Appendix E – Differences between FDA and IceCure Medical's Analyses

This Appendix details the differences between IceCure Medical and FDA's analyses in calculating the results of the ICE3 study and in conducting Systematic Literature Reviews for comparison to the results of the ICE3 study.

Differences in the ICE3 study results analyses

Number of patients included in the analysis population

The trial screened 212 patients, 206 of which were treated with cryoablation using the ProSense System. Of these, 12 patients were subsequently excluded from the study by the Data Safety Monitoring Board (DSMB) due to inclusion/exclusion criteria violations (N=9) or incomplete treatment (N=3), and their participation withdrawn prior to completing the study. These 12 patients were not included in the Primary Analysis Population used by IceCure Medical to calculate the primary endpoint (IBTR rate). Thus, the Primary Analysis Population contained 194 subjects. The patient disposition including the Primary Analysis Population is summarized in **Figure 4**.

Details of the 12 patients excluded by the DSMB as well as all other patients with major inclusion/exclusion criteria deviations are included in **Appendix D**. Seven of the 12 excluded subjects had lesions >1.5 cm in the largest dimension, one had prior lumpectomy and radiation in addition to baseline multifocal tumor, and one had DCIS 40% on baseline pathology. There were 44 total patients with inclusion/exclusion criteria deviations and 4 patients with inadequate treatment duration; however, only 12 were excluded from the Primary Analysis Set.

FDA includes these 12 patients in an analysis presented in the main Executive Summary (Full Analysis Set) because results from all treated patients account for deviations or non-compliance that may occur during real-world use of the treatment. Of these 12 patients, 5 had local recurrences.

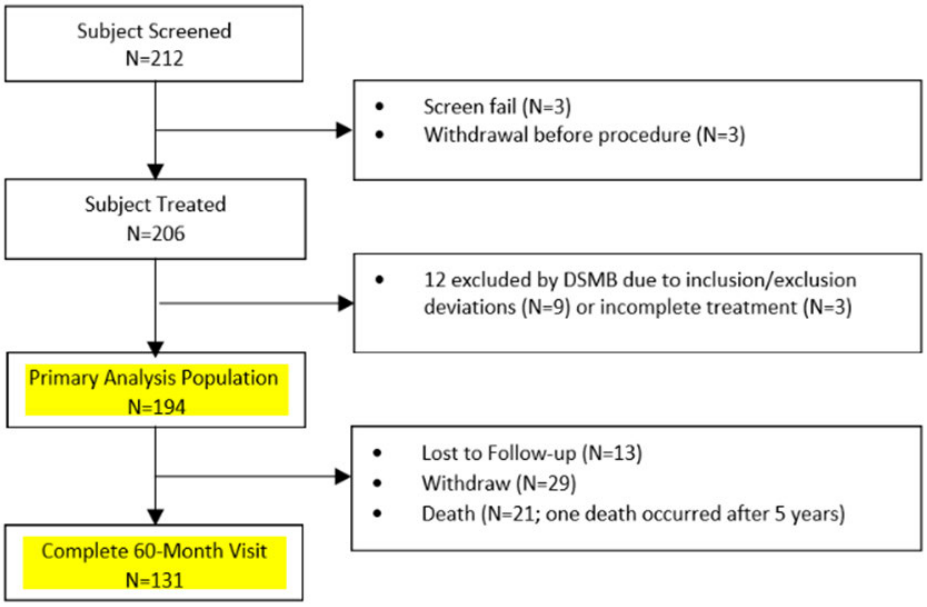

Figure 4. ICE3 patient disposition flowchart including the Primary Analysis Population.

