level; however, ER loss seems to be caused primarily by transcriptional inactivation of the ER gene as a result of abnormal DNA methylation within the coding domain or promoter region [123, 124]. In addition, loss of critical transcriptional factors such as ERBF-1 may also be involved in some ER-negative phenotypes [124], as may be hypoxia, which has been shown to down-regulate ER expression as well as ER function in breast cancer cells [125]. As hypoxia stimulates the hypoxia-inducible factor(HIP)-1α signaling pathway and can upregulate vascular endothelial growth factor (VEGF), an important angiogenic molecule [126, 127], it is interesting to note that HIP-1α expression was negatively correlated with ER expression level [125], and that VEGF was inversely associated with ER and PR in breast cancer patients [128]. A recent study has revealed that estrogen can directly regulate VEGF expression through the ERE [129], indicating that estrogens and antiestrogens may play a role in the regulation of VEGF. Thus, patients treated with endocrine therapy might benefit from antiangiogenic therapy aimed at suppressing or neutralizing the production of VEGF.

Furthermore, several ER variants, including base-pair insertion, transition and deletion as well as alternative splicing, may result in outlaw functions such as dominant-negative, which is itself transcriptionally inactive but prevents the function of normal ER or dominant-positive, which is transcriptionally active even in the absence of estrogen [130]. As missense mutations of the ER are not common in primary breast cancer, alternative RNA splicing of the ER generating truncated forms of the receptor protein, seems to be important. For example, a deletion of exon 5, which involves the hormone-binding domain, was cloned from an ER-negative, PR-positive tumor [130, 131]. To date, various alternatively spliced forms of the ER have been reported, but further studies are needed to assess whether these variants are expressed as proteins or have any major biological role.

Recently, a second ER gene, ERβ, was identified from rat prostate [132], and therefore the classical ER was renamed as ERα. ERα and ERβ are the products of independent genes located at 6q15.1 and 14q22-24, respectively; however, they are highly conserved at the DBD (96%), and LBD (60%) with reduced homology between other parts of the molecules, especially at the N-terminus [133]. ERβ binds estrogens with a similar affinity to ERα and activates the expression of reporter genes containing ERE in an estrogen-dependent manner [132, 134]. On the other hand, ERβ lacks a large portion of the C-terminal region of the ERα [133], domain F, which is important both in modulating the magnitude of gene transcription by estrogens and antiestrogens, and in determining the effectiveness of antiestrogens in suppressing estrogen-stimulated gene transcription [135]. In certain cells,

antiestrogens are less effective if ERα lacks the F domain, and thus ERβ may contribute to resistance to antiestrogens by interfering with their inhibition of tumor growth.

ERα and ERβ differ significantly in their tissue distribution and physiological function [136, 137]. ERβ is widely distributed in the organism, including prostate, ovary, central nervous system, cardiovascular system, lung, kidney, urogenital tract, mammary gland, colon, and immune system. In normal human mammary gland, ERα expression is restricted to the cell nuclei of epithelial cells lining ducts and lobules, whereas ERβ is the predominant form, and additional strong expression of ERβ is detected in the cell nuclei of stromal cells including myoepithelial cells [138]. Under normal conditions, the ratio of ERα to ERβ in the breast depends on plasma concentrations of estradiol, and elevated expression of ERα (as occurs in postmenopausal women) is a normal response to loss of estradiol and indicates nonproliferating cells [139]. In contrast, ERα is usually the predominant form in most breast cancers, although both ERα and ERβ are present [140]. This may be important in some breast cancers, as ERα and ERβ can form heterodimers when both are present in the same cell, resulting in receptor activity distinct from that of ER homodimers [141-143]. Interestingly, the ratio of ERα/ERβ gene expression appears to alter during carcinogenesis [140], suggesting that ERα-specific and ERβ-specific pathways may have distinct role in this process. In addition, a decrease in ERβ expression has been reported in the transition from benign breast lesion to carcinoma *in situ* [144]. Thus, the relative expression level of ERα and ERβ will be a key determinant of cellular responses to agonists and antagonists [141].

With respect to the clinical significance of ERβ, controversies have arisen regarding the function of ERβ in breast cancer as an indicator of good or poor prognosis [140, 145-147]. These studies were conducted at the mRNA level, however, it has been reported that mRNA levels of ERβ do not correlate with ERβ protein level [148], and using IHC, ERβ showed strong association with PR expression and well-differentiated carcinoma [147, 149, 150]. Another study using IHC found a significant association between ERβ positivity and increased disease-free survival [151]. So far, the differences in study population and methodology (e.g. antibodies or isotypes) make it difficult to compare the results, however, it is interesting to note that a large and systematic study by Homma et al [152] has found an association between positive staining for ERβ1-3 or ERβ1 and survival in breast cancer patients treated with tamoxifen. In this study, ERβ1 positivity was also associated with better survival in patients with ERα negative/PR negative/HER2 negative (triple negative) tumors.

