

AI-PLAN: Artificial Intelligence in Oncoplastic and Reconstructive Breast Surgery, From Guideline-Based Decisions to Personalized Surgical Planning

1. Introduction

Breast cancer represents the most frequently diagnosed malignancy among women and remains a major cause of cancer-related morbidity and mortality [1,2]. Over the past decades, improvements in screening, imaging, pathological and molecular characterization, systemic treatments, and radiotherapy have substantially changed both the prognosis and therapeutic management of the disease [2]. As survival has improved, increasing attention has progressively shifted from oncological outcomes alone toward treatment-related morbidity, quality of life, body image, and long-term functional and aesthetic results.

Breast surgery has consequently undergone a profound evolution. Radical procedures, historically considered necessary to achieve adequate local control, have progressively been replaced by more conservative and individualized approaches. Long-term randomized evidence demonstrated equivalent survival between mastectomy and breast-conserving surgery followed by radiotherapy in appropriately selected patients [3,4]. Breast-conserving surgery has therefore become a cornerstone of early-stage breast cancer treatment, while the development of oncoplastic techniques has further expanded its indications by allowing wider resections while preserving breast shape and symmetry [5].

At the same time, advances in reconstructive surgery have increased the number of patients eligible for immediate reconstruction following mastectomy and have broadened the range of available techniques. Contemporary breast surgical oncology therefore extends far beyond the simple distinction between breast-conserving surgery and mastectomy. Surgical treatment is increasingly

conceived as an individualized process aimed at achieving oncological radicality while minimizing morbidity and optimizing functional, aesthetic, and psychological outcomes.

Modern surgical planning requires the integration of multiple oncological, anatomical, technical, and patient-related variables. Tumor size and location, multifocality or multicentricity, biological subtype, clinical nodal status, response to neoadjuvant systemic therapy, breast volume, degree of ptosis, skin and subcutaneous tissue characteristics, previous radiotherapy, comorbidities, genetic predisposition, and patient preferences may all influence the final surgical strategy [2,5]. Two patients with apparently similar oncological characteristics may therefore require substantially different surgical approaches because of differences in tumor-to-breast volume ratio, breast anatomy, reconstructive possibilities, previous treatments, or individual expectations.

This complexity is particularly evident in oncoplastic and reconstructive breast surgery. The choice between breast-conserving surgery and mastectomy represents only the first step of a broader decision-making process. When breast conservation is feasible, the surgeon must determine the extent of resection, the most appropriate incision and the potential need for volume displacement or volume replacement techniques. Conversely, when mastectomy is indicated, additional decisions concern skin or nipple preservation, timing and type of reconstruction, implant positioning, or the potential use of autologous tissue. Each of these choices requires a balance between oncological safety, technical feasibility, anticipated aesthetic outcome, and patient preference.

The increasing complexity of breast cancer care has reinforced the central role of multidisciplinary management. Dedicated Breast Units integrate the expertise of breast surgeons, medical oncologists, radiologists, radiation oncologists, pathologists, geneticists, plastic surgeons, and specialized nurses in order to provide evidence-based and individualized treatment [6]. Nevertheless, even within multidisciplinary pathways, a certain degree of variability may persist across institutions and among surgeons, particularly when more than one oncologically acceptable surgical strategy is available.

Such variability may reflect differences in institutional protocols, surgical expertise, reconstructive resources, interpretation of patient-specific factors, and individual preferences.

In this scenario, artificial intelligence (AI) has emerged as a rapidly developing field with the potential to assist clinicians in processing increasingly complex and heterogeneous information. AI encompasses a wide range of computational methodologies capable of identifying patterns, integrating multiple variables, and generating predictions or structured outputs. Within breast cancer care, AI-based systems have been extensively investigated for imaging, pathological classification, prognosis, risk prediction, and treatment stratification [7,16]. Their potential applications have subsequently expanded toward anatomical assessment, preoperative planning, breast symmetry evaluation, reconstruction, and prediction of postoperative outcomes [7-9].

More recently, the development of generative AI and large language models (LLMs) has introduced a different paradigm of AI-assisted medicine. LLMs are computational models trained on extensive text corpora that are capable of processing natural language, integrating information from different domains, and generating context-dependent textual outputs. Rather than being designed exclusively for one predefined predictive task, these models can respond to complex prompts incorporating multiple clinical variables and can generate structured recommendations, explanations, or summaries [10,11].

Their potential role in healthcare extends across several domains, including medical documentation, information synthesis, education, patient communication, and clinical decision support [10,11]. However, LLM-generated outputs are inherently dependent on the information provided in the prompt, the characteristics of the underlying model, and the data on which the model was trained. In addition, inaccurate or fabricated information, lack of reproducibility, bias, and limited transparency remain relevant concerns. Consequently, clinical supervision and independent validation remain essential whenever such models are applied to medical decision-making [10-12].

The surgical field represents a particularly challenging environment for the implementation of LLMs. Surgical decision-making frequently combines evidence-based recommendations with technical considerations, anatomical assessment, institutional experience, and surgeon judgment. Recent literature has therefore emphasized that the role of LLMs in surgery should primarily be considered supportive, potentially assisting with preoperative planning, information synthesis, education, documentation, and selected decision-making processes rather than replacing surgeon expertise [11,12].

Preliminary studies have also explored the ability of LLMs to address intraoperative and reconstructive decision-making scenarios in plastic surgery, highlighting both their potential usefulness and the considerable variability in performance depending on the complexity and context of the clinical problem [13]. These findings are particularly relevant to breast surgery, where technical and aesthetic considerations are closely interconnected.

The application of AI to breast surgical planning therefore poses specific challenges. Some therapeutic decisions are predominantly driven by objective and relatively well-codified parameters. Biological subtype, clinical stage, nodal involvement, and response to systemic treatment, for example, play a major role in determining indications for neoadjuvant or adjuvant therapies. In these scenarios, the decision-making process can often be translated into structured clinical information and guideline-based algorithms.

In contrast, the choice of breast surgical procedure, incision, oncoplastic technique, or reconstructive strategy may depend on more nuanced factors, including breast morphology, tissue quality, degree of ptosis, previous treatments, anticipated cosmetic outcome, patient expectations, and surgeon experience. Some of these variables are inherently visual or subjective and may be difficult to completely represent through structured textual information alone.

This distinction between predominantly guideline-driven and experience-dependent decisions is particularly relevant when evaluating AI performance in breast surgery. The clinical value of an AI system should therefore not be assessed merely according to its ability to produce a technically plausible recommendation. More meaningful questions include whether AI-generated recommendations are consistent with those of an experienced multidisciplinary team, whether agreement differs according to the type of clinical decision being considered, and which components of surgical planning remain predominantly dependent on human expertise.

AI-driven technologies in breast and plastic surgery have already been explored for automated symmetry assessment, patient selection, surgical outcome evaluation, reconstructive planning, and clinical decision support [8,9,13,14]. At the same time, the development of explainable AI has highlighted the importance of making algorithmic outputs interpretable and auditable, particularly when they are intended to support decisions with direct implications for patient treatment [15]. Before integration into routine clinical practice, AI-based systems therefore require careful validation, assessment of generalizability, and appropriate clinical governance [12,15,17].

Against this background, the **BREAST AI-PLAN** project was developed as a prompt-driven AI workflow designed to generate structured therapeutic and surgical recommendations from routinely collected clinical information. Patient characteristics, breast anatomy, tumor biology, and clinical staging were integrated into standardized prompts to obtain recommendations regarding neoadjuvant treatment, type of breast surgery, surgical approach, reconstruction, implant strategy, adjuvant treatment, and expected aesthetic outcome.

The original BREAST AI-PLAN study was conducted as a retrospective single-center analysis of 30 women undergoing breast cancer surgery, comparing AI-generated recommendations with real-world multidisciplinary team decisions [18]. The primary analysis demonstrated heterogeneous levels of agreement across the different domains evaluated, with almost perfect concordance for neoadjuvant

treatment decisions and progressively lower agreement for surgical, reconstructive, and aesthetic variables [18]. These findings raised the hypothesis that the performance of AI-assisted decision support may be closely related to the extent to which a clinical decision is based on standardized parameters rather than on individualized anatomical, technical, and subjective considerations.

The present thesis builds upon the BREAST AI-PLAN experience and provides a broader surgical perspective on the interaction between artificial intelligence and individualized decision-making in oncoplastic and reconstructive breast surgery. Rather than considering AI performance as a single measure of accuracy, the study explores its concordance with clinical practice across different phases of the therapeutic pathway, with particular attention to the distinction between guideline-driven oncological decisions and more complex surgical and reconstructive choices.

The principal objective of this thesis is to assess the level of agreement between AI-assisted planning and multidisciplinary clinical decision-making across different components of the breast cancer treatment pathway.

The **secondary objective** is to identify the domains in which AI may provide reproducible and clinically meaningful decision support and, conversely, those in which anatomical complexity, aesthetic considerations, patient preferences, and surgeon expertise remain essential determinants of personalized surgical care.

2. BREAST CANCER AND MODERN SURGICAL MANAGEMENT

2.1 Biological and clinical determinants of surgical treatment

The surgical management of breast cancer is increasingly based on an individualized assessment integrating tumor biology, anatomical characteristics, disease extent, patient-related factors, and the anticipated response to systemic treatment. Although surgery remains a fundamental component of curative treatment for most patients with non-metastatic breast cancer, the type and extent of surgery can no longer be determined solely by tumor size or anatomical stage.

Modern breast cancer classification incorporates both traditional histopathological features and biological biomarkers that provide prognostic and predictive information. Histological subtype, tumor grade, estrogen receptor (ER) and progesterone receptor (PgR) expression, human epidermal growth factor receptor 2 (HER2) status, proliferative activity, and, when available, genomic signatures contribute to defining tumor behavior and treatment sensitivity [2].

From a practical surgical perspective, molecular phenotype is particularly relevant because it influences the sequence of treatment. Hormone receptor-positive/HER2-negative tumors, HER2-positive disease, and triple-negative breast cancer (TNBC) differ substantially in their probability of response to systemic therapy and therefore in the potential role of neoadjuvant treatment before surgery.

The distinction between Luminal A-like and Luminal B-like tumors is mainly based on hormone receptor expression, HER2 status, tumor grade, and proliferative activity. Luminal tumors are generally characterized by ER expression and are treated with endocrine therapy, although chemotherapy may also be indicated according to biological and clinical risk. HER2-positive tumors benefit from HER2-targeted treatment, whereas TNBC lacks expression of ER, PgR, and HER2 and is frequently characterized by a more aggressive clinical behavior and greater sensitivity to chemotherapy.

These biological differences have important implications for surgical planning. In patients with HER2-positive or triple-negative tumors, neoadjuvant systemic therapy is increasingly used even in operable disease because of the high probability of tumor response and because pathological response provides clinically relevant prognostic information. Tumor downstaging may convert patients initially considered unsuitable for breast-conserving surgery into candidates for a conservative approach [2].

Conversely, in selected patients with favorable hormone receptor-positive tumors, upfront surgery remains commonly appropriate. In these cases, the surgical strategy is primarily determined by the

relationship between tumor extent and breast anatomy, while final pathological information may subsequently guide adjuvant systemic treatment.

Tumor size and extent

Tumor size remains one of the main anatomical determinants of surgical treatment, but its relevance must be interpreted in relation to breast volume. The concept of the **tumor-to-breast volume ratio** is more informative for surgical planning than tumor diameter considered in isolation. A 3-cm tumor located within a large breast may be amenable to breast-conserving surgery with an acceptable cosmetic result, whereas a considerably smaller lesion in a small breast may require an oncoplastic procedure or, in selected circumstances, mastectomy.

Tumor location is equally important. Lesions located in the lower pole, medial quadrants, central breast, or near the nipple–areola complex may present greater reconstructive challenges following wide excision. Oncoplastic techniques were developed partly to overcome these limitations by combining oncological resection with immediate breast reshaping [5].

Multifocality and multicentricity may further influence the choice of procedure. Historically, the presence of multiple tumor foci frequently resulted in mastectomy. Contemporary management is more nuanced, and breast conservation may be feasible in selected cases when complete excision with negative margins can be achieved while maintaining an acceptable breast appearance. Therefore, the spatial distribution of disease rather than the mere presence of multiple foci has become increasingly relevant to decision-making.

Tumor biology

The growing integration of tumor biology into treatment planning has progressively modified the relationship between stage and surgery. Two tumors with identical dimensions may require different therapeutic pathways because of differences in subtype, response to systemic therapy, or genetic background.

ER and PgR expression remain fundamental predictive markers for endocrine responsiveness, while HER2 overexpression or amplification identifies patients who may benefit from HER2-directed

therapy. These biological markers are therefore central not only to systemic treatment but also to determining whether systemic therapy should precede surgery.

Ki-67 has traditionally been used as an indicator of cellular proliferation and may contribute to the distinction between tumors with more or less aggressive luminal characteristics. Nevertheless, its reproducibility may be affected by interobserver and methodological variability, and interpretation should always occur in conjunction with other pathological parameters.

Breast anatomy and patient-related factors

Surgical planning is also strongly influenced by anatomical factors that are not captured by conventional oncological staging systems.

Breast volume, shape, degree of ptosis, skin quality, thickness of subcutaneous tissue, and nipple–areola complex position may substantially modify the feasibility of different breast-conserving, oncoplastic, or reconstructive techniques.