Classification of recurrence
FDA included two additional cases of recurrence (9 total) compared to IceCure Medical (7 total) due to the following:

- One subject (b) (6) had a distant metastasis at approximately 37 months. The patient's CRF records a new breast lesion at 12 o'clock, 5 cm from the nipple, considered as separate from the index breast tumor treated with cryotherapy in the ICE3 trial. Per the DSMB determination, IceCure reported this event as a second primary breast cancer based upon location. Given that the new lesion was close to the index tumor and of the same histopathology and tumor receptor status as the index tumor, the FDA classified this lesion as an IBTR according to the protocol definition (Section 5.8.1 Local Recurrence): "*Local recurrence is defined as evidence of invasive or in situ breast cancer in the ipsilateral breast or chest wall.*"
- One subject (b) (6) had a report in the CRF at 60 month follow up that the investigator's review of the mammogram identified a density adjacent to the cryoablation site concerning for local recurrence. The investigator recommended diagnostic workup and possible biopsy of this lesion. However, the patient refused, declining any further workup. Given that the new lesion was adjacent to the cryoablation site and concerning for recurrent cancer, FDA classified this lesion as an IBTR according to the protocol definition (Section 5.8.1 Local Recurrence): "*Local recurrence is defined as evidence of invasive or in situ breast cancer in the ipsilateral breast or chest wall.*"

Kaplan-Meier (KM) method versus Cumulative Incidence Function (CIF)
FDA presented primary endpoint results from both the Kaplan-Meier (KM) method and Cumulative Incidence Function (CIF) method. The KM method was the pre-specified method in the ICE3 protocol. However, KM treats death as censored (same as lost-to-follow-up) which implies local recurrence could occur after death. In contrast, CIF considers local recurrence could not occur after death in the determination of IBTR rate.

Differences in determination of event time
Methods for determining event time in the Kaplan-Meier calculations for patients with events occurring at time points >60 months were not pre-specified in the ICE3 protocol. In the ICE3 study, four events included in IceCure Medical's KM calculation took place after the 60 months. For example, subject (b) (6) had a local recurrence found during the 60-month visit, but the actual visit date of the 60-month visit was 63.19 months. FDA used 60 months as the event time for such cases because the study is estimating a 5-year rate, and presumably the event occurred within the 5-year study period. IceCure Medical used the middle point between the 48-month visit and the 60-month visit as the event time to reflect that the event occurred at some point prior to the 60-month visit date. FDA notes that IceCure Medical's method leads to slightly lower event rates in the KM calculation. FDA and IceCure Medical's event times for all four events occurring after 60 months are detailed below in **Table 16**.

Table 16. Subjects with event times beyond 60 months and the event times used by IceCure Medical in the KM calculation of IBTR rate. FDA used 60 months as the event date for all four patients in the KM calculation of 5-year rates. Dates are presented in format YYYY/MM/DD.

Subject ID	Event	Procedure date	48-month visit	60-month visit	Event date	Months between procedure and 60-month visit	Months between procedure and event date	Event time used by IceCure (months)
(b) (6)	Local recurrence	2018/10/02	2023/01/09	2024/01/08	2024/01/10	63.19	63.26	57.5867 7686
(b) (6)	Second primary BC	2015/02/06	2019/03/01	2020/03/04	2020/03/04	60.93	60.93	55.1570 2479
(b) (6)	Second primary non-BC	2018/04/25	2022/06/08	2023/05/26	2023/05/26	61.03	61.03	55.5702 4793
(b) (6)	Second primary non-BC	2015/08/26	2019/08/27	2020/09/08	2020/09/08	60.42	60.42	54.5785 124

Differences in determination of censoring times

Censoring is common in clinical trials when observation of the patient is terminated before an event occurs, such as when a subject withdraws from the study or the patient is lost to follow-up. Censoring subjects in the Kaplan-Meier calculation removes the subject from the denominator (i.e., the number of individuals still at risk). Methods for determining censoring time in the Kaplan-Meier calculations were not pre-specified in the ICE3 protocol. In FDA's KM calculation, the censoring time is considered to be the last date of a subject being event-free, because the exact event time is not known, but we do know that it is greater than the last date of being event-free. IceCure Medical's approach to determining censoring time for subjects without an event of interest is described in **Table 17**. IceCure Medical considers subjects to be within the pool of "subjects at risk" until they complete an end of study form or until an event of interest occurs. For example, when calculating the IBTR rate for subjects without a local recurrence:

- FDA considers subjects with death (19 subjects) to be censored at the date of the last follow-up before death; however, IceCure Medical considers these subjects to be censored at 60 months. For example, subject (b) (6) died about 27.45 months after the procedure. FDA considers the subject to be censored at 11.97 months when the last follow-up occurred at 12-month visit, whereas IceCure Medical considered this subject to be censored at 60 months.
- FDA considers that subjects lost-to-follow-up (LTFU, 13 subjects) were censored at the date of the last follow-up before LTFU; however, IceCure Medical considers these subjects to be censored at the date on the End of Study form. For example, subject (b) (6) had the last follow-up at the 12-month visit and was LTFU thereafter. FDA considers the censor date to be 11.19 months, when the last follow-up occurred, whereas IceCure Medical considered this subject to be censored at 60 months.

- FDA considers that subjects who withdrew from the study (26 subjects) were censored at the date of the last follow-up before withdrawal; however, IceCure Medical considers these subjects to be censored at the date on the End of Study form. For example, subject (b) (6) had the last follow-up at the 6-month visit. FDA considers the censored date to be 6.13 months, when the last follow-up occurred, whereas IceCure Medical considered this subject to be censored at 8.4 months.

FDA notes that IceCure Medical's approach to determining censoring time when using the KM method to estimate 5-year rates for local recurrence, "distant metastases", and DFS does not account for the fact that an event of interest could occur any time between the last follow-up and IceCure Medical's censoring time. IceCure Medical's approach keeps more subjects "at risk" longer in the KM calculation, and will therefore increase the reported event-free rate, such as DFS, and lower event rates, such as local recurrence and "distant metastases" rates.

Table 17. IceCure Medical's determination of censoring time for subjects without an event of interest in the KM calculation.

Outcome	Subjects who withdrew or were lost to follow up (LTFU) were censored at	Subjects who died were censored at
IBTR	date on End of Study form or 60 months*	60 months
"Distant Metastases"	date on End of Study form or 60 months*	60 months
DFS (protocol definition)	date on End of Study form or 60 months*	N/A (death considered an event)
DFS (NCI definition interpreted by IceCure Medical)	date on End of Study form or 60 months*	60 months

* Six subjects who withdrew or were LTFU were censored at 60 months, as their dates on the End of Study form exceeded 60 months.

Combined impact of KM calculation differences

Due to the above-described differences in the handling of event times and censoring methods in the KM calculations, even for the same number of events, FDA and IceCure Medical's KM calculations result in different event rates. FDA's calculation of IBTR rate using IceCure Medical's count of 7 patients with recurrence (out of 194 in the Primary Analysis Set) results in a rate of 5.2% (95% CI 2.5-10.7%) by the KM method whereas IceCure Medical reports an IBTR rate of 4.3% (2.1-8.7%).

Disease-Free Survival definition

The definition of Disease-Free Survival (DFS) has been inconsistent across breast cancer trials in the literature, with different events included or excluded as disease events. This makes it difficult to compare the results of the ICE3 study with different trials in the literature. The ICE3 protocol defines DFS as the time from the date of complete ablation of the primary tumor until the first disease event, where the disease event is defined as local (DCIS or invasive), regional, or distant breast cancer recurrence, second primary cancer, DCIS or invasive contralateral breast cancer, or death due to any cause. FDA presents the results of DFS using this definition in the main Executive Summary.