Among several variant forms of ERβ, ERβcx, which is truncated at the C-terminal region but has an extra 26 amino acids due to alternative splicing, shows very poor ability to bind to estrogen and has not or only low ERE binding ability [153]. Importantly, ERβcx is capable of heterodimerization with ERα and has a dominant negative effect on ERα function [153, 154]. Moreover, Saji et al [155] found that ERβcx might be one cause of the ERα positive/PR negative phenotype of cancer cells, and that ERβcx expression along with a low expression of PR correlated with poor response to tamoxifen.

In addition to activity at the ERE, the ER can mediate gene transcription through binding to AP-1 protein complexes at AP-1 enhancer sites. Interestingly, differential ligand activation of ERα and ERβ can occur at AP-1 sites [156]. At AP-1 site, ligands including 17β-estradiol, tamoxifen, and raloxifene activated transcription with ERα, whereas 17β-estradiol did not with ERβ. Moreover, tamoxifen and raloxifene were potent transcriptional activators with ERβ at the AP-1 site. On the other hand, at the ERE site, both ERα and ERβ showed the same activation profiles with these ligands. Thus, the two ERs signal in different ways depending on ligand and response element. These findings suggesting a possible role for ERβ at the AP-1 site warrant further investigations.

PR are ligand-activated transcription factors that act in concert with intracellular signaling pathways as "sensors" of multiple growth factor inputs to hormonally regulated tissues, such as the breast. The biological response to progesterone is mediated by two, nearly identical, forms of PR, PR-A and PR-B. These two isoforms are transcribed from distinct estrogen-inducible promoters within a single-copy PR gene, and the only difference between them is that the first 164 amino acids of PR-B are absent from PR-A. PR-A functions as a transcriptional repressor of progesterone-responsive promoters, whereas PR-B functions as a transcriptional activator of the same genes [157], indicating that their balance may affect endocrine response in breast cancer. To our knowledge, only a few studies have investigated the clinical significance of the expression of these PRs [158, 159]. In a study by Bamberger et al [158], expression of both isoforms was correlated with ER expression, and most breast tumors expressed PR-A at levels equal to or higher than those of PR-B [158]. Interestingly, PR-A expression at a level less than or equal to that of PR-B was associated with G1/G2 grading, whereas a PR-A expression greater than that of PR-B was associated with a more undifferentiated phenotype (G3 grading). Another study investigated PR-A and PR-B expression using immunoblot analysis in ER-positive, PR-positive patients with node-positive breast cancer, and found that tamoxifen-treated patients with high PR-A/PR-B ratios were more likely to relapse than

those with lower ratios [159]. Thus, PR-A/PR-B ratio may identify a subgroup that benefit poorly from endocrine therapy.

III. Multiparameter gene expression analysis

Gene expression profiling recently has been introduced into the clinical literature during the last decade. The molecular signatures has provided us with insight into the heterogeneity of breast cancer and hold promise for improving diagnosis, for the prediction of recurrence and in aiding selection of therapies for individual patients [160, 161]. For example, using complementary DNA microarray, Perou et al [162, 163] classified breast cancer into distinct sub-categories; luminal A (ER+, HER2-), luminal B (ER+, HER2+), HER2-overexpressing (ER-, HER2+), basal-like (ER-, HER2-, EGFR+, and/or cytokeratin 5/6+), and normal breast-like subgroups. Luminal subgroups share features with luminal epithelial cells arising from the inner layer of the duct lining, normal breast-like subgroup is characterized by high expression of basal epithelial-cell genes and low expression of luminal epithelial-cell genes. Basal-like subgroup shares features with normal breast basal epithelial cells, and often exhibits p53 mutation and low expression of BRCA1 tumor suppressor genes. This phenotype is common among BRCA1 carriers and sporadic triple-negative tumors. Importantly, when analyzed in a similarly treated group of patients, a significant difference in overall survival was found between these subtypes [163, 164].

Until now, several technologies have been developed to generate molecular signatures, including cDNA and oligonucleotide arrays and multiplex PCR technologies, and a number of studies specifically investigated molecular signatures in breast cancer, primarily focused on associations between particular sets of genes with altered expression and survival [161]. Here, we introduce several gene-expression profiling models including the 21-gene recurrence score (Oncotype DX), the 70-gene prognostic signature (MammaPrint).

The 21-gene recurrence score is a RT-PCR assay that measures the expression of 21 genes—sixteen cancer-related genes and five reference genes—in RNA extracted from formalin-fixed paraffin-embedded tissue samples from primary breast cancer [165]. These genes were selected from a much larger set of genes following the analysis of retrospective test sets of clinical material from several sources, including specimens from a cooperative group trial in which patients with ER-positive, node-negative breast cancer received tamoxifen versus tamoxifen plus chemotherapy (NSABP B-20) [166]. The levels of expression of the 21 genes are manipulated by an empirically derived, prospectively defined mathematical

algorithm to calculate an recurrence score (RS), which is then used to assign a patient to one of three groups by estimated risk of distant recurrence: low, intermediate, and high. The assay is intended to estimate risk of recurrence of patients with hormone receptor–positive, node-negative, stage I or II breast cancer. It has been suggested that tamoxifen-treated patients with an excellent estimated prognosis may be spared adjuvant chemotherapy. In addition, patients with a high RS appear to achieve a higher proportional benefit from adjuvant CMF/MF chemotherapy than those with low or intermediate RSs. This model was validated by the analysis of specimens and data from a second set of patients with node-negative, ER-positive breast cancer treated only with tamoxifen, who were enrolled in the NSABP clinical trial B-14 [167].