The patient's general condition is equally relevant. Age, BMI, smoking status, vascular comorbidities, previous breast or thoracic radiotherapy, previous surgery, and genetic susceptibility may affect both the indication for surgery and the risk of postoperative complications.

Smoking and severe vascular disease may impair tissue perfusion and increase the risk of skin-flap or nipple–areola complex necrosis after mastectomy and reconstruction. Similarly, very high BMI may increase the incidence of wound complications, infection, seroma, reconstructive failure, and anesthetic morbidity.

Previous radiotherapy represents another important determinant of reconstructive strategy because irradiated tissues may demonstrate reduced vascularity, fibrosis, and impaired healing. In this setting, implant-based reconstruction may be associated with a higher risk of complications and capsular contracture, and autologous reconstruction may therefore be considered in selected patients.

Genetic predisposition

Pathogenic variants in breast cancer susceptibility genes may significantly influence surgical decision-making. BRCA1 and BRCA2 mutation carriers have an increased lifetime risk of developing

breast cancer and an increased risk of contralateral disease. Consequently, surgical planning may include discussion of bilateral mastectomy or contralateral risk-reducing surgery in appropriately selected patients.

However, the presence of a pathogenic variant should not automatically lead to mastectomy. Breast-conserving therapy remains oncologically appropriate for selected mutation carriers, provided that patients are carefully counselled regarding future breast cancer risk, surveillance requirements, and the potential alternatives of risk-reducing surgery.

Patient preference and shared decision-making

Patient preferences represent an essential component of modern breast surgery. In many clinical situations, more than one surgical option can provide appropriate oncological control. The final treatment should therefore reflect not only technical feasibility and oncological safety but also the patient's values, perception of risk, expectations regarding body image, and attitude toward reconstruction or future surveillance.

Shared decision-making is particularly important when discussing breast-conserving surgery versus mastectomy. Long-term randomized studies have demonstrated that mastectomy does not provide a survival advantage over breast-conserving treatment followed by radiotherapy in appropriately selected patients [3,4].

The role of the surgeon is therefore not simply to identify whether an intervention is technically feasible, but to translate oncological, anatomical, and reconstructive information into a treatment strategy consistent with both evidence and patient priorities.

This multidimensional process explains why breast surgical planning represents an attractive but challenging field for artificial intelligence. Several determinants are objectively measurable and potentially suitable for algorithmic processing, whereas others depend on visual assessment, surgical experience, institutional resources, and patient preference. This distinction forms a central conceptual basis for the analysis presented in this thesis.

2.2 Neoadjuvant systemic therapy and its impact on surgical planning

Neoadjuvant systemic therapy (NST) has become an integral component of contemporary breast cancer management. Initially used mainly for locally advanced or inoperable disease, its indications have progressively expanded to operable breast cancer, particularly in biologically aggressive subtypes such as HER2-positive and triple-negative breast cancer (TNBC) [2].

From a surgical perspective, NST has substantially modified the traditional sequence of treatment. Surgery is no longer determined exclusively by the anatomical extent of disease at diagnosis but may also be influenced by the magnitude and pattern of response to systemic therapy. Tumor downstaging may improve the feasibility of breast-conserving surgery, while pathological response provides additional prognostic information and may influence subsequent systemic treatment.

Rationale and biological selection

One of the principal surgical objectives of NST is reduction of tumor burden. Early randomized evidence demonstrated that preoperative chemotherapy can increase the proportion of patients eligible for breast conservation without compromising long-term oncological outcomes compared with the same systemic treatment administered postoperatively [19].

NST also provides an *in vivo* assessment of tumor sensitivity. Pathological complete response (pCR), particularly the absence of residual invasive disease in the breast and axillary lymph nodes, is associated with favorable outcomes, although its prognostic significance varies according to tumor subtype. The CTNeoBC pooled analysis

demonstrated the strongest association between pCR and long-term outcome in HER2-positive and triple-negative disease [20].

Tumor biology therefore plays a major role in patient selection for preoperative treatment. HER2-positive tumors may achieve substantial responses to chemotherapy combined with HER2-directed therapy, including dual HER2 blockade [21]. Furthermore, residual HER2-positive invasive disease after neoadjuvant treatment may guide postoperative escalation to trastuzumab emtansine, as demonstrated by the KATHERINE trial [22].

TNBC similarly represents an important neoadjuvant setting because of its relatively high chemosensitivity and the prognostic relevance of pathological response. The introduction of immunotherapy has further modified treatment in selected high-risk patients, with KEYNOTE-522 demonstrating improved event-free survival following the addition of pembrolizumab to neoadjuvant chemotherapy [23]. Conversely, residual disease after neoadjuvant chemotherapy identifies patients at increased risk of recurrence and may influence postoperative systemic treatment, as illustrated by the CREATE-X trial [24].

The role of NST is more heterogeneous in hormone receptor-positive/HER2-negative disease. Many patients with favorable endocrine-sensitive tumors remain appropriate candidates for upfront surgery, whereas preoperative chemotherapy or endocrine therapy may be considered when downstaging is clinically advantageous [2].

Consequently, the indication for NST is not determined by tumor size alone but results from the integration of biological subtype, clinical stage, nodal status, expected treatment sensitivity and the potential surgical benefit of tumor reduction.

Impact on breast surgical planning

One of the most relevant consequences of NST is the possibility of reducing the extent of breast surgery. In selected patients with an unfavorable initial tumor-to-breast volume ratio, substantial tumor response may convert an initial indication for mastectomy into a feasible breast-conserving procedure.

The extent and distribution of residual disease remain fundamental. Radiological response does not necessarily correspond perfectly to pathological response, and tumors may shrink concentrically or leave scattered residual foci. Surgical planning should therefore integrate the pretreatment extent of disease with post-treatment imaging and clinical assessment rather than relying exclusively on the apparent residual lesion.

Localization of the original tumor site is particularly important when a major response is anticipated. Placement of a radiopaque clip or other marker before systemic therapy can allow identification of the original tumor bed when the lesion becomes difficult to detect after treatment. The surgical objective is consequently to achieve adequate oncological excision while avoiding unnecessary removal of the entire pretreatment breast volume.

Post-treatment mammography, ultrasound and, in selected cases, magnetic resonance imaging may provide complementary information regarding residual disease. These

findings must be interpreted in relation to baseline imaging, particularly in patients initially presenting with extensive disease, multifocality or tumors located in anatomically challenging areas.

Oncoplastic techniques may further expand the possibility of breast conservation following NST. When residual disease requires a relatively wide excision, appropriate glandular reshaping or volume replacement may allow oncologically adequate resection while limiting postoperative deformity.

Therefore, an initial indication for mastectomy does not necessarily remain unchanged after effective systemic treatment. Conversely, an excellent biological or radiological response should not automatically result in breast conservation when disease distribution, breast anatomy, genetic factors or expected cosmetic outcome make such an approach inappropriate.

Axillary and postoperative implications

NST can also downstage axillary disease, particularly in HER2-positive and triple-negative tumors, allowing less extensive axillary surgery in appropriately selected responders. Contemporary post-neoadjuvant axillary strategies are discussed in detail in Section 2.7.

The pathological findings obtained at surgery may additionally influence postoperative systemic treatment. Residual disease after NST can identify patients requiring treatment escalation, as demonstrated in HER2-positive and HER2-negative disease [22,24]. Surgery therefore has a dual function after neoadjuvant treatment: it provides

local control while simultaneously generating pathological information that contributes to subsequent risk stratification and treatment selection.

This close interaction between systemic treatment and operative strategy reinforces the multidisciplinary nature of modern breast surgery. Treatment sequence, lesion localization, response assessment and surgical technique must be planned as components of a continuous therapeutic pathway.

Relevance to AI-assisted decision-making

The decision to administer neoadjuvant treatment represents a particularly relevant domain for AI-assisted clinical decision support. Compared with decisions concerning incision placement, oncoplastic technique or reconstruction, the indication for NST relies to a greater extent on structured and codifiable variables such as tumor subtype, clinical stage, nodal status and tumor dimensions.

This may partly explain the very high concordance observed in BREAST AI-PLAN. Agreement between multidisciplinary assessment and AI-generated recommendations for neoadjuvant treatment was 96.67%, with a Cohen's κ of 0.9112, corresponding to almost perfect agreement [18].

These findings support the hypothesis that AI-assisted systems may reproduce clinical decision-making more consistently when recommendations are based predominantly on explicit guideline-driven criteria. Conversely, agreement may decrease when surgical planning becomes increasingly dependent on anatomical interpretation, technical expertise, aesthetic judgment and patient preference.

2.3 Evolution of breast-conserving surgery

Breast cancer surgery has progressively evolved from radical procedures toward increasingly conservative and individualized treatment. Randomized trials conducted by the Milan group and the National Surgical Adjuvant Breast and Bowel Project demonstrated that breast-conserving surgery followed by radiotherapy can provide long-term survival comparable with mastectomy in appropriately selected patients [3,25].

These findings represented a major conceptual shift in breast surgical oncology. The objective of surgery progressively moved from maximizing the extent of resection toward achieving adequate local control with the minimum necessary removal of healthy tissue. Radiotherapy remains an integral component of breast-conserving treatment in most patients and substantially reduces the risk of ipsilateral recurrence [26].

From tumor size to individualized surgical feasibility

Eligibility for breast-conserving surgery is no longer determined exclusively by absolute tumor diameter. Contemporary planning considers the relationship between tumor extent and breast volume, lesion location, disease distribution, ability to obtain negative margins, breast morphology, expected cosmetic outcome and response to systemic treatment.

The tumor-to-breast volume ratio is particularly relevant. A relatively large lesion may be removed conservatively from a large breast with limited deformity, whereas a smaller tumor in a small breast or unfavorable anatomical location may result in an unacceptable defect.

Tumor location also influences operative strategy. Central, medial and lower-pole lesions may be more challenging because wide excision can substantially alter breast contour or nipple–areola position. These anatomical considerations contributed to the development of oncoplastic breast surgery.

Oncoplastic breast-conserving surgery

Oncoplastic surgery combines oncological resection with techniques of breast reshaping and should be considered an extension of breast-conserving surgery rather than a separate oncological treatment [5].

Its introduction has expanded the indications for breast preservation by allowing wider excisions in patients with unfavorable tumor-to-breast volume ratios or tumors located in cosmetically challenging areas. The anticipated surgical defect becomes part of the preoperative plan, rather than a deformity addressed only after tumor removal.

Oncoplastic procedures can broadly be divided into volume-displacement and volume-replacement approaches. Volume-displacement techniques reshape the residual breast parenchyma, whereas volume-replacement techniques introduce tissue to compensate for the excised volume. Selection depends on breast size, ptosis, tumor location, expected resection volume, tissue quality and surgeon expertise.

The increasing use of these techniques has therefore transformed breast conservation from a relatively standardized lumpectomy into a more complex reconstructive decision. Incision placement, extent of glandular mobilization and the potential need for contralateral symmetrization may all need to be considered before tumor excision.

Oncological adequacy nevertheless remains the primary objective. Complete tumor removal with appropriate negative margins must never be compromised by cosmetic considerations. Contemporary margin principles and localization of non-palpable lesions are discussed separately in Section 2.6.

Breast conservation after neoadjuvant treatment

Neoadjuvant systemic therapy has further expanded the potential for breast preservation. Tumor reduction may improve the tumor-to-breast volume ratio and allow selected patients initially considered candidates for mastectomy to undergo breast-conserving surgery.

As discussed in Section 2.2, however, the final operation cannot be based exclusively on post-treatment tumor size. Baseline disease extent, residual disease distribution, imaging findings, localization of the original tumor bed, breast anatomy and expected cosmetic outcome must all be considered.

Breast-conserving surgery should therefore be regarded as a dynamic strategy in which the definitive procedure may change according to treatment response and individual anatomical characteristics.

Aesthetic and patient-centered outcomes

The goals of modern breast-conserving surgery extend beyond local control. Breast shape, symmetry, nipple position, scarring, body image and quality of life have become increasingly relevant outcomes.

Poor cosmetic results may follow excessive tissue removal, unfavorable scar placement, radiotherapy-related changes or inadequate reshaping. Appropriate

preoperative planning and oncoplastic techniques can reduce these risks, although aesthetic outcome remains inherently more subjective than oncological endpoints.

This distinction is relevant to individualized surgical planning because apparently similar tumors may require different procedures according to breast morphology, expected resection volume and patient priorities.

Relevance to AI-assisted surgical planning

Breast-conserving surgery represents a particularly informative field for AI-assisted decision support because it combines codifiable oncological information with anatomical and technical variables that are more difficult to standardize.

Factors such as tumor size, biological subtype, disease extent and previous treatment can be represented within structured datasets. In contrast, the optimal incision, extent of tissue redistribution and anticipated cosmetic result depend heavily on direct anatomical assessment, three-dimensional breast morphology, technical experience and patient preference.

AI systems may therefore assist in identifying oncologically appropriate treatment pathways, while the detailed operative strategy is likely to remain more dependent on surgeon judgment and direct clinical assessment.