After the study was completed, IceCure Medical reanalyzed the DFS based on the National Cancer Institute (NCI) definition of DFS as the length of time that the patient survives after completing primary treatment for a cancer without any signs or symptoms of that cancer. IceCure Medical's interpretation of the NCI definition for DFS excluded second primary non-breast cancer events and deaths not due to BC from disease events but censored deaths due to other causes at 60 months. As recognized in FDA guidance, Clinical Trial Endpoints for the Approval of Cancer Drugs and Biologics, *"this method can introduce bias in the attribution of the cause of death."* FDA typically considers deaths from all causes included in the definition of DFS in order to minimize bias. In addition, IceCure Medical censored those subjects who died due to other causes at 60 months in the Kaplan-Meier calculations as detailed in "Differences in determination of censoring times" above.

Determination of the Indicated and LUMINA-aligned Subpopulations

IceCure Medical conducted multiple additional subpopulation analyses compared with FDA. FDA analyzed only two subpopulations as defined in the main Executive Summary: Indicated Subpopulation (N=120) and LUMINA-aligned Subpopulation (N=48). The primary difference between the primary analysis set of the ICE3 study and these subpopulations was the exclusion of patients who did not receive hormone therapy, and the exclusion of patients with a nuclear score of 3 or unreported Nottingham Grade sub-scores. The LUMINA-aligned subpopulation further excluded patients who received radiation therapy and without Ki-67<14%. IceCure Medical performed additional subpopulation analyses that did not exclude patients due to Nottingham grade sub-score information. We acknowledge that the proposed IFU and the LUMINA study do not specify criteria for the Nottingham sub-score components. However, nuclear and mitotic score criteria (specifically, nuclear and mitotic scores must be ≤2) were used in the definition of the subpopulations because these criteria were defined in the inclusion/exclusion criteria of the ICE3 study. Although some patients were enrolled in the ICE3 study in violation of the protocol criteria, these patients were excluded from the subpopulations due to their protocol deviations. FDA was unable to independently confirm the nuclear grade of 19 patients whose nuclear grade was not reported. Of those with reported values, FDA identified 12 patients with nuclear grade 3 or 2-3. For some of these patients, the nuclear grade inclusion criterion was not required at the time of enrollment due to modifications to the study protocol.

Differences in Systematic Literature Review Methodology

The information presented in this section highlights the differences in methodology between FDA's Systematic Literature Review (SLR) and IceCure Medical's SLR, which were used to estimate an IBTR rate for patients similar to the intended patient population defined in the proposed Indications for Use. For full details on the methodology used in FDA's Systematic Literature Review (SLR), please refer to **Appendix G**.

Both FDA's and IceCure Medical's SLRs adhered to PRISMA guidelines for conducting and reporting systematic reviews. For bias assessment, both used validated tools: the FDA SLR employed the Cochrane Risk of Bias 2 tool for RCTs and ROBINS-I for observational studies, while the IceCure Medical's SLR used a custom data appraisal tool. Both SLRs also used random effects meta-analysis models to account for between-study heterogeneity. Both reviews focused on IBTR at 5 years as a primary outcome. IceCure Medical defined IBTR as recurrent in situ or

invasive carcinoma in the ipsilateral breast without clinical-radiologic evidence of regional or distant disease.

However, due to significant differences in search strategy and article selection for the meta-analysis, FDA identified a different set of relevant papers than IceCure Medical in the overall search and in the final meta-estimate. IceCure Medical's SLR included 12 studies, of which 11 studies were included in the meta-estimate. FDA's SLR included 25 studies from which qualitative information was drawn, while only 5 studies were ultimately used for the meta-analysis. These differences are detailed further below.

Search Terms and Strategies

IceCure Medical's SLR and the FDA's SLR utilized different search strategies and terms, which impacted the breadth and specificity of the literature captured. This variance in search methodology is crucial as it directly influences the comprehensiveness and relevance of the systematic review results. FDA searched PubMed/MEDLINE and Embase using targeted search strings. IceCure Medical's search included PubMed, Ovid/Medline, Embase, and ClinicalTrials.gov. Indexing in databases like PubMed and Embase relies on controlled vocabulary terms (like MeSH in PubMed) to categorize studies. These terms help in retrieving literature that is accurately tagged by subject area, methodology, and focus. However, not all studies are indexed uniformly or comprehensively. Discrepancies in how a study is indexed (e.g., missing relevant terms or inconsistent use of synonyms) can lead to its omission from search results if the search terms used are too narrow or not aligned with the indexing terms.