The 70-gene signature, so called Netherlands signature (MammaPrint) is a gene expression profiling platform, requiring a fresh sample of tissue that is composed of a minimum of 30% malignant cells [168]. In the development, primary tumors from patients with node-negative primary breast cancer were analyzed on oligonucleotide microarrays, and the data were subjected to supervised classification to establish a 70-gene RNA expression profile that correlated with a relatively short interval to distant metastases. The signature largely consists of genes regulating proliferation plus those involved in invasion, metastasis, stromal integrity, and angiogenesis. This assay was then validated in stage I, II primary breast cancer patients younger than age 53 years [168].

The Rotterdam 76-gene signature consists of a 76-gene microarray assay that does not overlap with either the 21-gene recurrence score or 70-gene signature [169, 170], and was specifically studied in node-negative breast cancer patients. The breast cancer gene expression ratio (two-gene expression ratio) is a quantitative RT-PCR–based assay that measures the ratio of the HOXB6 and IL17BR genes [171, 172]. This assay was developed to predict poor outcome in ER-positive patients treated with tamoxifen [173, 174]. The wound-response gene signature model is developed based on the hypothesis that fibroblast play a similar role in wound healing and tumor progression [175].

IV. Molecular detection of breast cancer cells

To date, morphological examination remains the gold-standard for cancer diagnosis. However, inaccurate diagnosis may result from inadequate preparation of the specimen, unskilled evaluation and differences in interpretation of morphological criteria. Therefore, it would be of great value to have a more sensitive method for cancer diagnosis that is less subject to evaluation bias. Metastatic spread of breast cancer cells occurs through the

blood and lymphatic circulation, and these cells settle in a specific organ, and develop into secondary tumors [176]. Clinically, micrometastasis may not be detectable at the time of diagnosis of primary breast cancer, even though it may have already occurred. This includes mini-residual and quiescent, dormant cells or stem cell [176, 177]. Currently, tumor metastasis is diagnosed by clinical manifestations and imaging studies, such as bone scans, chest radiography, liver ultrasonographys, or computed tomography scans, together with serum marker assays. However, clinical data are not sufficient to support the use of circulating tumor markers such as carcinoembryonic antigen (CEA) and carbohydrate antigen 15-3, p53 or circulating extracellular domain of HER2 for routine clinical decision-making [160]. Therefore, the ability to rule out the presence of micrometastasis at any stage of the clinical management protocol, whether before, during, or after therapy, would provide a useful monitoring and diagnostic tool.

In this respect, PCR and RT-PCR, which exponentially amplify a previously undetectable amount of nucleic acid to a detectable level, are promising high-sensitivity techniques. The use of these methods offers a number of advantages compared with cytological and IHC techniques, as both suffer from poor sensitivity [177, 178]. In addition, PCR-based methods can be designed to detect a cell using a specific target. Today, gene-expression profiling of single cells was reported to be feasible [179, 180].

1. Molecular diagnosis of breast cancer

To detect small numbers of cancer cells, not only in samples from fine needle aspiration or core needle, but also in nipple discharge or in samples that are cytologically undetermined to be malignant, PCR-based methods seem to be promising for accurate diagnosis. For example, one study investigated *HER2* mRNA in nipple discharge from breast cancer [181]. In this study, amplification of the *HER2* gene was more frequent in carcinoma *in situ* than in invasive types. However, mRNA in discharge may not be a suitable target for molecular detection, as it may be rapidly degraded in the collected fluid. Therefore, the best approach for nipple discharge seems to be PCR. In a recent study using PCR for seven markers of LOH, at least one LOH was found in either supernatant or cell pellet in 70% of breast cancer patients with nipple discharge [182]. This method showed higher sensitivity than smear cytology or quantification of CEA concentration of the discharge. However, some papillomas also showed LOH [183]. In a similar study using methylation-specific PCR, methylated alleles of the *cyclin D2*, *RAR-β* and *Twist* genes were frequently detected in cells collected from breast ducts, but rarely in ductal lavage fluid from healthy ducts [184]. These genes are

hypermethylated in more than 30% of breast cancers, but unmethylated in healthy mammary epithelial cells, mammary stroma, and white blood cells [185-187]. In association with this, it is interesting to note that in the study by Fackler et al [188], the quantitative multiplex methylation-specific PCR doubled the sensitivity of detection of cancer compared with cytology. In addition, telomerase detection in cells derived from fine needle aspirates of breast showed high sensitivity and specificity for cancer detection [73, 189]. Among the several gene alterations in cancer, telomerase is unique as it is detected in almost all kinds of cancer tissues with a very high positive rate. Thus, molecular analysis of nipple fluid appears to be promising, not only for diagnosis of breast cancer but also for screening of high risk women [187, 190].