2.4 Mastectomy in contemporary practice

Mastectomy remains an essential component of contemporary breast cancer surgery despite the progressive expansion of breast-conserving and oncoplastic techniques. Its role, however, has changed substantially over time. Modern mastectomy is increasingly conceived not only as removal of the breast gland but as part of an

individualized surgical pathway integrating oncological safety, preservation of the breast envelope when appropriate, reconstructive possibilities and patient preference. The indication for mastectomy should not be determined by tumor size alone. Contemporary decision-making considers disease extent and distribution, feasibility of obtaining negative margins with breast conservation, tumor-to-breast volume ratio, previous treatments, genetic predisposition, expected aesthetic outcome, reconstructive options and patient preference [2].

Mastectomy may be appropriate in patients with extensive or diffuse disease, widespread malignant microcalcifications, persistent positive margins despite appropriate attempts at conservation, or an unfavorable tumor-to-breast volume ratio not amenable to oncoplastic correction. It may also be considered when radiotherapy cannot be safely delivered or when the patient prefers mastectomy after appropriate counselling.

Importantly, mastectomy should not be regarded as intrinsically more curative than breast-conserving treatment. In appropriately selected patients, long-term randomized evidence has demonstrated equivalent survival between mastectomy and breast-conserving surgery followed by radiotherapy [3,4].

Evolution of the mastectomy technique

The increasing integration of immediate reconstruction has progressively changed the technical objectives of mastectomy. Conventional mastectomy removes the breast gland together with most of the overlying skin and the nipple–areola complex, whereas

skin-sparing and nipple-sparing techniques preserve elements of the native breast envelope when oncologically appropriate.

Skin-sparing mastectomy preserves most of the breast skin while removing the gland and nipple–areola complex, facilitating immediate reconstruction. Nipple-sparing mastectomy extends this concept by preserving the nipple–areola complex and can provide favorable aesthetic and psychosocial outcomes in appropriately selected patients [28-30].

Oncological safety remains the principal prerequisite. Direct involvement of the nipple–areola complex contraindicates its preservation, while tumor location, disease extent, imaging characteristics and retroareolar assessment contribute to patient selection [28]. Available evidence supports the oncological safety of nipple-sparing mastectomy in appropriately selected patients, with low reported rates of recurrence involving the preserved nipple–areola complex [29,31].

Anatomical and technical considerations

Preservation of the breast skin and nipple–areola complex introduces specific technical challenges. Complete removal of breast tissue must be balanced against preservation of adequate skin-flap vascularity.

Flap viability may be influenced by skin and subcutaneous tissue characteristics, previous surgery, smoking, vascular disease, BMI and previous radiotherapy. The nipple–areola complex is particularly vulnerable to ischemic complications, making careful patient selection and incision planning essential.

Incision choice is also relevant because different approaches may influence surgical exposure, perfusion, scar position and subsequent reconstruction. Breast size, degree of ptosis, tumor location, previous scars and reconstructive strategy may all affect the optimal approach.

These factors illustrate an important distinction between the indication for mastectomy and the technical execution of mastectomy. The former can often be derived from relatively structured oncological information, whereas the latter frequently depends on direct anatomical assessment and surgical experience.

Integration with breast reconstruction

When reconstruction is planned, mastectomy and reconstruction should ideally be considered as components of the same preoperative strategy rather than as independent sequential decisions.

Immediate reconstruction may provide important aesthetic and psychological advantages, but it is not appropriate for every patient. The decision should integrate oncological requirements with anatomy, comorbidities, tissue quality, previous treatments, anticipated post-mastectomy radiotherapy and patient preference.

Reconstruction may be performed using implant-based or autologous techniques, with the specific choice discussed in detail in Section 2.5. From the perspective of mastectomy planning, the fundamental principle is that preservation of the skin or nipple should be undertaken only when both oncological appropriateness and reconstructive feasibility are established.

The possible need for post-mastectomy radiotherapy adds further complexity because radiation can adversely affect tissue quality and reconstructive outcomes. Nevertheless, anticipated radiotherapy does not represent an absolute contraindication to immediate reconstruction, and treatment should be individualized through multidisciplinary planning [33].

Risk-reducing mastectomy

Mastectomy may also be considered for risk reduction in patients with substantial inherited susceptibility to breast cancer, particularly carriers of pathogenic variants in high-penetrance genes such as BRCA1 or BRCA2.

In patients with unilateral breast cancer, genetic information may influence the discussion of breast conservation, unilateral mastectomy or bilateral surgery. However, mutation status alone should not determine the procedure. Age, estimated contralateral cancer risk, prognosis of the index tumor, surveillance options, reconstructive consequences and patient preference must also be considered.

Risk-reducing surgery therefore represents another area in which shared decision-making is essential and where more than one clinically acceptable strategy may exist.

Relevance to AI-assisted surgical planning

Mastectomy provides a useful example of the distinction between relatively codifiable oncological decisions and more individualized surgical choices.

The indication for mastectomy may often be informed by structured variables such as disease extent, multifocality, previous treatment and feasibility of breast conservation.

In contrast, decisions concerning nipple preservation, incision placement and

reconstructive strategy require increasingly detailed assessment of breast morphology, skin quality, tissue perfusion, ptosis, previous scars and expected aesthetic outcome.

These variables are only partially represented by structured textual information and may depend substantially on physical examination and surgical expertise.

This distinction is consistent with the BREAST AI-PLAN findings. While agreement between AI and multidisciplinary assessment was very high for neoadjuvant therapy, concordance was lower for reconstructive planning, with an observed agreement of 57.14% and a Cohen's κ of 0.3588 [18].

Mastectomy and reconstruction therefore provide a clinically relevant model for exploring the boundary between structured AI-assisted decision support and individualized surgical judgment.

2.5 Reconstructive breast surgery: principles and patient selection

Breast reconstruction is an integral component of contemporary breast cancer surgery and should ideally be considered during initial treatment planning. The objective is to restore breast shape while maintaining oncological safety and minimizing morbidity, taking into account patient anatomy, comorbidities, previous treatments and individual preferences.

Reconstruction may be performed using implants, autologous tissue or a combination of both. Implant-based reconstruction can be completed as a single-stage direct-to-implant (DTI) procedure or through a two-stage approach using a temporary tissue expander followed by a definitive implant [32]. Autologous reconstruction represents an alternative, particularly in selected patients in whom implant-based procedures may

be less suitable because of previous radiotherapy, tissue characteristics or patient preference.

Implant-based reconstruction and patient selection

When implant reconstruction is considered, the surgeon must determine whether definitive reconstruction can safely be completed during the same procedure or whether a staged approach is preferable.

DTI reconstruction may be appropriate when the mastectomy skin envelope is adequately preserved and flap vascularity is satisfactory. Conversely, an expander-based strategy may provide greater flexibility when tissue viability is uncertain or when gradual adaptation of the breast envelope is desirable [32].

Implant position represents a further component of reconstructive planning. Subpectoral reconstruction provides muscular implant coverage, whereas prepectoral reconstruction avoids elevation of the pectoralis major muscle. The feasibility of a prepectoral approach depends particularly on the quality, thickness and vascularity of the mastectomy flaps.

Patient selection therefore remains fundamental. BMI, smoking, vascular disease and previous radiotherapy may influence healing and reconstructive complications. Anatomical variables including breast volume, degree of ptosis, skin envelope and subcutaneous tissue thickness may further determine the feasibility of DTI reconstruction, implant positioning and nipple-preserving procedures.

Radiotherapy and patient preference

The potential need for post-mastectomy radiotherapy should also be considered during reconstructive planning. Radiation-related changes may adversely affect tissue quality and increase complications in implant-based reconstruction. Nevertheless, anticipated radiotherapy does not automatically preclude immediate reconstruction, and the optimal strategy should be individualized within a multidisciplinary pathway [33].

Patient preference is equally important. Different reconstructive options may be oncologically and technically acceptable while producing different consequences in terms of scars, recovery, donor-site morbidity, future procedures and aesthetic expectations. Reconstruction should therefore be considered a shared decision rather than a purely technical choice.

Relevance to BREAST AI-PLAN

Reconstructive planning represents a particularly challenging domain for AI-assisted decision support because it combines measurable clinical variables with anatomical and subjective factors.

The BREAST AI-PLAN workflow included several parameters relevant to reconstruction, including BMI, smoking status, previous radiotherapy, vascular disease, BRCA status, breast volume, breast density and subcutaneous tissue thickness.

The AI system was asked to provide recommendations regarding reconstruction, DTI feasibility and implant characteristics [18].

However, factors such as skin elasticity, flap perfusion, degree of ptosis, previous scars, three-dimensional breast morphology and patient expectations cannot be fully represented through structured textual information.

Consistent with this complexity, agreement between AI-generated and clinical reconstructive planning was 57.14%, with a Cohen's κ of 0.3588, corresponding to fair agreement [18].

These findings suggest that structured AI may assist in organizing reconstructive variables, but detailed operative planning remains strongly dependent on direct clinical assessment, surgical experience and patient preference.

2.6 Surgical margins and localization of non-palpable lesions

Achieving complete tumor excision with negative margins is a fundamental objective of breast-conserving surgery. At the same time, excessive removal of healthy tissue may compromise cosmetic outcome and increase the need for reconstructive correction. Contemporary surgical planning therefore requires a balance between oncological adequacy and preservation of breast volume.

For invasive breast carcinoma treated with breast-conserving surgery and whole-breast irradiation, the accepted definition of an adequate negative margin is “no ink on tumor” [27]. Wider microscopic margins have not been shown to provide additional clinically meaningful reduction in ipsilateral breast tumor recurrence and may result in unnecessary re-excision and loss of healthy breast tissue.

Positive margins generally require further surgical management, whereas close but negative margins should be interpreted within the overall clinical context. Tumor histology, disease distribution, imaging findings and the presence of an extensive intraductal component may all influence management. In ductal carcinoma in situ

(DCIS), the relationship between mammographic microcalcifications, pathological extent and margin assessment is particularly relevant.

Localization of non-palpable lesions

The increasing detection of non-palpable breast cancers through screening and high-resolution imaging has made accurate lesion localization an essential component of breast-conserving surgery.

The objective of localization is to allow reliable identification and removal of the target while minimizing unnecessary excision of surrounding healthy tissue. Wire-guided localization has historically represented a widely used approach, but it has several practical limitations, including the need for radiological placement close to the time of surgery, patient discomfort, possible wire displacement and constraints on incision placement.

Alternative wire-free techniques include radioactive seeds, magnetic markers and radiofrequency identification systems. Comparative evidence indicates that several of these approaches can achieve high localization and surgical success rates while offering logistical advantages and greater flexibility in operative scheduling [34]. Selection of the optimal technique depends on lesion characteristics, available technology, institutional workflow and local radiological and surgical expertise.

Localization is particularly important in patients undergoing neoadjuvant systemic therapy. When substantial tumor response is anticipated, placement of a clip or marker before treatment may permit identification of the original tumor site even when the lesion becomes difficult to visualize after therapy.

Specimen assessment and oncoplastic surgery

Accurate specimen orientation is essential to allow pathological identification of individual margins. In non-palpable lesions, intraoperative specimen radiography may also confirm removal of the imaging target, microcalcifications, clip or localization device and may guide selective additional excision when required.

This coordination between imaging, localization, surgical technique and pathological assessment becomes particularly important in oncoplastic surgery. Wide resection followed by immediate glandular rearrangement can alter the position of the original tumor bed and may make subsequent re-excision more difficult. Precise preoperative localization and specimen orientation should therefore form part of the operative plan when extensive tissue mobilization is anticipated.

Margin management should consequently be considered a multidisciplinary process involving the surgeon, radiologist and pathologist rather than an isolated postoperative pathological assessment.

Relevance to AI-assisted surgical planning

Margin management and lesion localization occupy an intermediate position between highly standardized and highly individualized surgical decisions.

Several relevant variables can be represented through structured data, including lesion size, imaging visibility, presence of microcalcifications, previous marker placement and planned type of surgery. In contrast, selection of the optimal localization strategy and intraoperative management of margins may depend on direct interpretation of imaging, institutional resources and real-time surgical findings.

AI-based systems may therefore assist in organizing information relevant to the feasibility of breast conservation, but detailed localization strategy and intraoperative margin management remain closely dependent on radiological assessment and operative judgment.

This domain further illustrates that the potential usefulness of AI in breast surgery is related to the extent to which the information underlying a decision can be represented through reproducible structured variables rather than through visual interpretation and technical expertise.

2.7 Axillary surgery and surgical de-escalation

Axillary management has undergone one of the most important processes of surgical de-escalation in breast cancer treatment. Historically, axillary lymph node dissection (ALND) was routinely performed for both staging and regional disease control. However, ALND is associated with clinically relevant morbidity, including lymphedema, sensory disturbances, shoulder dysfunction and chronic pain.

The introduction of sentinel lymph node biopsy (SLNB) represented a major change in this paradigm. In clinically node-negative patients, SLNB provides accurate axillary staging with substantially lower morbidity than routine ALND and consequently became the standard surgical staging procedure for most patients with early invasive breast cancer [35].

More recently, axillary surgery has progressively shifted from a procedure-based approach toward a more selective strategy in which the extent of surgery depends on

whether additional nodal information or treatment is expected to modify patient management.

Omission of completion axillary dissection

The ACOSOG Z0011 trial demonstrated that completion ALND could be safely omitted in selected patients with cT1–T2 breast cancer, clinically negative axillae and one or two metastatic sentinel lymph nodes undergoing breast-conserving surgery, radiotherapy and systemic treatment [36].