IceCure Medical's search terms, while comprehensive, were potentially limited by not incorporating a wider range of synonymous terms and controlled vocabulary that could capture more specific studies relevant to their criteria. For example, using broad terms like "Early Stage" and "Breast Cancer" might miss studies specifically indexed under sub-categories like "Stage I" or "Low-Risk Breast Cancer." Moreover, not including specific outcome measures like "ipsilateral breast tumor recurrence" or detailed pathological criteria (e.g., "Nottingham Grade" or "Ki-67") in the search could lead to excluding studies that focus precisely on those aspects but are indexed under these specific terms. FDA's search strategy incorporated a diverse range of terms and controlled vocabulary to minimize the risk of missing studies due to indexing limitations. This approach is particularly important when developing meta-estimates for performance goals, as missing even a small subset of relevant studies can skew the results.

Please find FDA's PRISMA flow chart **(Figure 5)** and IceCure Medical's initial PRISMA flow chart (**Figure 6**) below.

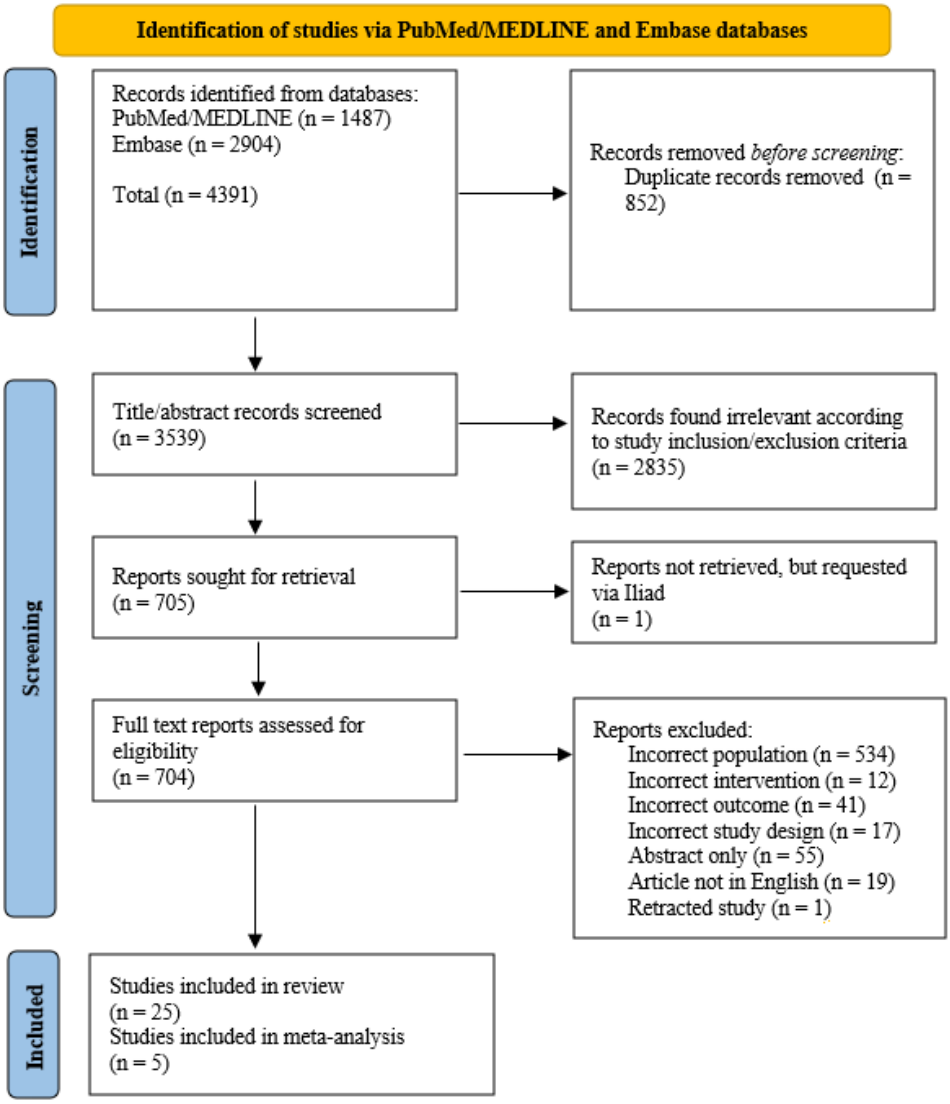

Figure 5. PRISMA flow diagram from FDA's SLR.

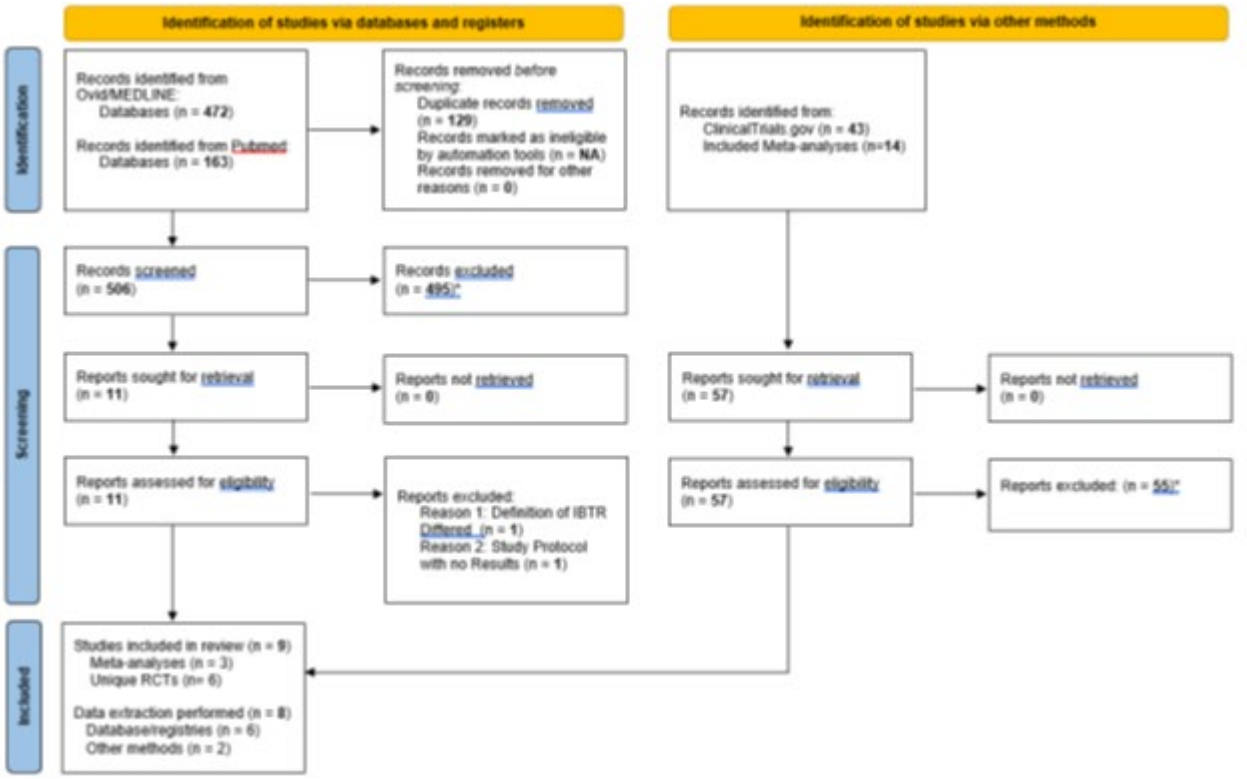

Figure 6. PRISMA flow diagram from IceCure Medical's Initial SLR.
Please note the number of "Identified" studies was later updated with 113 additional studies (100 additional studies identified via search of databases and 13 additional studies identified via ClinicalTrials.gov and 3 of these new studies were included in the SLR after screening for a total of 12 studies.)