Increased quantities of DNA or nucleosomes have been found in the plasma or serum of breast cancer patients [191-200] (Table 3). Interestingly, this circulating extracellular DNA may exhibit tumor-related alterations. PCR for LOH, p53 mutation, and methylation of the first exon of $p16^{INK4a}$ found that 66% of breast cancer patients had molecular alterations in plasma DNA [192]. In a similar study using different markers for microsatellite instability and LOH, two studies independently demonstrated that plasma or serum DNA of breast cancer patients displayed the same gene alterations as those found within the primary tumor [191, 193]. Of clinical importance, the presence of plasma DNA of tumor origin correlated with pathological parameters associated with a poor prognosis, histological features of highly malignant lesions [192]. Moreover, circulating DNA was shown to undergo quantitative changes in cancer patients after radiation or chemotherapy [194-196]. Thus, DNA derived from plasma or serum could be used for molecular diagnosis of cancer, and monitoring the disease course [197, 201]. Moreover, it is interesting to note that telomerase RNA (hTERT) could be detectable in plasma from cancer patients [202].

2. Clinical significance of micrometastasis in breast cancer

In contrast to hematological malignancies, solid tumors rarely have specific genetic changes. To overcome this limitation, the majority of studies have used RT-PCR for molecular detection of micrometastasis [177, 203]. Table 4 lists the molecular markers used most frequently to detect micrometastasis in breast cancer. The major tissues of interest for molecular detection of breast cancer micrometastasis and disseminated or circulating tumor cells (DTCs, CTCs, respectively) are LNs [204-249], BM [193, 213, 245, 250-288] and PB [193, 207, 212, 218, 228, 235, 242, 245, 250, 254, 255, 257, 258, 260-273, 276, 278, 282, 285, 287, 289-352]. In addition, cancer cells are detected in body fluids such as cerebrospinal fluid (CSF) and pleural

Table 3. Representative studies of circulating extracellular DNA.

Author ,Year	Targets	Sample source	Positive rates in breast cancer		Positive rates in controls	Remarks
			Primary, NMD	Metastatic disease		
RT-PCR based analysis of plasma/serum sample						
Chen, 1999 [191]	LOH	P, S	18.4%*(49)	66.7%* (3)	0%* (>10[1])	Useful
Silva, 1999 [192]	LOH	P	61% (38)	-	0% (13)	Identical to tumor DNA
	MI	P	5% (38)	-	0% (13)	Not present in tumor DNA
	Methylation of p16 ^{NK4a}	P	14% (43)	-	0% (3)	
	p53	P	5% (62)	-	0% (5)	
Shaw, 2000 [193]	LOH	P	15% (39)	63% (16)	0% (9)	Sensitive
	MI	P	10% (39)	38% (16)	0% (9)	Sensitive
Rykova, 2004 [200]	Methylation[2]	P	33-47%(17)		0-1%(6)	Promising
RIA, ELISA based analysis of plasma/serum sample						
Leon, 1977 [194]	Free DNA(RIA)	S	Increased**		7% (55)	Increased, especially in metastatic diseases.
Holdenrieder, 1999 [195]	Nucleosome (ELISA)	S	84% (43)***		4% (50)	Sensitive, useful for monitoring
Kuroi, 2001 [196]	Nucleosome (ELISA)	P	68.7% (99)	84.6% (26)	4.5% (111)	Predictive of response

Abbreviations: ELISA, enzyme linked immunosorbent assay; LOH, loss of heterozygosity; MI, microsatellite instability; NMD, Non metastatic disease; P, plasma; RIA, radioimmunoassay; S, serum.

() Number of patients.

1) At least 10 controls/microsatellite markers were investigated.

2) APC, RASSF1A, RAFβ, CDH1, CDH13 gene promoter methylation.

*Percentage of positive cases for at least one locus, **Metastatic disease was included, ***Breast cancer, NOS.

Table 4. Molecular markers frequently used in breast cancer.

Molecular markers	Comments
Carcinoembryonic antigen (CEA)	Glycoprotein molecule with an oncofetal expression pattern
Cytokeratin 19	An epithelial marker for detection of cancer cells in mesenchymal organs by RT-PCR, as the type of cytoplasmic intermediate filaments primarily exist in epithelial cells and cancer cell derived from epithelia, but is normally not expressed in cells of non-epithelial origin
MUC1	Mucins are the major protein constituents of the intestinal mucous gel which coats and protects epithelial surface. Among several different mucins, MUC1, a membrane-bound form, is a core protein of polymorphic epithelial mucin, and it is uniformly and highly expressed in epithelial tumors including breast cancer, and rarely expressed in nonepithelial tissues
Maspin	Homologous to the serine superfamily of protease
Mammaglobin (MGB/MGB1)	A mammary specific member of the uteroglobin gene family, and encode a glycoproteins. The human MGB gene is localized on 11q12.3-13.1. In non-malignant tissues, expression is strictly limited to the mammary epithelium, and is highly expressed in primary breast cancer cells
Mammaglobin B (MGB2)	A recently cloned member of uteroglobin gene family, and highly homologous to mammaglobin

Table 5. Representative studies of micrometastases in lymph nodes.