This finding challenged the traditional assumption that identification of metastatic disease in the sentinel node should automatically result in complete axillary clearance.

The IBCSG 23-01 trial further supported de-escalation in patients with sentinel-node micrometastases, demonstrating that omission of ALND did not compromise long-term disease control while reducing surgical morbidity [37].

More recently, the SENOMAC trial extended this principle to a broader population of clinically node-negative patients with one or two sentinel-node macrometastases, supporting omission of completion ALND in appropriately selected patients receiving contemporary multidisciplinary treatment [38].

Consequently, nodal positivity remains clinically relevant, but the need for additional axillary surgery increasingly depends on nodal burden, type of breast surgery, radiotherapy, systemic therapy and the potential impact of further pathological information on treatment decisions.

Axillary radiotherapy as an alternative to surgery

Regional control may also be achieved without completion ALND in selected sentinel-node-positive patients.

The AMAROS trial compared ALND with axillary radiotherapy and demonstrated excellent regional control with both approaches, while axillary radiotherapy was associated with lower rates of treatment-related arm morbidity [39].

These findings reinforce the broader principle that effective regional disease control does not necessarily require maximal surgical clearance. Selection between surgery and radiotherapy should therefore be incorporated into the overall multidisciplinary treatment strategy.

Toward omission of sentinel lymph node biopsy

The most recent phase of axillary de-escalation has questioned whether surgical nodal staging is necessary at all in selected clinically node-negative patients.

The SOUND trial evaluated patients with small breast cancers and negative preoperative axillary ultrasonography and demonstrated that omission of axillary surgery was non-inferior to SLNB in appropriately selected women undergoing breast-conserving treatment [40].

The INSEMA trial subsequently provided further randomized evidence supporting omission of SLNB in selected clinically node-negative patients undergoing breast-conserving surgery. Avoidance of axillary surgery did not compromise invasive disease-free survival and was associated with lower arm morbidity and improved functional outcomes [41].

These studies do not imply that axillary staging has become universally unnecessary. Rather, they emphasize the importance of selecting patients in whom nodal pathological information is unlikely to alter systemic or radiation treatment.

Axillary management after neoadjuvant treatment

Neoadjuvant systemic therapy introduces additional complexity because both breast and nodal disease may be downstaged.

Patients who are clinically node-negative before treatment may undergo SLNB after systemic therapy in appropriately selected situations. In patients presenting with proven nodal involvement who subsequently achieve clinical nodal response, contemporary strategies increasingly aim to reduce the extent of axillary surgery while maintaining reliable assessment of residual disease [2].

Post-neoadjuvant axillary planning therefore requires integration of the initial nodal burden, treatment response, residual disease and the implications of nodal information for subsequent systemic and radiation therapy.

This setting illustrates the transition from fixed surgical procedures toward conditional treatment pathways in which the extent of surgery depends on both baseline disease and response to treatment.

Implications for individualized surgical planning and AI

The evolution of axillary surgery reflects a broader transition from procedure-driven to information-driven treatment. The relevant clinical question is increasingly whether additional axillary surgery or pathological information will meaningfully influence subsequent management.

This makes the axilla conceptually relevant to AI-assisted decision support. Tumor stage, clinical nodal status, biological subtype, type of breast surgery and previous systemic treatment can all be represented through structured variables and incorporated into guideline-oriented decision pathways.

However, axillary surgery was not evaluated as an independent concordance endpoint in BREAST AI-PLAN. Clinical nodal status was included among the variables provided to the AI system, but the predefined study outputs concerned neoadjuvant treatment, planned breast surgery, surgical approach, reconstruction, DTI feasibility, implant type, adjuvant treatment and aesthetic assessment [18]. Therefore, no specific AI-MDT performance for axillary management can be inferred from the available results.

Nevertheless, axillary de-escalation provides an important conceptual example for the present thesis because recommendations in this field have changed rapidly over time. An AI-based clinical decision-support system must therefore rely not only on structured data but also on current evidence, precisely defined eligibility criteria and appropriate clinical context.

Axillary surgery consequently illustrates both the potential value of AI in navigating increasingly complex treatment pathways and the continuing need for evidence updating and multidisciplinary human oversight.

2.8 Multidisciplinary decision-making in breast cancer

Modern breast cancer treatment is inherently multidisciplinary. Surgical planning cannot be considered independently from tumor biology, systemic therapy, radiotherapy, imaging, pathology, genetic information and patient preference. The increasing complexity of treatment has therefore transformed breast surgery from an isolated operative decision into one component of an integrated therapeutic pathway.

Multidisciplinary team (MDT) meetings provide the organizational framework in which these different domains are combined. A typical breast MDT includes breast surgeons, medical oncologists, radiation oncologists, radiologists and pathologists, with additional input from plastic surgeons, geneticists, nuclear medicine specialists and breast care nurses when required.

The importance of specialist multidisciplinary care is reflected in European standards for dedicated breast units. EUSOMA quality indicators emphasize standardized pathways, specialist expertise and systematic monitoring of the quality of breast cancer diagnosis and treatment [6,42]. Observational evidence has also associated implementation of multidisciplinary breast cancer care with improved consistency of treatment and reduced variability in outcomes between institutions [43].

Integration of structured and individualized information

Breast cancer MDT decision-making combines variables that differ substantially in their degree of standardization.

Tumor histology, hormone receptor status, HER2 status, tumor dimensions and clinical nodal status can generally be represented through predefined categories and often contribute directly to guideline-based treatment pathways.

Other aspects of surgical planning require greater interpretation. Selection between breast conservation and mastectomy may depend on tumor-to-breast volume ratio, disease distribution and anticipated cosmetic outcome. Similarly, oncoplastic and reconstructive planning requires assessment of breast morphology, ptosis, skin quality, previous scars, tissue characteristics and patient expectations.

The MDT therefore does not simply apply guidelines. Rather, it integrates evidence-based recommendations with the characteristics of the individual patient and the technical feasibility of different surgical options.

Variability and standardization

Multidisciplinary decision-making is not completely reproducible. Institutional protocols, available reconstructive techniques, surgeon expertise and local resources may influence the final recommendation. In addition, when several oncologically acceptable alternatives exist, patient preference may appropriately lead to different treatment choices.

Documentation can represent an additional source of variability. Multicenter evaluation of breast cancer MDT activity has identified missing or discordant information in clinically relevant domains despite overall high consistency between MDT discussion and medical records [44].

This issue is particularly relevant in surgery because several determinants of operative planning may be incompletely represented in structured documentation. Breast shape, ptosis, skin elasticity, previous scars and patient aesthetic priorities may be evident during direct clinical assessment but not fully captured in a written dataset.

Standardization should therefore support, rather than replace, individualized clinical judgment.

The MDT as the clinical reference in BREAST AI-PLAN

In BREAST AI-PLAN, multidisciplinary clinical planning constituted the reference against which AI-generated recommendations were compared [18].

The structured clinical forms included demographic characteristics, comorbidities, breast anatomy, tumor biology and staging, which were converted into standardized prompts generating predefined treatment recommendations.

Agreement with the MDT should not be interpreted as proof that an AI recommendation was independently “correct.” Rather, concordance measures the extent to which the AI system reproduced real-world specialist decision-making using the information provided to it.

This distinction is important because agreement and clinical validity are not synonymous. High concordance may indicate that a decision is highly standardized and reproducible, whereas disagreement may reflect missing information, multiple acceptable treatment options or a greater contribution of human judgment.

In the published BREAST AI-PLAN study, concordance was highest for neoadjuvant treatment and lower for reconstructive and aesthetic decisions [18]. This pattern

supports the distinction between decisions based predominantly on structured guideline-driven variables and those requiring more complex anatomical, technical or subjective assessment.

From multidisciplinary care to AI-assisted decision support

The increasing amount of clinical, pathological, radiological and molecular information considered during MDT discussion provides a potential rationale for AI-assisted decision support.

Structured AI systems may help organize information, highlight guideline-relevant variables and generate preliminary recommendations for subsequent specialist review. Their role should nevertheless remain complementary to multidisciplinary clinical assessment.

The BREAST AI-PLAN workflow relied predominantly on structured textual clinical information rather than direct interpretation of imaging [18]. Consequently, physical examination, three-dimensional breast anatomy, patient–surgeon interaction, intraoperative findings and institutional expertise were not fully represented within the AI input.

The MDT therefore remains particularly important in situations characterized by uncertainty, multiple acceptable therapeutic alternatives or preference-sensitive decisions.

The interaction between structured evidence and individualized expert judgment provides the conceptual transition to the following chapter, which examines the potential and limitations of artificial intelligence in breast surgical decision-making.

3. ARTIFICIAL INTELLIGENCE IN SURGERY AND BREAST CANCER

Artificial intelligence in breast cancer and breast surgery

Artificial intelligence (AI) encompasses computational approaches designed to process complex information and generate predictions, classifications or other structured outputs. In breast cancer, early applications focused predominantly on imaging and pathology, while more recent studies have explored risk prediction, treatment selection, aesthetic assessment and surgical planning [7-9,16].

The development of generative AI and large language models (LLMs) has introduced a different form of clinical decision support. Unlike conventional predictive models developed for a predefined task, LLMs can process information presented in natural language and generate context-dependent responses to structured clinical questions [10,11].

Breast surgery represents a particularly challenging setting for these technologies because operative planning requires simultaneous integration of oncological, anatomical and patient-related variables. Tumor biology and stage can often be represented through standardized categories, whereas breast morphology, tissue quality, anticipated cosmetic outcome, reconstructive feasibility and patient preference may require direct clinical interpretation.

Existing applications of AI in breast and reconstructive surgery include anatomical assessment, breast symmetry evaluation, prediction of aesthetic outcomes and assistance with reconstructive planning [7-9,13,14]. These applications are particularly

relevant to oncoplastic surgery, in which oncological adequacy must be integrated with the expected anatomical and cosmetic consequences of resection.

The potential value of AI therefore depends not only on whether a model can generate a clinically plausible recommendation, but also on whether the information required for that decision can be adequately represented in the data provided to the system.

Prompt-driven clinical decision support

LLM-based decision support is strongly dependent on the prompt used to provide clinical information and define the requested output. Ambiguous or incomplete inputs may result in variable recommendations, whereas structured prompts can improve consistency and facilitate comparison between cases.

For clinical applications, standardization of both input and output is therefore particularly important. Each patient should ideally be represented using comparable clinical variables, while the questions posed to the model should be predefined and clinically meaningful.

BREAST AI-PLAN was developed according to this principle. Patients were represented using structured information concerning demographic characteristics, comorbidities, breast anatomy, tumor biology and clinical stage. These variables were incorporated into a guideline-oriented prompt requesting eight predefined outputs: suitability for neoadjuvant therapy, planned breast surgery, surgical approach, reconstruction, direct-to-implant feasibility, implant type, adjuvant therapy and expected aesthetic result [18].

This approach differs from asking a general-purpose AI system to provide an open-ended treatment recommendation. By standardizing the information provided and the required outputs, BREAST AI-PLAN attempted to transform a conversational technology into a more reproducible clinical decision-support workflow.

The study consequently evaluated agreement between predefined AI-generated recommendations and multidisciplinary clinical planning rather than the simple ability of the model to produce plausible medical text [18].

Limitations of generative AI in surgical decision-making

Several limitations remain relevant when LLMs are applied to clinical medicine.

First, AI-generated recommendations are necessarily dependent on the information contained in the input. Relevant variables that are not included in the prompt cannot be assumed to influence the model appropriately. This represents a particular limitation in surgery, where clinical examination, visual assessment, tissue characteristics and intraoperative findings may substantially affect treatment.

Second, LLMs can generate fluent and apparently authoritative responses that may nevertheless be inaccurate or unsupported. Linguistic plausibility therefore cannot be considered equivalent to clinical reliability, and expert verification remains necessary [10-12].

Model characteristics, training data and prompt formulation may also influence output. Performance observed with one specific workflow cannot automatically be generalized to another model, prompt structure or clinical setting. Bias and limited transparency

represent additional concerns, particularly when particular patient groups or surgical approaches are under-represented in the information underlying a system [15].

These limitations are particularly relevant in preference-sensitive areas such as breast conservation and reconstruction. A technically feasible recommendation may still be inappropriate if the patient's expectations, priorities or tolerance for additional procedures have not been adequately represented.

BREAST AI-PLAN illustrates several of these issues. The workflow relied predominantly on structured textual clinical information rather than direct AI interpretation of imaging, and the published study identified the absence of multimodal data and the possibility of transcription bias among its limitations [18].

Future systems may incorporate imaging, pathological and genomic information, standardized clinical photography or three-dimensional anatomical data, and patient-reported outcomes. Nevertheless, broader data integration does not eliminate the need for clinical supervision. AI should therefore currently be considered a decision-support tool rather than an autonomous decision-maker.

Rationale and aims of the present study

Modern breast surgery offers a useful model for investigating AI-assisted decision support because different clinical decisions vary substantially in their degree of standardization.

Some treatment choices, particularly the indication for neoadjuvant systemic therapy, depend largely on codifiable variables such as tumor subtype, clinical stage and nodal status. In contrast, selection of a surgical approach, reconstructive strategy or

anticipated aesthetic outcome may require greater integration of breast anatomy, tissue characteristics, technical expertise and patient preference.

