text{IHC} \leq 10\%)$, and ER-high-positive cases are defined as $10\% < \text{IHC}$. Abbreviations: IDC (Invasive Ductal Carcinoma), ILC (Invasive Lobular Carcinoma), CNB (Core Needle Biopsy), EXC (Excisional Biopsy). Blank cells represent missing variables in the data.

File name: Supplementary Data 2

File description: Source data of Fig. 2a

File name: Supplementary Data 3

File description: Source data of Fig. 4

File name: Supplementary Data 4

File description: Source data of Fig. 6