Author, Year	Targets	Positive rates in breast cancer patients		Positive rates in LNs from controls	Remarks
		Histologically positive LNs	Histologically negative LNs		
Noguchi, 1994 [204]	MUC1	100% (9)	14.6% (41)	0% (10)	Sensitive
Schoenfeld, 1994 [205]	CK19	100% (18)	10% (39)	0% (11)	By ethidium brimide staining
	CK19	-	28% (39)	0% (11)	By Southern blotting and hybridization
Mori, 1995 [206]	CEA	100% (6)	0% (5)	0% (5)	Included various types of cancer
Hoon, 1996 [207]	βhCG	90% (10)	25% (8)	0% (8)	Sensitive and specific
Noguchi, 1996 [208]	CK19	90% (20*)	14 .2% (106*)	-	Useful to detect micrometastasis
Noguchi, 1996 [209]	CK19	100% (10)	9% (53)	0% (10)	Useful to detect micrometastasis
	MUC1	100% (10)	6% (53)	0% (10)	Useful to detect micrometastasis
Schoenfeld , 1996 [210]	CK19	-	20% (530)	0% (28)	Improve detection of micrometastasis
Lockett, 1997 [211]	Multimarker [1]	94% (16)	48% (29)	-	Useful
Bostick, 1998 [212]	CEA	70%(10)	42% (12)	100% (3)	No diagnostic value
	CK19	80% (10)	67% (12)	100% (3)	No diagnostic value
	CK20	20% (10)	8% (12)	0% (3)	Low sensitivity, high specificity
	GA733.2	70% (10)	92% (12)	100% (3)	No diagnostic value
	MUC1	70% (10)	83% (12)	100% (3)	No diagnostic value

Table 5. Continued

Gerhardt, 1998 [213]	HER2[2]	83.3% (6)	-	-	
Lockett, 1998 [214]	Multimarker[3]	-	40% (35)	-	Useful
Lockett, 1998 [215]	Multimarker[3]	92% (24*)	40% (37*)	0% (9*)	Useful
Aihara, 1999 [216]	MGB2	100% (11)	31% (45)	0% (15)	Useful to detect micrometastasis
Merrie, 1999 [217]	CK8	81% (313)	-	96% (54)	No diagnostic value
	CK16	23% (313)	-	18% (146)	No diagnostic value
	CK19	71% (313)	-	67% (61)	No diagnostic value
	Maspin	55% (90)	-	48% (113)	No diagnostic value
	MUC1	18% (159)	-	49% (90)	No diagnostic value
	ST3	70% (342)	-	64% (64)	No diagnostic value
Watson, 1999 [218]	CK19	100% (11*)	100% (3*)	-	Less specific
	MGB1	91% (11*)	0% (3*)	-	Promising than CK19
Leygue, 1999 [219]	MGB1	100% (13)	0% (7)	-	Correlated with H &E staining
Masuda, 2000 [220]	CEA	-	31.0% (129*)	-	Correlated with prognosis
Kataoka, 2000 [221]	CEA	100% (17)	25% (48)	0% (30)	Improve detection of micrometastasis
	MGB1	70.6% (17)	20.8% (48)	0% (30)	Improve detection of micrometastasis
Ooka, 2000 [222]	CEA	35.7% (14)	1.2% (163)	0 (16)	Less sensitive
	CK20	-	-	0% (16)	Not suitable because of low expression in tumor
	MAGE1	-	-	0% (16)	Not suitable because of low expression in tumor

Table 5. Continued

Ooka, 2000 [222]	CEA	35.7% (14)	1.2% (163)	0 (16)	Less sensitive
	CK20	-	-	0% (16)	Not suitable because of low expression in tumor
	MAGE1	-	-	0% (16)	Not suitable because of low expression in tumor
	MAGE3	-	-	0% (16)	Not suitable because of low expression in tumor
	MGB1	100% (14)	30.1% (163)	0% (16)	Specific
	MGB2	100% (14)	17.8% (163)	0% (16)	Specific
	PSA	57.1% (14)	43.7% (163)	0% (16)	Less sensitive
Mitas, 2001 [223]	Multimarker[4]	-	38% (21)		Applicable as a panel
Branagan, 2002 [224]	MGB1	17.2% (29)	8% (119)	-	
Mitas, 2002 [225]	CEA	-	4.5% (22)	-	
	PIP	-	27.3% (22)	-	
	PSE	82% (22)	14% (22)	2% (51)	Informative

Abbreviations: CEA, carcinoembryonic antigen; CK, cytokeratin; GA733.2, gastrointestinal tumor-associated antigen-733.2; βhCG, β human dhorionic gonadotropin; LN, lymph node; MGB1, mammaglobin; MGB2, mammaglobin B; PIP, prolactin inducible protein; PSA, prostate-specific antigen; PSE, prostate-specific Ets; ST3, stromelysin 3.

() Number of LNs.

1) Combination of 4 genes including CK19, c-myc, HER2, PIP.