The present thesis builds upon the BREAST AI-PLAN experience previously reported by Cossu et al. [18]. The dataset and primary concordance analysis have already been published; the present work places these findings within the broader context of contemporary oncoplastic and reconstructive breast surgical decision-making.

The principal aim was to evaluate agreement between AI-generated recommendations and multidisciplinary clinical planning, with particular attention to the proposed surgical procedure. Secondary analyses examined concordance for neoadjuvant treatment, surgical approach, reconstructive strategy, adjuvant treatment and aesthetic assessment.

The underlying hypothesis was that agreement would be greater for decisions based predominantly on explicit and standardized clinical variables and lower for decisions requiring greater anatomical interpretation, technical judgment, aesthetic assessment or patient-specific considerations.

The purpose was therefore not to determine whether artificial intelligence could replace multidisciplinary decision-making, but to identify which components of breast surgical planning may be most suitable for structured AI support and which remain predominantly dependent on human clinical judgment.

4. MATERIALS AND METHODS

Study design and setting

This was a retrospective, single-center observational study conducted at the Breast Surgery Unit of IRCCS Azienda Ospedaliera Metropolitana – plesso Ospedale Policlinico San Martino, Genoa, Italy.

The study evaluated agreement between multidisciplinary clinical planning and recommendations generated through the BREAST AI-PLAN artificial intelligence workflow in women undergoing surgery for breast cancer.

Thirty consecutive women treated between January 2024 and May 2025 were included. For each patient, the multidisciplinary treatment plan was compared with the AI-generated recommendation obtained from the corresponding standardized clinical information, allowing a paired within-case analysis.

The dataset and primary concordance analysis have previously been reported by Cossu et al. [18]. The present thesis uses the same cohort and expands the interpretation of these findings within the context of contemporary oncoplastic and reconstructive breast surgery.

Study population

Inclusion criteria were female sex, age ≥ 18 years, histologically confirmed invasive ductal or lobular carcinoma or ductal carcinoma in situ (DCIS), clinical tumor stage $\leq cT3$, clinical nodal status cN0–cN1, and suitability for breast surgery with or without reconstruction.

Exclusion criteria were metastatic disease (M1), locally advanced inoperable disease (T4), uncontrolled comorbidities, contraindications to surgery or anesthesia, BMI > 40

kg/m² or <16 kg/m², and active heavy smoking associated with a high risk of tissue necrosis [18].

Given the limited sample size, the study was considered exploratory and hypothesis-generating rather than a definitive validation study.

Clinical variables and data collection

Each patient was characterized using a standardized clinical form containing variables relevant to oncological and surgical planning.

Patient-related variables included age, BMI, smoking status, previous thoracic radiotherapy, vascular disease and BRCA1/2 status. Breast-related variables included breast volume, breast density and subcutaneous tissue thickness. Tumor-related variables included histological type, estrogen receptor (ER) and progesterone receptor (PgR) status, HER2 status, Ki-67 proliferation index, molecular subtype, and clinical cT and cN stage.

The dataset also incorporated information regarding the feasibility of breast conservation versus mastectomy and potential reconstructive options [18].

This combination of oncological, anatomical and patient-related information formed the structured input used for AI-assisted planning.

Reference multidisciplinary clinical planning

For each patient, the treatment strategy established through specialist assessment and multidisciplinary discussion constituted the clinical reference against which the AI-generated recommendation was compared.

The term reference clinical decision is used rather than “gold standard”, since multidisciplinary planning represents expert real-world decision-making but may also reflect institutional practice, surgeon experience and patient-specific considerations. Agreement with the multidisciplinary plan was therefore interpreted as concordance with clinical decision-making rather than proof that a given AI recommendation represented the uniquely correct treatment.

BREAST AI-PLAN workflow

The structured clinical information available for each patient was transcribed into a guideline-oriented prompt submitted to the AI assistant.

The prompt requested a standardized response addressing eight predefined domains:

1. suitability for neoadjuvant therapy;
2. planned breast surgery;
3. surgical approach;
4. type of reconstruction;
5. feasibility of direct-to-implant reconstruction;
6. implant type;
7. adjuvant therapy;
8. expected aesthetic result according to a four-point BCCT.core-aligned scale.

Standardization of both input variables and requested outputs was intended to reduce variability associated with open-ended prompting and to permit direct comparison between AI-generated and clinical recommendations [18].

The workflow relied predominantly on structured textual clinical information rather than direct AI interpretation of imaging. It should therefore be considered a text-based structured decision-support system rather than a multimodal AI workflow [18].

Classification of surgical recommendations

To allow quantitative comparison, planned surgical procedures were converted into predefined categorical codes:

- A: targeted lumpectomy plus sentinel lymph node biopsy;
- B: targeted lumpectomy plus intraoperative specimen radiography;
- C: simple mastectomy plus sentinel lymph node biopsy;
- D: mastectomy with reconstruction;
- E: targeted lumpectomy plus axillary lymph node dissection;
- F: oncoplastic surgery.

Combined categories, including A+D, A+E, C+F and D+F, were retained according to the original classification.

Surgical approach was categorized as I, periareolar; II, inframammary fold; III, Lazy-S; and IV, other.

Reconstructive procedures were classified as:

1. tissue expander;
2. definitive subpectoral implant;
3. definitive prepectoral implant;
4. surgical flap;
5. level I oncoplastic surgery;

6. level II–III oncoplastic surgery [18].

Study outcomes

The principal analysis of interest in the present thesis was agreement between AI-generated and multidisciplinary recommendations regarding the planned surgical procedure.

Additional analyses assessed concordance for neoadjuvant therapy, surgical approach, reconstructive strategy, adjuvant therapy and aesthetic assessment.

Direct-to-implant feasibility and implant characteristics were included among the predefined AI outputs but were not reported as separate concordance estimates in the published analysis. Accordingly, no additional κ analyses for these variables were introduced in the present thesis [18].

Aesthetic assessment

Expected aesthetic outcome was classified using a four-category scale aligned with the BCCT.core approach, ranging from 1 (poor) to 4 (excellent).

Because aesthetic assessment represents an ordinal variable, agreement was evaluated using weighted Cohen's κ . Both linear and quadratic weighting were applied to account for the different magnitude of disagreement between adjacent and more distant categories [18].

Statistical analysis

Descriptive statistics were used to summarize the study population. Continuous variables were reported as mean $\pm$ standard deviation (SD) or median and interquartile

range (IQR), as appropriate. Categorical variables were reported as absolute numbers and percentages.

Agreement between multidisciplinary clinical planning and AI-generated recommendations was assessed using Cohen's kappa coefficient (κ), which accounts for agreement expected by chance. Observed and expected agreement were calculated for each evaluated domain, and κ coefficients were reported together with their standard errors and inferential statistics.

For aesthetic assessment, weighted κ coefficients were calculated using both linear and quadratic weights.

Kappa coefficients were interpreted according to Landis and Koch [45]:

- <0.00 : poor agreement;
- 0.00 – 0.20 : slight agreement;
- 0.21 – 0.40 : fair agreement;
- 0.41 – 0.60 : moderate agreement;
- 0.61 – 0.80 : substantial agreement;
- 0.81 – 1.00 : almost perfect agreement.

Given the small cohort and exploratory study design, κ estimates were interpreted according to their magnitude and clinical context and were not considered evidence of definitive model validation.

5. RESULTS

Patient characteristics

A total of 30 women undergoing breast cancer surgery were included in the analysis. Mean age was **59.2±9.8 years**, while median BMI was **24.2 kg/m² (IQR 20.9–26.8)**. Ten patients (33.3%) were active smokers and four (13.3%) had previously undergone thoracic radiotherapy. Peripheral vascular disease was present in one patient (3.3%), and one patient (3.3%) was BRCA2 positive.

Breast volume was heterogeneous, although intermediate volume categories predominated. Thirteen patients (43.3%) had a breast volume corresponding to approximately 500 cc, six (20.0%) to 650 cc, and five (16.7%) to 800 cc. Mean subcutaneous breast thickness was **10.8±3.2 mm**. Microcalcifications were present in four patients (13.3%).

The majority of patients were clinically node-negative, with **27 patients (90.0%) classified as cN0** and three (10.0%) presenting clinical nodal involvement.

Regarding tumor biology, Luminal A and Luminal B tumors represented the majority of the cohort, accounting for **43.3% and 26.7%**, respectively. HER2-positive tumors represented 16.7% of cases and triple-negative breast cancers 13.3%.

Table 1. Main baseline characteristics of the study population

Characteristic	Study population (n=30)
Age, years	59.2 ± 9.8
BMI, kg/m ²	24.2 [20.9–26.8]
Active smoking	10 (33.3%)
Previous radiotherapy	4 (13.3%)
Peripheral vascular disease	1 (3.3%)
BRCA2-positive	1 (3.3%)
Microcalcifications	4 (13.3%)
cN0	27 (90.0%)
cN+	3 (10.0%)
Luminal A	13 (43.3%)

Characteristic	Study population (n=30)
Luminal B	8 (26.7%)
HER2-positive	5 (16.7%)
Triple-negative	4 (13.3%)

Continuous variables are reported as mean $\pm$ standard deviation or median [interquartile range], as appropriate.

Neoadjuvant treatment

The multidisciplinary clinical assessment indicated neoadjuvant treatment in seven patients (23.3%), compared with eight patients (26.7%) according to the AI-generated recommendations.

This domain showed the highest agreement of all evaluated variables. Observed agreement between the clinical and AI classifications was **96.67%**, compared with an expected agreement of 62.44%. Cohen's κ was **0.9112 (SE=0.1819)**, corresponding to almost perfect agreement. The concordance was statistically significant ($Z=5.01$; $p<0.001$).

Planned surgery

Agreement regarding the planned surgical procedure was observed in **70.00%** of cases, compared with an expected agreement of 28.33%. Cohen's κ was **0.5814 (SE=0.0972)**, corresponding to moderate agreement. The observed concordance was significantly greater than that expected by chance ($Z=5.98$; $p<0.001$).

Differences were also observed in the distribution of individual surgical recommendations. Targeted lumpectomy with sentinel lymph node biopsy (code A) was proposed in 46.7% of clinical evaluations and 40.0% of AI evaluations.

Conversely, mastectomy with reconstruction (code D) represented **20.0% of clinical recommendations compared with 43.3% of AI recommendations**.

Thus, although overall agreement for planned surgery was moderate, the distribution of procedures was not identical between clinical and AI planning.

Surgical approach

For surgical approach, the observed agreement was **60.00%**, compared with an expected agreement of 30.33%. Cohen's κ was **0.4258 (SE=0.1125)**, corresponding to moderate agreement. This result was statistically significant ($Z=3.79$; $p<0.001$).

The periareolar approach was the most frequently selected option in both classifications, accounting for 43.3% of clinical decisions and 40.0% of AI recommendations. The inframammary approach was selected in 10.0% of cases by both assessments, whereas the Lazy-S approach was proposed more frequently by AI (23.3% vs 13.3%).

Reconstructive planning

Agreement regarding reconstruction was lower than that observed for planned surgery or surgical approach. The observed concordance was **57.14%**, compared with an expected agreement of 33.16%, resulting in a Cohen's κ of **0.3588 (SE=0.0906)**, corresponding to fair agreement. The result remained statistically significant ($Z=3.96$; $p<0.001$).

Differences were evident in the distribution of reconstructive recommendations. A definitive prepectoral implant was proposed in **3.3% of clinical assessments compared with 23.3% of AI assessments**, whereas a definitive subpectoral implant

was selected in 13.3% and 3.3%, respectively. Surgical flap reconstruction represented the most frequent reconstructive category in both assessments, although it was more common in clinical planning (60.0% vs 43.3%).

These data describe the areas of disagreement but are not interpreted further here; their potential relationship with anatomical assessment and the limitations of structured textual input is considered in the Discussion.

Adjuvant treatment

Agreement regarding adjuvant treatment was observed in **64.29%** of evaluable cases, compared with an expected agreement of 49.62%. Cohen's κ was **0.2911** (**SE=0.0961**), corresponding to fair agreement. This result was statistically significant ($Z=3.03$; $p=0.0012$).

Adjuvant treatment therefore showed lower chance-corrected agreement than neoadjuvant therapy, planned surgery and surgical approach.

Aesthetic assessment

Aesthetic scoring showed a different pattern from the other evaluated domains.

Using linear weights, observed agreement was **80.95%**, but expected agreement was also high at 77.81%, resulting in a weighted κ of only **0.1418** (**SE=0.1464**). With quadratic weighting, observed agreement increased to **91.27%**, while expected agreement was 89.60%, producing a weighted κ of **0.1608** (**SE=0.1763**).

Both coefficients corresponded to slight agreement and neither reached statistical significance (linear weights: $Z=0.97$, $p=0.1665$; quadratic weights: $Z=0.91$, $p=0.1809$).

Most aesthetic evaluations were concentrated in the higher categories. Clinicians assigned scores of 3 or 4 in 83.3% of the entire cohort, while the AI assigned these scores in 90.0%. Two cases lacked an aesthetic assessment in both classifications.