Study Selection

IceCure Medical's SLR and the FDA's SLR used different approaches to study selection. FDA's SLR used stringent inclusion/exclusion criteria focusing on a specific low-risk population, particularly in the meta-analysis. FDA's approach targeted patients aged >50 years with specific breast cancer characteristics (tumor size ≤1.5-2 cm, Nottingham grade 1-2, ER+, PR+, HER2-, Ki-67<14%, clinically node-negative). Literature study populations were matched as closely as possible for those articles selected into the meta-analysis with 98-100% alignment. This is a significant difference compared with IceCure Medical's SLR methodology, which pre-specified a requirement for only 75% alignment with the meta-analysis inclusion/exclusion criteria. IceCure Medical weighted articles in the meta-analysis according to their relative alignment with the inclusion/exclusion criteria. However, weighting based on a relative match can bias the results of a meta-estimate by incorrectly attributing the observed events to a misaligned population.

A comparison of patient characteristics between the SLRs and the ICE3 study are presented in **Table 18** and discussed in further detail below. The criteria described in the table for the SLRs are representative rather than absolute.

Patient characteristics

We note that the patient age target of >50 years in FDA's SLR is younger than the proposed IFU (≥60 years) and the ICE3 enrollment criteria (age ≥50 (local IRB); age ≥ 60 (WCG IRB)). We note that the patient age target in FDA's SLR of >50 years is younger than the proposed IFU (≥60 years) and the average age of the ICE3 population (mean age 74.9 ± 6.9 years; median age 74.5 years). We also note that Ki-67 is not included as a criterion in the IFU and was removed from the enrollment criteria of the ICE3 protocol during the study. Thus, Ki-67 was not used as an inclusion criterion in FDA's SLR, but where Ki-67 was reported to be >14%, these studies were excluded from the meta-analysis to facilitate an estimate more representative of low-risk patients.

IceCure Medical's strategy included patients ≥50 years old and defined low-risk, early-stage breast cancer as T1 (tumor size <2 cm), node-negative (N0), local (M0), ER/PR positive, and HER2 negative. We note that some patients were included with relatively higher risk factors for recurrence than the indicated population, such as lobular carcinoma, high tumor grade, multifocal tumors, and lymphovascular invasion, which could inflate the IBTR rate.

The inclusion of younger patients relative to the proposed IFU of the ProSense System in both FDA and IceCure Medical's SLRs may impact recurrence rate. Some studies have shown that older patients may have a relatively lower risk of recurrence due to a likelihood of death by competing risks and the fact that older patients tend to present with less aggressive disease.[44]

Adjuvant Therapies
A key difference between FDA and IceCure Medical's SLRs was the inclusion of articles with use of specific adjuvant therapies. Both FDA and IceCure Medical included use of adjuvant radiotherapy, per the proposed IFU of the ProSense System. However, as in the proposed IFU, FDA had no requirements related to the use of adjuvant radiotherapy, while IceCure Medical explicitly excluded use of adjuvant radiotherapy. IceCure Medical's inclusion criteria specified BCS without adjunctive radiation, allowing for other adjunctive treatments such as endocrine therapy or chemotherapy. They excluded trials where all patients received radiation therapy or mastectomy.