2) Alternatively spliced HER2 was used as target.

3) Combination of CK19, c-myc, PIP.

4) Combination of MGB1, MGB2, CK19, MUC1, CEA, HER2, PIP.

* Number of patients.

Table 6. Representative studies of micrometastases in bone marrow.

Author, Year	Targets	Positive rates in breast cancer		Positive rates in controls	Remarks
		Primary, NMD	Metastatic disease		
Datta, 1994 [250]	CK19	-	75% (8)	6.7% (30)	Sensitive and specific
Gerhard, 1994 [251]	CEA	66.7% (6)*		0% (56)	
Brown, 1995 [252]	DF3	11.1% (9)	-	0% (4)	
	CK18	-	-	100% (4)	Not usuful
Fields, 1996 [253]	CK19	55% (33)	82% (50)	3.7% (214)	Sensitive and associated with prognosis
Krüger, 1996 [254]	CK19	35% (20)	75% (4)	0% (5)	Sensitive and specific
Luppi, 1996 [255]	Maspin	-	Pos**	0% (4)	Sensitive and specific
Moscinski, 1996 [256]	CK19	13.3% (30)	73% (30***)	3.8% (52)	Highly sensitive and specific
Schoenfeld, 1997 [257]	CK19	35% (65)	80% (5***)	-	Improve detection of micrometastasis
Wulf, 1997 [258]	PTHrP	26.5% (34)	-	12% (25)	Less specific
Zippelius, 1997 [259]	CEA	-	-	26.3%	
	CK18	-	-	71.4%	Not useful
	DPII	-	-	100% (5)	Not useful
	EGP-40	-	-	100% (53)	Not useful
	HER2	-	-	71.4% (7)	Not useful
	HER3	-	-	85.7% (7)	Not useful
	PSA	0% (10)	-	0% (53)	
	PSM	-	-	44.4% (9)	Not useful
Gerhardt, 1998 [213]	HER2[1]	-	100% (6)	-	
Vannucchi, 1998 [260]	CK19	48% (33)	-	13.9% (43)	Correlated with positivity in PBST
Shammas, 1999 [261]	CK19	25% (12)	-	-	Sensitive, but overlapping

Table 6. Continued

Slade, 1999 [262]	CK19	83% (23)		-	60% (30)	Highly sensitive, more likely to be positive in BM
Zhong, 1999 [263]	CEA	27.6% (181)		-	0% (8)	Limitation due to heterogeneity
Zhong, 1999 [264]	GA733.2	100% (33)		-	100% (8)	Not specific
Zhong, 1999 [265]	CK19	40.9% (115)		-	0% (8)	
Berois, 2000 [266]	CEA	29% (42)		-	-	More sensitive in BM than in PB
	CK19	48% (42)		-	-	More sensitive in BM than in PB
Ikeda, 2000 [267]	CK19	34.2% (117)		-	-	Prognostic value
Shaw, 2000 [193]	CK19	56% (32)		-	-	
Zhong, 2000 [268]	CK19[2)]	69.2% (26)		-	0% (8)	Sensitive and specific
Ooka, 2001 [269]	MGB	29.7% (111)		-	0%	Correlated with prognosis
Shammas, 2001 [270]	CK19	20.6% (141)		-	4% (48)	Serial sampling may be useful
Silva, 2001 [271]	CK20		100% (1)*		100% (2)	Not reliable
Silva, 2002 [272]	MGB	40% (5)		86% (7)	0% (15)	Sensitive and specific
Stathopoulou, 2002 [273]	CK19	63% (27)		74% (46)	-	

Abbreviations: CEA, carcinoembryonic antigen; CK, cytokeratin; DF3, human breast carcinoma-assocaited antigen; DPI, desmoplakin I; EGP, epithelial glycoprotein; GA733.2, gastrointestinal tumor-associated antigen-733.2; MGB, mammaglobin; NMD, non metastatic disease; PBST, peripheral blood stem cell; PSA, prostate-specific antigen; PSM, prostate-specific membrane antigen; PTHrP, parathyroid hormone related protein.

() Number of patients.

1) Alternatively spliced HER2 was used as target.

2) Cells were separated using immunomagnetic beads.

* Breast cancer, NOS, ** Positive in 3 BM, NOS, *** Patients with bone marrow metastasis.

Table 7. Representative studies of circulating tumor cells in peripheral blood.