Table 2. Agreement between multidisciplinary clinical planning and AI-generated recommendations

Domain	Observed agreement	Expected agreement	Cohen's κ	Interpretation	p-value
Neoadjuvant therapy	96.67%	62.44%	0.9112	Almost perfect	<0.001
Planned surgery	70.00%	28.33%	0.5814	Moderate	<0.001
Surgical approach	60.00%	30.33%	0.4258	Moderate	<0.001
Reconstruction	57.14%	33.16%	0.3588	Fair	<0.001
Adjuvant therapy	64.29%	49.62%	0.2911	Fair	0.0012
Aesthetic score – linear weights	80.95%	77.81%	0.1418	Slight	0.1665
Aesthetic score – quadratic weights	91.27%	89.60%	0.1608	Slight	0.1809

Weighted Cohen's κ was used for aesthetic assessment because of the ordinal nature of the score.

Overall pattern of concordance

Overall, the level of agreement between multidisciplinary planning and AI-generated recommendations varied substantially according to the clinical domain evaluated.

The highest concordance was observed for neoadjuvant therapy ($\kappa=0.9112$), followed by planned surgery ($\kappa=0.5814$) and surgical approach ($\kappa=0.4258$). Agreement was lower for reconstruction ($\kappa=0.3588$) and adjuvant therapy ($\kappa=0.2911$).

Aesthetic assessment represented a distinct finding: despite high crude observed agreement, weighted κ remained low because expected agreement was also very high.

Taken together, these results demonstrate substantial heterogeneity in the ability of the AI workflow to reproduce different components of multidisciplinary breast cancer planning. The clinical significance of this gradient and its relationship with the degree of standardization, anatomical complexity and subjectivity inherent to each decision are explored in the following Discussion.

6. DISCUSSION

The present study evaluated the agreement between a prompt-driven artificial intelligence workflow and multidisciplinary clinical planning across different domains of breast cancer treatment. The principal finding was not a uniform level of concordance across all decisions, but rather a marked variation according to the type of clinical question being addressed.

Agreement was almost perfect for neoadjuvant treatment, moderate for planned surgery and surgical approach, fair for reconstructive and adjuvant planning, and only slight for aesthetic assessment after correction for chance. This pattern suggests that the ability of an AI system to reproduce multidisciplinary recommendations may depend substantially on the degree to which the relevant clinical decision can be expressed through explicit, standardized and reproducible variables.

Importantly, agreement should not be interpreted as equivalent to clinical accuracy. The multidisciplinary plan represented the clinical reference in this study, but not an objective or universally applicable gold standard. Surgical decision-making may legitimately differ according to institutional protocols, surgeon expertise, available

reconstructive techniques and patient preference. Consequently, the present results primarily describe the ability of the AI workflow to reproduce real-world clinical planning using the information provided to it.

Neoadjuvant treatment and guideline-driven decisions

The highest level of concordance was observed for neoadjuvant treatment, with an observed agreement of 96.67% and a Cohen's κ of 0.9112, corresponding to almost perfect agreement.

This finding is consistent with the relatively structured nature of neoadjuvant treatment selection. Clinical stage, nodal status and biological subtype represent major determinants of treatment sequencing and can be readily included within standardized datasets and guideline-oriented prompts [2,18]. In HER2-positive and triple-negative breast cancer in particular, contemporary treatment pathways frequently rely on clearly defined biological and staging criteria [20-24].

The high concordance observed in this domain should therefore not necessarily be interpreted as evidence that AI possesses superior clinical reasoning. Rather, it may indicate that neoadjuvant treatment selection is particularly amenable to structured decision support because many of its determinants are explicitly represented within current treatment algorithms.

This interpretation is relevant to the broader application of artificial intelligence in clinical medicine. AI-based decision support may provide its most reproducible outputs when the clinical problem can be translated into clearly defined inputs and

when the relationship between those inputs and treatment recommendations is relatively standardized.

Planned surgery and surgical approach

Agreement decreased when the analysis moved from systemic treatment selection to operative planning. Planned surgery showed 70.0% observed agreement and a κ of 0.5814, while surgical approach showed 60.0% agreement and a κ of 0.4258. Both corresponded to moderate agreement.

This difference is clinically plausible. The choice between breast conservation and mastectomy depends on oncological factors but cannot be reduced to tumor stage or diameter alone. Tumor-to-breast volume ratio, disease distribution, breast morphology, feasibility of achieving adequate margins, anticipated cosmetic outcome, previous treatments, genetic risk and patient preference may all contribute to the final procedure [2,5].

Furthermore, the distribution of individual surgical recommendations was not identical between AI and clinical assessment. Mastectomy with reconstruction was proposed substantially more frequently by the AI workflow than by the clinical team, accounting for 43.3% and 20.0% of recommendations, respectively.

The published BREAST AI-PLAN study proposed that this finding might partly reflect a conservative bias within the information underlying the AI model [18].

However, other explanations are also possible. Structured textual data may incompletely represent the feasibility of breast conservation, particularly when oncoplastic techniques can permit wider resections than would be apparent from

tumor dimensions alone. Similarly, subtle elements of breast anatomy or surgeon experience may make a conservative approach feasible even when this cannot be fully encoded in the prompt.

The difference should therefore not be interpreted as demonstrating that either the AI or clinical recommendation was intrinsically superior. Rather, it highlights an important limitation of concordance studies in surgical decision-making: disagreement may reflect an AI error, incomplete representation of the clinical scenario, institutional practice, or the existence of more than one reasonable surgical strategy.

A similar consideration applies to surgical access. Although incision placement can be influenced by tumor location and planned procedure, the final choice frequently depends on breast shape, previous scars, nipple position, reconstructive requirements and the surgeon's anticipated tissue rearrangement. These factors are considerably more difficult to reproduce through structured textual variables alone.

Reconstructive planning

Reconstruction represented one of the domains with lower concordance, with an observed agreement of 57.14% and a κ of 0.3588.

This result is particularly relevant to the central hypothesis of the present thesis.

Reconstructive breast surgery combines objective variables with factors that remain highly dependent on direct clinical assessment. BMI, smoking status, previous radiotherapy, breast volume and subcutaneous tissue thickness can all be incorporated into structured data. However, skin elasticity, degree of ptosis,

mastectomy-flap vascularity, three-dimensional breast morphology, previous scars and patient expectations are more difficult to quantify.

Differences in the distribution of reconstructive recommendations support this interpretation. A definitive prepectoral implant was proposed considerably more frequently by AI than in clinical planning, whereas surgical flap reconstruction was more frequent among clinical recommendations.

These findings should again be interpreted cautiously. The study was not designed to determine which reconstructive strategy produced superior clinical outcomes, and no postoperative complication or patient-reported outcome data were used to validate the recommendations. Rather, the observed differences illustrate how reconstructive planning may depend on variables that were only partially represented in the structured dataset.

Direct-to-implant feasibility and implant characteristics were included among the predefined AI outputs, but separate concordance coefficients for these variables were not reported in the original analysis [18]. No independent conclusions regarding AI performance for these individual components can therefore be derived from the present dataset.

The relatively lower agreement for reconstruction is consistent with the broader concept that reconstructive decision-making remains strongly influenced by surgeon expertise and patient preference. Even among experienced surgeons, more than one reconstructive strategy may be reasonable for the same patient, particularly when oncological safety can be achieved using different technical approaches.

Adjuvant treatment

Adjuvant treatment showed an observed agreement of 64.29% but a κ of only 0.2911, corresponding to fair agreement.

This relatively low chance-corrected concordance contrasts with the almost perfect agreement observed for neoadjuvant therapy. One possible explanation is that postoperative systemic treatment can depend on information that may not be fully available during preoperative planning.

Definitive pathological tumor size, nodal burden, treatment response, margin status and, in selected patients, genomic assays may influence adjuvant recommendations.

When these variables are absent from a preoperative structured dataset, a model can only formulate recommendations on the basis of the information provided to it.

The finding therefore reinforces an important principle of AI-assisted clinical decision support: the quality and completeness of the input define the boundaries of the output. A recommendation generated from incomplete preoperative information should not be interpreted as equivalent to a definitive postoperative treatment plan.

Aesthetic assessment

Aesthetic assessment produced a different statistical pattern. Crude observed agreement was high, reaching 80.95% with linear weighting and 91.27% with quadratic weighting. However, expected agreement was also very high, resulting in weighted κ values of only 0.1418 and 0.1608, respectively.

This apparent discrepancy illustrates why percentage agreement alone may be misleading. Most assessments were concentrated in the higher aesthetic categories,

increasing the probability that the two classifications would agree by chance. Cohen's κ adjusts for this expected agreement and consequently provides a substantially more conservative estimate of concordance [45].

Beyond the statistical effect of category distribution, aesthetic outcome is inherently difficult to represent using text-based clinical data. Breast symmetry, scar position, nipple position, contour, tissue quality and patient perception contribute to the final cosmetic result. These factors are dynamic and three-dimensional and may change following healing and radiotherapy.

Furthermore, the aesthetic score used in BREAST AI-PLAN represented an expected preoperative assessment rather than direct independent analysis of postoperative photographic outcomes. The workflow did not perform autonomous computer-vision evaluation of breast images. Consequently, the present findings should not be interpreted as validation of AI-based postoperative cosmetic assessment.

Future integration of standardized photography, three-dimensional imaging or validated computer-vision tools may improve the ability of AI systems to evaluate aesthetic outcomes [7-9]. Nevertheless, the subjective and patient-centered component of aesthetic evaluation is unlikely to be completely replaced by an automated score.

Clinical significance of the observed concordance

Taken together, the results suggest a gradient between decisions that are predominantly guideline-driven and those requiring progressively greater anatomical, technical and subjective interpretation.

At one extreme, neoadjuvant treatment selection can often be derived from highly structured biological and staging information. At the other, reconstructive and aesthetic decisions depend on information that is difficult to reduce to standardized textual variables.

Planned surgery and surgical approach occupy an intermediate position, combining objective oncological information with direct anatomical assessment and operative expertise.

This gradient represents the most clinically relevant finding of the present study.

Rather than asking whether AI is globally “good” or “poor” at breast surgical planning, it may be more useful to identify **which components of the decision-making process are sufficiently structured to benefit from AI assistance and which still require substantial human interpretation.**

Cohen's κ measures agreement and reproducibility; it does not demonstrate clinical benefit, oncological safety or superiority of one decision over another. The present study did not evaluate margin positivity, reoperation, postoperative complications, recurrence, patient satisfaction or quality of life. Statistical significance of agreement should therefore not be equated with evidence that implementation of AI would improve patient outcomes.

Within these limitations, structured AI may nevertheless have potential applications as a clinical support tool. It could assist in systematic case synthesis, highlight relevant guideline-based variables, provide a standardized preliminary assessment before MDT discussion and support resident education [18]. These potential

applications remain hypotheses generated by the present workflow rather than outcomes directly tested in this study.

Strengths and limitations

The main strength of BREAST AI-PLAN lies in its structured approach. The same categories of clinical information were provided for each patient, the requested outputs were predefined, and AI-generated recommendations were directly compared with real-world multidisciplinary decisions. This design reduced some of the variability associated with completely open-ended interaction with generative AI [18].

However, several limitations should be acknowledged.

The cohort included only 30 patients from a single institution and should therefore be considered exploratory. The results may reflect local surgical practice and cannot be assumed to generalize to other breast units.

The study relied predominantly on structured textual information. The AI workflow did not independently interpret mammography, ultrasound, magnetic resonance imaging or three-dimensional breast morphology. Relevant information available to the surgeon through physical examination, including skin quality, ptosis, scars and tissue characteristics, could therefore be incompletely represented.

Manual transcription of clinical data into structured prompts also introduces the possibility of input or transcription bias [18].

In addition, generative AI outputs may depend on prompt wording, model characteristics and model version. Reproducibility therefore requires careful

documentation of the exact workflow, and results obtained with one AI configuration cannot automatically be extrapolated to other systems.

The multidisciplinary clinical decision used as the reference also has limitations.

MDT planning represents expert real-world practice but may incorporate institutional preferences and does not necessarily identify a single objectively correct treatment.

Consequently, disagreement with the reference should not automatically be classified as AI error.

Finally, the study evaluated concordance rather than clinical outcomes. No prospective assessment was performed to determine whether AI-assisted planning improves operative safety, oncological outcomes, cosmetic results, MDT efficiency or patient satisfaction.

Future perspectives

Future research should move from retrospective concordance studies toward prospective and multicenter validation. Larger cohorts would permit more precise estimation of agreement and identification of specific clinical scenarios in which AI recommendations are more or less reliable.

Multimodal systems integrating structured clinical information with imaging, pathology, genomic data, three-dimensional breast assessment and patient-reported outcomes may provide a more complete representation of the information used during surgical planning [18].

Future studies should also evaluate repeatability across different prompts and model versions, characterize clinically relevant disagreement rather than reporting

agreement alone, and compare performance across surgeons with different levels of experience.

Most importantly, clinical implementation should maintain a human-in-the-loop model. Artificial intelligence may assist with information organization and standardized reasoning, but final decisions in breast surgery require integration of evidence with anatomy, technical feasibility and patient values.

The present findings therefore support a role for AI as an adjunct to multidisciplinary breast cancer care rather than as a substitute for specialist clinical judgment.

7. CONCLUSIONS

The present study explored the potential role of prompt-driven artificial intelligence in breast cancer surgical planning by comparing AI-generated recommendations with multidisciplinary clinical decisions across different domains of treatment.

The results showed that concordance varied substantially according to the nature of the clinical decision. Agreement was highest for neoadjuvant treatment, a domain largely based on structured biological and staging criteria, while it progressively decreased for planned surgery, surgical approach and reconstructive planning.