FDA finds the exclusion of adjuvant radiotherapy from IceCure Medical's SLR to be a key limitation due to the following reasons:

1) BCS followed by radiotherapy is the standard of care for women with small breast cancers who wish to avoid mastectomy.
2) Use of radiotherapy achieves good local control of disease. Randomized, controlled, peer-reviewed studies, e.g., CALGB 9343 and PRIME II, concluded that omission of the use of adjuvant radiotherapy was associated with an increased incidence of local recurrence.
3) Certain guidelines, like the NCCN guidelines, recommend considering radiation omission only for a very select subpopulation of patients. Despite these guidelines, older women (≥70 years of age) with early-stage hormone-receptor–positive breast cancer are still often receiving radiotherapy after breast-conserving surgery. Local recurrence is often treated by mastectomy and associated with considerable psychological effects. In the current era of shared patient–physician decision making, the notion that omission of radiotherapy may result in a higher incidence of local recurrence may be unacceptable to many patients.

Table 18. Comparison of SLR patient characteristics with the ICE3 study enrollment criteria and the indicated subpopulation patient characteristics.

ICE3 Study Enrollment Criteria	ICE3 Study Indicated Subpopulation	FDA SLR	IceCure Medical SLR
Invasive breast cancer (excludes lobular, microinvasion, extensive intraductal component, multifocal, multicentric, multifocal calcifications)	Invasive breast cancer (excludes lobular, microinvasion, extensive intraductal component, multifocal, multicentric, multifocal calcifications) [1]	Invasive breast cancer (excludes lobular, microinvasion, extensive intraductal component, multifocal, multicentric, where noted)	Invasive breast cancer
Age ≥ 50 (Local IRB) Age ≥ 60 (WCG IRB) The mean age was 74.9 years and the median age was 74.5 years.	Age ≥ 60 years	Age >50 years	Age ≥ 50 years
Nottingham grade 1-2; specifically, nuclear and mitotic scores must be ≤2	Nottingham grade 1-2; specifically, nuclear and mitotic scores must be ≤2 [2]	Nottingham grade 1-2	No criteria
Node negative	Node negative	Node negative	Node negative
ER positive PR positive or negative HER2 negative	ER positive and/or PR positive HER2 negative	ER positive PR positive HER2 negative	ER positive PR positive HER2 negative
Note[3]	No criteria	Ki-67 ≥ 14% excluded	No criteria
Tumor size ≤1.5 cm	Tumor size ≤1.5 cm	Tumor size ≤1.5-2 cm[4]	Tumor size <2 cm
Adjuvant therapy not specified as enrollment criteria. In the ICE3 Full Analysis Set, 77% received Hormone therapy; 15% received radiation therapy.	Must receive adjuvant endocrine therapy	Must receive adjuvant endocrine therapy	Must receive adjuvant endocrine therapy
	Radiation therapy not specified in the IFU. 18% received radiation therapy in the indicated subpopulation of ICE3.	No Radiation versus Radiation assessed via sensitivity analysis	Radiation and chemotherapy use not specified; no study included had radiation treatment arms

[1] Because the indicated subpopulation is a subset of the ICE3 study, it uses the same criteria for the breast cancer classification; however, the proposed IFU defines the intended patient population as having *"infiltrating ductal carcinoma (excluding lobular carcinoma, extensive intraductal component, or evidence of lymphovascular invasion)."* Lymphovascular invasion was not part of the ICE3 protocol exclusion criteria but is part of the IFU. Multifocal/multicentric disease and multifocal calcifications are part of the ICE3 criteria and not part of the IFU.

[2] The proposed IFU does not specify Nottingham sub-score components. However, the criterion in the table related to nuclear and mitotic scores was used in the definition of the intended subpopulation because this was defined in the inclusion/exclusion criteria of the ICE3 study. Although some patients were enrolled in the ICE3 study in violation of the protocol criteria, these patients were requested to be excluded from the subpopulation analysis due to their protocol deviations.

[3] Ki-67 was initially defined as an inclusion criterion in the ICE3 protocol but was later removed.

[4] Given that T1 staging is defined by tumor size less than 2 cm, articles in which T1 was used to describe the patient population rather than specific tumor size may include tumors up to 2 cm in size.