Author, Year	Targets	Sample source	Positive rates in breast cancer		Positive rates	Remarks
		Patients : Controls	Primary, NMD	Metastatic disease	in controls	
Datta, 1994 [250]	CK19	B, PBSC : B	0% (8)	21.1% (19)	2.5% (39)	Sensitive and specific
Hoon, 1996 [207]	βhCG	B : B	80% (10)**		0% (28)	
Krüger, 1996 [254]	CK19	PBSC : B	37.5% (16)	60% (10)	0% (10)	Sensitive and specific
Krüger, 1996 [289]	CK19	PBSC : -	40% (10*)	-	-	Mobilized by GCSF
Lehrer, 1996 [290]	PSA	B : B	23.1% (78)	-	0% (42)	Useful
Luppi, 1996 [255]	CK19	B : B	33.3% (9)	25.0% (20)	29.4% (17)	
	Maspin	B : B	0% (9)	20% (20)	0% (17)	Sensitive and specific
Mori, 1996 [291]	CEA	B : B	25% (8)	100% (1)	0% (22)	Useful
Mapara, 1997 [292]	EGFR	PBSC : B	-	58.8% (17*)	-	Better than CK19
Schoenfeld, 1997 [257]	CK19	B : B	25% (75)	80% (5)	-	Improve detection of micrometastasis
Traystman, 1997 [293]	CK19	B : B	25% (20)	63.6% (11)	0% (22)	Sensitive
Wulf, 1997 [258]	PTHrP	B : B	30% (30)	-	0% (30)	Very sensitive
Bostick, 1998 [212]	CEA	B : B	NA	NA	46% (13)	High false positive
	CK19	- : B	-	-	77% (13)	High false positive
	CK20	B : B	-	-	0% (13)	
	GA733.2	B : B	-	-	54% (13)	High false positive
	MUC1	- : B	-	-	46% (13)	High false positive
Leitzel, 1998 [294]	EGFR	B : B	0% (13)	22% (18)	0% (23)	Sensitive and specific
Luke, 1998 [295]	CK19	B-IM : B-IM	-	67% (15)	0% (39)	More sensitive than GCDFP
	GCDFP	B-IM : B-IM	-	27% (15)	13% (39)	Less sensitive than CK19
Vannucchi, 1998 [260]	CK19	PBSC : PBSC	57.5% (33)	-	3.3% (30)	Mobilized by GCSF

Table 7. Continued

Lambrechts, 1999 [296]	CK19	B : B	-	94.3% (35)	95% (20)	Not suitable
Shammas, 1999 [261]	CK19	B : -	-	25% (12)	-	Sensitive, but overlapping
Slade,1999 [262]	CK19	B : B	70% (23)	76% (37)	51% (45)	Highly sensitive
Soria, 1999 [297]	Telomerase	B-IM : B-IM	-	84% (25)	0% (9)	Useful
Watson, 1999 [218]	MGB	PBSC : -	-	60% (15)	-	
Zach, 1999 [298]	MGB	B : B	28% (18)	49% (43)	0% (27)	Useful
Zhong, 1999 [263]	CEA	- : B	-	-	0% (96)	
Zhong, 1999 [264]	GA733.2	- : B	-	-	40% (40)	Not specific
Zhong, 1999 [265]	CK19	PBSC : B	7.7% (13)	37.5% (16)	0% (96)	
Berois, 2000 [266]	CEA	B : B	3% (37)	-	0% (20)	More sensitive in BM than in PB
	CK19	B : B	35% (37)	-	0% (20)	More sensitive in BM than in PB
	MUC1	- : B	-	-	60% (20)	High false positive
de Cremoux, 2000 [299]	MUC1	B-IM : B-IM	43.3% (97)	52.2% (23)	11% (28)	Sensitive and specific
Grünewald, 2000 [300]	CK19	B : B	48% (133)	-	39% (31)	Not specific
	EGFR	B : B	10% (133)	-	25% (31)	
	MGB	B : B	8% (133)	-	0% (31)	Specific marker for spread of tumor
Ikeda, 2000 [267]	CK19	- : B	-	-	0% (15)	
Kahn, 2000 [301]	CK19	B : B	34.6% (81)	71.4% (28)	0 (45)	Correlated with prognosis
Sabbatini, 2000 [302]	Maspin	B : B	4.8% (21)	0% (8)	-	Mobilized by chemotherapy
Shaw, 2000 [193]	CK19	B : -	-	67% (18)	-	Sensitive
Smith, 2000 [303]	CK19	B : -	-	49.6% (145*)	-	Correlated with response
Zhong, 2000 [268]	CK19	- : B	-	-	0% (117)	
Aerts, 2001 [304]	CK19	B : B	31.6% (19)	71.4% (14)	8.7% (23)	
An, 2001 [305]	CEA	B : B	38% (32)	-	0% (11)	Correlated with prognosis

Table 7. Continued

Houghton, 2001 [306]	MGB	B : B	-	62.5% (32)	0% (11)	Microarray study
	Multimarker[1]	B-IM : B-IM	-	84.4% (32)	0% (11)	
Hu, 2001 [307]	βhCG	B : B	12.4% (72)**		0% (30)	Correlated with stage
	CK19	B : B	9.7% (72) **		10% (30)	
	CK20	B : B	2.8% (72) **		0% (30)	Not useful
Matin, 2001 [308]	Cluster[2]	B : B	77% (13)	-	19% (15)	
Ooka, 2001 [269]	MGB	- : B	-	-	0%	
Silva, 2001 [271]	CK20	B : B	100% (2)***		74% (31)	Not reliable
		PBSC : PBSC	100% (2)***		80% (5)	Not reliable
Shammas, 2001 [270]	CK19	B : B	1.3% (74)	-	4% (48)	Not useful
Silva, 2002 [272]	MGB	B : B	22% (65)	38% (13)	0% (47)	Sensitive and specific
Stathopoulou, 2002 [273]	CK19	B : B	29.7% (148)	52% (46)	3.7% (54)	Correlated with prognosis