Aesthetic assessment showed high crude agreement but only slight chance-corrected concordance.

These findings suggest that AI-assisted decision support may be particularly suitable for clinical questions that can be translated into explicit, standardized and reproducible variables. In contrast, decisions requiring direct anatomical evaluation,

technical expertise, assessment of tissue quality, aesthetic judgment and integration of patient preferences remain more difficult to reproduce through structured textual information alone.

BREAST AI-PLAN should therefore be considered a **clinical decision-support framework rather than an autonomous decision-making system**. Its potential value lies in organizing complex clinical information, supporting standardized reasoning, facilitating multidisciplinary case preparation and providing an educational tool for trainees [18]. Final surgical decisions must nevertheless remain under specialist multidisciplinary supervision.

The present findings should be interpreted within the limitations of a small, single-center, retrospective cohort and a predominantly text-based workflow. Concordance with multidisciplinary planning does not demonstrate superiority, clinical safety or improved patient outcomes.

Future studies should evaluate AI-assisted surgical planning prospectively in larger multicenter populations and incorporate multimodal information, including imaging, pathological and genomic data, three-dimensional anatomical assessment and patient-reported outcomes. Particular attention should also be given to reproducibility across models and prompts, transparency of recommendations and continuous clinical oversight.

Overall, this study suggests that the most realistic near-term role of artificial intelligence in breast surgery is not to replace surgical judgment, but to support it

where clinical information can be structured, standardized and critically reviewed within a human-led multidisciplinary pathway.

References

1. Bray F, Laversanne M, Sung H, Ferlay J, Siegel RL, Soerjomataram I, Jemal A. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2024 May-Jun;74(3):229-263. doi: 10.3322/caac.21834. Epub 2024 Apr 4. PMID: 38572751.
2. Loibl S, André F, Bachelot T, Barrios CH, Bergh J, Burstein HJ, Cardoso MJ, Carey LA, Dawood S, Del Mastro L, Denkert C, Fallenberg EM, Francis PA, Gamal-Eldin H, Gelmon K, Geyer CE, Gnant M, Guarneri V, Gupta S, Kim SB, Krug D, Martin M, Meattini I, Morrow M, Janni W, Paluch-Shimon S, Partridge A, Poortmans P, Puztai L, Regan MM, Sparano J, Spanic T, Swain S, Tjulandin S, Toi M, Trapani D, Tutt A, Xu B, Curigliano G, Harbeck N; ESMO Guidelines Committee. Electronic address: clinicalguidelines@esmo.org. Early breast cancer: ESMO Clinical Practice Guideline for diagnosis, treatment and follow-up. *Ann Oncol*. 2024 Feb;35(2):159-182. doi: 10.1016/j.annonc.2023.11.016. Epub 2023 Dec 13. PMID: 38101773.
3. Fisher B, Anderson S, Bryant J, Margolese RG, Deutsch M, Fisher ER, Jeong JH, Wolmark N. Twenty-year follow-up of a randomized trial comparing total

- mastectomy, lumpectomy, and lumpectomy plus irradiation for the treatment of invasive breast cancer. *N Engl J Med*. 2002 Oct 17;347(16):1233-41. doi: 10.1056/NEJMoa022152. PMID: 12393820.
4. Litière S, Werutsky G, Fentiman IS, Rutgers E, Christiaens MR, Van Limbergen E, Baaijens MH, Bogaerts J, Bartelink H. Breast conserving therapy versus mastectomy for stage I-II breast cancer: 20 year follow-up of the EORTC 10801 phase 3 randomised trial. *Lancet Oncol*. 2012 Apr;13(4):412-9. doi: 10.1016/S1470-2045(12)70042-6. Epub 2012 Feb 27. PMID: 22373563.
 5. Clough KB, Kaufman GJ, Nos C, Buccimazza I, Sarfati IM. Improving breast cancer surgery: a classification and quadrant per quadrant atlas for oncoplastic surgery. *Ann Surg Oncol*. 2010 May;17(5):1375-91. doi: 10.1245/s10434-009-0792-y. Epub 2010 Feb 6. PMID: 20140531.
 6. Biganzoli L, Marotti L, Hart CD, Cataliotti L, Cutuli B, Kühn T, Mansel RE, Ponti A, Poortmans P, Regitnig P, van der Hage JA, Wengström Y, Rosselli Del Turco M. Quality indicators in breast cancer care: An update from the EUSOMA working group. *Eur J Cancer*. 2017 Nov;86:59-81. doi: 10.1016/j.ejca.2017.08.017. Epub 2017 Sep 28. PMID: 28963914.
 7. Seth I, Bulloch G, Joseph K, Hunter-Smith DJ, Rozen WM. Use of Artificial Intelligence in the Advancement of Breast Surgery and Implications for Breast Reconstruction: A Narrative Review. *J Clin Med*. 2023 Aug 6;12(15):5143. doi: 10.3390/jcm12155143. PMID: 37568545; PMCID: PMC10419723.

8. Kenig N, Monton Echeverria J, Chang Azancot L, De la Ossa L. A Novel Artificial Intelligence Model for Symmetry Evaluation in Breast Cancer Patients. *Aesthetic Plast Surg*. 2024 Apr;48(7):1500-1507. doi: 10.1007/s00266-023-03554-1. Epub 2023 Aug 17. PMID: 37592148.
9. Gorgy A, Xu HH, Hawary HE, Nepon H, Lee J, Vorstenbosch J. Integrating AI into Breast Reconstruction Surgery: Exploring Opportunities, Applications, and Challenges. *Plast Surg (Oakv)*. 2024 Nov 6;34(1):37-46. doi: 10.1177/22925503241292349. PMID: 39545210; PMCID: PMC11559540.
10. Lee P, Bubeck S, Petro J. Benefits, Limits, and Risks of GPT-4 as an AI Chatbot for Medicine. *N Engl J Med*. 2023 Mar 30;388(13):1233-1239. doi: 10.1056/NEJMSr2214184. PMID: 36988602.
11. Pressman SM, Borna S, Gomez-Cabello CA, Haider SA, Haider CR, Forte AJ. Clinical and Surgical Applications of Large Language Models: A Systematic Review. *J Clin Med*. 2024 May 22;13(11):3041. doi: 10.3390/jcm13113041. PMID: 38892752; PMCID: PMC11172607.
12. Rengers TA, Thiels CA, Salehinejad H. Academic Surgery in the Era of Large Language Models: A Review. *JAMA Surg*. 2024 Apr 1;159(4):445-450. doi: 10.1001/jamasurg.2023.6496. PMID: 38353991.
13. Gomez-Cabello CA, Borna S, Pressman SM, Haider SA, Forte AJ. Large Language Models for Intraoperative Decision Support in Plastic Surgery: A Comparison between ChatGPT-4 and Gemini. *Medicina (Kaunas)*. 2024 Jun

- 8;60(6):957. doi: 10.3390/medicina60060957. PMID: 38929573; PMCID: PMC11205293.
14. De Pellegrin L, Weinzierl A, Kappos EA, Lindenblatt N, Zucal I, Harder Y. Evaluating artificial intelligence in decision-making for surgical treatment of benign breast conditions. *J Plast Reconstr Aesthet Surg*. 2025 Jun;105:189-195. doi: 10.1016/j.bjps.2025.03.057. Epub 2025 Apr 3. PMID: 40305890.
15. Ozmen BB, Almeida VFA, Mathur P, Schwarz GS. Explainable artificial intelligence: enhancing decision-making in plastic surgery. *J Plast Reconstr Aesthet Surg*. 2025 Sep;108:90-92. doi: 10.1016/j.bjps.2025.07.012. Epub 2025 Jul 16. PMID: 40716326.
16. Yan S, Li J, Wu W. Artificial intelligence in breast cancer: application and future perspectives. *J Cancer Res Clin Oncol*. 2023 Nov;149(17):16179-16190. doi: 10.1007/s00432-023-05337-2. Epub 2023 Sep 1. PMID: 37656245; PMCID: PMC11797688.
17. Kiwan O, Al-Kalbani M, Rafie A, Hijazi Y. Artificial intelligence in plastic surgery, where do we stand? *JPRAS Open*. 2024 Sep 14;42:234-243. doi: 10.1016/j.jpra.2024.09.003. PMID: 39435018; PMCID: PMC11491964.
18. Cossu M, Cuniolo L, Diaz R, Oliva M, Cornacchia C, Murelli F, Depaoli F, Gipponi M, Margarino C, Boccardo C, Franchelli S, Pesce M, Allievi R, D'Agraves AC, Abdallah S, Fregatti P. BREAST AI-PLAN: Prompt-Driven AI Assistance for Breast Surgery Planning - A Retrospective Single-center

Study. *In Vivo*. 2025 Nov-Dec;39(6):3271-3277. doi: 10.21873/invivo.14126.

PMID: 41167704; PMCID: PMC12588214.

19. Fisher B, Brown A, Mamounas E, Wieand S, Robidoux A, Margolese RG, Cruz AB Jr, Fisher ER, Wickerham DL, Wolmark N, DeCillis A, Hoehn JL, Lees AW, Dimitrov NV. Effect of preoperative chemotherapy on local-regional disease in women with operable breast cancer: findings from National Surgical Adjuvant Breast and Bowel Project B-18. *J Clin Oncol*. 1997 Jul;15(7):2483-93. doi: 10.1200/JCO.1997.15.7.2483. PMID: 9215816.
20. Cortazar P, Zhang L, Untch M, Mehta K, Costantino JP, Wolmark N, Bonnefoi H, Cameron D, Gianni L, Valagussa P, Swain SM, Prowell T, Loibl S, Wickerham DL, Bogaerts J, Baselga J, Perou C, Blumenthal G, Blohmer J, Mamounas EP, Bergh J, Semiglazov V, Justice R, Eidtmann H, Paik S, Piccart M, Sridhara R, Fasching PA, Slaets L, Tang S, Gerber B, Geyer CE Jr, Pazdur R, Ditsch N, Rastogi P, Eiermann W, von Minckwitz G. Pathological complete response and long-term clinical benefit in breast cancer: the CTNeoBC pooled analysis. *Lancet*. 2014 Jul 12;384(9938):164-72. doi: 10.1016/S0140-6736(13)62422-8. Epub 2014 Feb 14. Erratum in: *Lancet*. 2019 Mar 9;393(10175):986. doi: 10.1016/S0140-6736(18)32772-7. PMID: 24529560.
21. van Ramshorst MS, van der Voort A, van Werkhoven ED, Mandjes IA, Kemper I, Dezentjé VO, Oving IM, Honkoop AH, Tick LW, van de Wouw AJ, Mandigers CM, van Warmerdam LJ, Wesseling J, Vrancken Peeters MT, Linn SC, Sonke GS; Dutch Breast Cancer Research Group (BOOG). Neoadjuvant

- chemotherapy with or without anthracyclines in the presence of dual HER2 blockade for HER2-positive breast cancer (TRAIN-2): a multicentre, open-label, randomised, phase 3 trial. *Lancet Oncol.* 2018 Dec;19(12):1630-1640. doi: 10.1016/S1470-2045(18)30570-9. Epub 2018 Nov 6. PMID: 30413379.
- 22.von Minckwitz G, Huang CS, Mano MS, Loibl S, Mamounas EP, Untch M, Wolmark N, Rastogi P, Schneeweiss A, Redondo A, Fischer HH, Jacot W, Conlin AK, Arce-Salinas C, Wapnir IL, Jackisch C, DiGiovanna MP, Fasching PA, Crown JP, Wülfing P, Shao Z, Rota Caremoli E, Wu H, Lam LH, Tesarowski D, Smitt M, Douthwaite H, Singel SM, Geyer CE Jr; KATHERINE Investigators. Trastuzumab Emtansine for Residual Invasive HER2-Positive Breast Cancer. *N Engl J Med.* 2019 Feb 14;380(7):617-628. doi: 10.1056/NEJMoa1814017. Epub 2018 Dec 5. PMID: 30516102.
- 23.Schmid P, Cortes J, Dent R, Pusztai L, McArthur H, Kümmel S, Bergh J, Denkert C, Park YH, Hui R, Harbeck N, Takahashi M, Untch M, Fasching PA, Cardoso F, Andersen J, Patt D, Danso M, Ferreira M, Mouret-Reynier MA, Im SA, Ahn JH, Gion M, Baron-Hay S, Boileau JF, Ding Y, Tryfonidis K, Aktan G, Karantza V, O'Shaughnessy J; KEYNOTE-522 Investigators. Event-free Survival with Pembrolizumab in Early Triple-Negative Breast Cancer. *N Engl J Med.* 2022 Feb 10;386(6):556-567. doi: 10.1056/NEJMoa2112651. PMID: 35139274.
- 24.Masuda N, Lee SJ, Ohtani S, Im YH, Lee ES, Yokota I, Kuroi K, Im SA, Park BW, Kim SB, Yanagita Y, Ohno S, Takao S, Aogi K, Iwata H, Jeong J, Kim

- A, Park KH, Sasano H, Ohashi Y, Toi M. Adjuvant Capecitabine for Breast Cancer after Preoperative Chemotherapy. *N Engl J Med*. 2017 Jun 1;376(22):2147-2159. doi: 10.1056/NEJMoa1612645. PMID: 28564564.
25. Veronesi U, Cascinelli N, Mariani L, Greco M, Saccozzi R, Luini A, Aguilar M, Marubini E. Twenty-year follow-up of a randomized study comparing breast-conserving surgery with radical mastectomy for early breast cancer. *N Engl J Med*. 2002 Oct 17;347(16):1227-32. doi: 10.1056/NEJMoa020989. PMID: 12393819.
26. Early Breast Cancer Trialists' Collaborative Group (EBCTCG); Darby S, McGale P, Correa C, Taylor C, Arriagada R, Clarke M, Cutter D, Davies C, Ewertz M, Godwin J, Gray R, Pierce L, Whelan T, Wang Y, Peto R. Effect of radiotherapy after breast-conserving surgery on 10-year recurrence and 15-year breast cancer death: meta-analysis of individual patient data for 10,801 women in 17 randomised trials. *Lancet*. 2011 Nov 12;378(9804):1707-16. doi: 10.1016/S0140-6736(11)61629-2. Epub 2011 Oct 19. PMID: 22019144; PMCID: PMC3254252.
27. Moran MS, Schnitt SJ, Giuliano AE, Harris JR, Khan SA, Horton J, et al. Society of Surgical Oncology-American Society for Radiation Oncology consensus guideline on margins for breast-conserving surgery with whole-breast irradiation in stages I and II invasive breast cancer. *Ann Surg Oncol*. 2014;21(3):704-716.

28. Weber WP, Haug M, Kurzeder C, Bjelic-Radisic V, Koller R, Reitsamer R, Fitzal F, Biazus J, Brenelli F, Urban C, Paulinelli RR, Blohmer JU, Heil J, Hoffmann J, Matrai Z, Catanuto G, Galimberti V, Gentilini O, Barry M, Hadar T, Allweis TM, Olsha O, Cardoso MJ, Gouveia PF, Rubio IT, de Boniface J, Svensjö T, Bucher S, Dubsky P, Farhadi J, Fehr MK, Fulco I, Ganz-Blättler U, Günthert A, Harder Y, Hauser N, Kappos EA, Knauer M, Landin J, Mechera R, Meani F, Montagna G, Ritter M, Saccilotto R, Schwab FD, Steffens D, Tausch C, Zeindler J, Soysal SD, Lohsiriwat V, Kovacs T, Tansley A, Wyld L, Romics L, El-Tamer M, Pusic AL, Sacchini V, Gnant M. Oncoplastic Breast Consortium consensus conference on nipple-sparing mastectomy. *Breast Cancer Res Treat*. 2018 Dec;172(3):523-537. doi: 10.1007/s10549-018-4937-1. Epub 2018 Sep 4. PMID: 30182349; PMCID: PMC6245050.
29. Galimberti V, Morigi C, Bagnardi V, Corso G, Vicini E, Fontana SKR, Naninato P, Ratini S, Magnoni F, Toesca A, Kouloura A, Rietjens M, De Lorenzi F, Vingiani A, Veronesi P. Oncological Outcomes of Nipple-Sparing Mastectomy: A Single-Center Experience of 1989 Patients. *Ann Surg Oncol*. 2018 Dec;25(13):3849-3857. doi: 10.1245/s10434-018-6759-0. Epub 2018 Sep 17. PMID: 30225833.
30. Wei CH, Scott AM, Price AN, Miller HC, Klassen AF, Jhanwar SM, Mehrara BJ, Disa JJ, McCarthy C, Matros E, Cordeiro PG, Sacchini V, Pusic AL. Psychosocial and Sexual Well-Being Following Nipple-Sparing Mastectomy

and Reconstruction. Breast J. 2016 Jan-Feb;22(1):10-7. doi: 10.1111/tbj.12542. PMID: 26782950; PMCID: PMC4843778.

31. Zaborowski AM, Roe S, Rothwell J, Evoy D, Geraghty J, McCartan D, Prichard RS. A systematic review of oncological outcomes after nipple-sparing mastectomy for breast cancer. J Surg Oncol. 2023 Mar;127(3):361-368. doi: 10.1002/jso.27115. Epub 2022 Oct 8. PMID: 36208279.
32. Basta MN, Gerety PA, Serletti JM, Kovach SJ, Fischer JP. A Systematic Review and Head-to-Head Meta-Analysis of Outcomes following Direct-to-Implant versus Conventional Two-Stage Implant Reconstruction. Plast Reconstr Surg. 2015 Dec;136(6):1135-1144. doi: 10.1097/PRS.0000000000001749. PMID: 26595013.
33. Weber WP, Shaw J, Pusic A, Wyld L, Morrow M, King T, Mátrai Z, Heil J, Fitzal F, Potter S, Rubio IT, Cardoso MJ, Gentilini OD, Galimberti V, Sacchini V, Rutgers EJT, Benson J, Allweis TM, Haug M, Paulinelli RR, Kovacs T, Harder Y, Gulluoglu BM, Gonzalez E, Faridi A, Elder E, Dubsky P, Blohmer JU, Bjelic-Radisic V, Barry M, Hay SD, Bowles K, French J, Reitsamer R, Koller R, Schrenk P, Kauer-Dorner D, Biazus J, Brenelli F, Letzkus J, Saccilotto R, Joukainen S, Kauhanen S, Karhunen-Enckell U, Hoffmann J, Kneser U, Kühn T, Kontos M, Tampaki EC, Carmon M, Hadar T, Catanuto G, Garcia-Etienne CA, Koppert L, Gouveia PF, Lagergren J, Svensjö T, Maggi N, Kappos EA, Schwab FD, Castrezana L, Steffens D, Krol J, Tausch C, Günthert A, Knauer M, Katapodi MC, Bucher S, Hauser N, Kurzeder C, Mucklow R,

Tsoutsou PG, Sezer A, Çakmak GK, Karanlik H, Fairbrother P, Romics L, Montagna G, Urban C, Walker M, Formenti SC, Gruber G, Zimmermann F, Zwahlen DR, Kuemmel S, El-Tamer M, Vrancken Peeters MJ, Kaidar-Person O, Gnant M, Poortmans P, de Boniface J. Oncoplastic breast consortium recommendations for mastectomy and whole breast reconstruction in the setting of post-mastectomy radiation therapy. *Breast*. 2022 Jun;63:123-139. doi: 10.1016/j.breast.2022.03.008. Epub 2022 Mar 18. PMID: 35366506; PMCID: PMC8976143.

34. Davey MG, O'Donnell JPM, Boland MR, Ryan ÉJ, Walsh SR, Kerin MJ, Lowery AJ. Optimal localization strategies for non-palpable breast cancers -A network meta-analysis of randomized controlled trials. *Breast*. 2022 Apr;62:103-113. doi: 10.1016/j.breast.2022.02.004. Epub 2022 Feb 8. PMID: 35151049; PMCID: PMC8844725.
35. Veronesi U, Viale G, Paganelli G, Zurrida S, Luini A, Galimberti V, Veronesi P, Intra M, Maisonneuve P, Zucca F, Gatti G, Mazzarol G, De Cicco C, Vezzoli D. Sentinel lymph node biopsy in breast cancer: ten-year results of a randomized controlled study. *Ann Surg*. 2010 Apr;251(4):595-600. doi: 10.1097/SLA.0b013e3181c0e92a. PMID: 20195151.
36. Giuliano AE, Ballman KV, McCall L, Beitsch PD, Brennan MB, Kelemen PR, Ollila DW, Hansen NM, Whitworth PW, Blumencranz PW, Leitch AM, Saha S, Hunt KK, Morrow M. Effect of Axillary Dissection vs No Axillary Dissection on 10-Year Overall Survival Among Women With Invasive Breast

- Cancer and Sentinel Node Metastasis: The ACOSOG Z0011 (Alliance) Randomized Clinical Trial. JAMA. 2017 Sep 12;318(10):918-926. doi: 10.1001/jama.2017.11470. PMID: 28898379; PMCID: PMC5672806.
37. Galimberti V, Cole BF, Viale G, Veronesi P, Vicini E, Intra M, Mazzarol G, Massarut S, Zgajnar J, Taffurelli M, Littlejohn D, Knauer M, Tondini C, Di Leo A, Colleoni M, Regan MM, Coates AS, Gelber RD, Goldhirsch A; International Breast Cancer Study Group Trial 23-01. Axillary dissection versus no axillary dissection in patients with breast cancer and sentinel-node micrometastases (IBCSG 23-01): 10-year follow-up of a randomised, controlled phase 3 trial. Lancet Oncol. 2018 Oct;19(10):1385-1393. doi: 10.1016/S1470-2045(18)30380-2. Epub 2018 Sep 5. PMID: 30196031.
38. de Boniface J, Filtenborg Tvedskov T, Rydén L, Szulkin R, Reimer T, Kühn T, Kontos M, Gentilini OD, Olofsson Bagge R, Sund M, Lundstedt D, Appelgren M, Ahlgren J, Norenstedt S, Celebioglu F, Sackey H, Scheel Andersen I, Hoyer U, Nyman PF, Vikhe Patil E, Wieslander E, Dahl Nissen H, Alkner S, Andersson Y, Offersen BV, Bergkvist L, Frisell J, Christiansen P; SENOMAC Trialists' Group; SENOMAC Trialists' Group. Omitting Axillary Dissection in Breast Cancer with Sentinel-Node Metastases. N Engl J Med. 2024 Apr 4;390(13):1163-1175. doi: 10.1056/NEJMoa2313487. PMID: 38598571.
39. Bartels SAL, Donker M, Poncet C, Sauvé N, Straver ME, van de Velde CJH, Mansel RE, Blanken C, Orzalesi L, Klinkenbijn JHG, van der Mijle HCJ, Nieuwenhuijzen GAP, Veltkamp SC, van Dalen T, Marinelli A, Rijna H, Snoj

M, Bundred NJ, Merkus JWS, Belkacemi Y, Petignat P, Schinagl DAX, Coens C, van Tienhoven G, van Duijnhoven F, Rutgers EJT. Radiotherapy or Surgery of the Axilla After a Positive Sentinel Node in Breast Cancer: 10-Year Results of the Randomized Controlled EORTC 10981-22023 AMAROS Trial. *J Clin Oncol*. 2023 Apr 20;41(12):2159-2165. doi: 10.1200/JCO.22.01565. Epub 2022 Nov 16. PMID: 36383926.

40. Gentilini OD, Botteri E, Sangalli C, Galimberti V, Porpiglia M, Agresti R, Luini A, Viale G, Cassano E, Peradze N, Toesca A, Massari G, Sacchini V, Munzone E, Leonardi MC, Cattadori F, Di Micco R, Esposito E, Sgarella A, Cattaneo S, Busani M, Dessena M, Bianchi A, Cretella E, Ripoll Orts F, Mueller M, Tinterri C, Chahuan Manzur BJ, Benedetto C, Veronesi P; SOUND Trial Group. Sentinel Lymph Node Biopsy vs No Axillary Surgery in Patients With Small Breast Cancer and Negative Results on Ultrasonography of Axillary Lymph Nodes: The SOUND Randomized Clinical Trial. *JAMA Oncol*. 2023 Nov 1;9(11):1557-1564. doi: 10.1001/jamaoncol.2023.3759. PMID: 37733364; PMCID: PMC10514873.

41. Reimer T, Stachs A, Veselinovic K, Kühn T, Heil J, Polata S, Marmé F, Müller T, Hildebrandt G, Krug D, Ataseven B, Reitsamer R, Ruth S, Denkert C, Bekes I, Zahm DM, Thill M, Golatta M, Holtschmidt J, Knauer M, Nekljudova V, Loibl S, Gerber B. Axillary Surgery in Breast Cancer - Primary Results of the INSEMA Trial. *N Engl J Med*. 2025 Mar 13;392(11):1051-1064. doi: 10.1056/NEJMoa2412063. Epub 2024 Dec 12. PMID: 39665649.

42. Rubio IT, Marotti L, Biganzoli L, Aristei C, Athanasiou A, Campbell C, Cardoso F, Cardoso MJ, Coles CE, Eicher M, Harbeck N, Karakatsanis A, Offersen BV, Pijnappel R, Ponti A, Regitnig P, Santini D, Sardanelli F, Spanic T, Varga Z, Vrancken Peeters MJTFD, Wengström Y, Wyld L, Curigliano G. EUSOMA quality indicators for non-metastatic breast cancer: An update. *Eur J Cancer*. 2024 Feb;198:113500. doi: 10.1016/j.ejca.2023.113500. Epub 2023 Dec 26. PMID: 38199146.
43. Kesson EM, Allardice GM, George WD, Burns HJ, Morrison DS. Effects of multidisciplinary team working on breast cancer survival: retrospective, comparative, interventional cohort study of 13 722 women. *BMJ*. 2012 Apr 26;344:e2718. doi: 10.1136/bmj.e2718. PMID: 22539013; PMCID: PMC3339875.
44. Pons-Tostivint E, Daubisse-Marliac L, Grosclaude P, Oum Sack E, Goddard J, Morel C, Dunet C, Sibrac L, Lagadic C, Bauvin E, Bergé Y, Bernard-Marty C, Vaysse C, Lacaze JLL; EvaSein Group. Multidisciplinary team meeting and EUSOMA quality indicators in breast cancer care: A French regional multicenter study. *Breast*. 2019 Aug;46:170-177. doi: 10.1016/j.breast.2019.06.001. Epub 2019 Jun 7. PMID: 31226572.
45. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977 Mar;33(1):159-74. PMID: 843571.