Abbreviations: B, Blood; B-IM, immunomagnetic separation of cancer cells from blood; CEA, carcinoembryonic antigen; CK, cytokeratin; EGFR, epidermal growth factor receptor; GA733.2, gastrointestinal tumor-associated antigen-733.2; GCDFP, gross cystic disease fluid protein; βhCG, β human chorionic gonadotropin; LOH, loss of heterozygosity; MGB, mammaglobin; MI, microsatellite instability; NMD, non metastatic disease; PBSC, peripheral blood stem cell; PSA, prostate-specific antigen ; PTHrP, parathyroid hormone related protein.

() Number of patients.

1) Combination of MGB1, B305D, B311D, B533S, B726P, GABAπ.

2) A group of SRP19, CD44, TRP-2-8b, Maspin, HSIX1, Gro α, Myosin light chain, MDM2, ZZ38, β–tubulin, N33, Laminin α3.

* Number of patients or samples, ** Metastatic disease was included, *** Breast cancer, NOS.

Table 8. Representative studies of disseminated tumor cells in pleural effusion, cerebrospinal fluid.

Author, Year	Targets	Positive rates in breast cancer according to cytological diagnosis			Positive rates in control	Remarks
		Positive	Negative	Suspected		
Cerebrospinal fluid						
Datta, 1994 [250]	CK19	100% (1)	-	-	0% (2)	Detectable
Pleural effusion						
de Matos Granja, 2002 [353]						
	LOH	38% (24)*	0% (3)*	36% (11)*	-	Useful

Abbreviations: CK, cytokeratin; LOH, loss of heterozygosity.

() No of patients.

* Percentage of positive cases for at least one locus.

effusion of patients with breast cancer [250, 353-357], or axillary drainage fluid after surgery [358-361]. The representative studies at the dawn of a new golden age of molecular detection and characterization of micrometastasis, DTCs and CTCs in breast cancer are listed in table 5-8.

LN metastasis is usually determined by histologic analysis of one or a few hematoxylin and eosin sections from each LN, and axillary LN status is the most powerful predictor of patient outcome. However, patients without axillary LN metastasis are not completely devoid of risk for relapse. Therefore, considerable efforts have been made to find prognostic markers such as HER2, p53 and cathepsin D in breast cancer tissue, although their clinical significance is less conclusive because of limited and conflicting data. Another concern is false negative pathological findings. In this respect, a study of particular interest indicated that the serial sectioning technique revealed micrometastases in 9% of breast cancer patients who were diagnosed as node-negative by routine histological examination, and these patients had a poor disease-free and overall survival [362]. In agreement with this study, several studies have shown that the serial sectioning technique with or without IHC can detect micrometastases [363-366]. Thus, routine histological examination may underestimate the true incidence of metastasis. Unfortunately, the serial section technique is not practical as a routine method, as it is cumbersome and time-consuming. On the other hand, RT-PCR based methods are highly sensitive, require relatively less time to perform and are cost effective [214, 367]. So far, several studies have demonstrated that micrometastases could be detected at a significant frequency in histologically negative LNs, and it is interesting to note that micrometastases in LNs detected by RT-PCR for CEA were associated with reduction of both disease-free survival and overall survival [220].

Recent studies have shown that the sentinel lymph node (SLN), which is the first drainage LN from the site of the breast cancer, can be used to predict nodal status [368-370]. The fundamental concept underlying sentinel node mapping is that the lymphatic effluent of a tumor drains initially to a sentinel node (or to a few SLNs) before other nodes in the group receive tumoral drainage. SLN biopsy may allow patients with breast cancer to avoid the morbidity of formal axillary clearance [370, 371]. For intraoperative assessment, the recent European Working Group for Breast Screening Pathology guidelines advocate multilevel assessment of grossly or intraoperatively negative sentinel nodes with levels separated by a maximum of 1 mm [372]. In this respect, several reports have suggested that RT-PCR based methods may be useful for more accurate diagnosis of metastasis in SLN [221, 222, 224, 234, 237, 241, 244, 246-248, 373]. For perioperative

diagnosis, the rapid molecular test assay appears to be a favorable method for analysis [367, 374, 375].

Detection of BM micrometastases may have a role in monitoring treatment, predicting prognosis, and understanding tumor biology. At the time the primary breast cancer is diagnosed, metastatic bone lesions are usually beyond detection by conventional methods such as radiography, or skeletal scintigraphy. Furthermore, conventional histological methods could find cancer cells in only a few cases. However, the addition of immunohistological methods has led to improvements in the ability to detect breast cancer cells in BM aspirates or biopsy samples. Importantly, immunologic detection of epithelial cells in the BM correlated with a significantly shorter disease-free survival in breast cancer [177, 376-380]. Similarly, BM micrometastases detected by RT-PCR for mammaglobin (MGB) were correlated with early distant recurrence of breast cancer [269]. Moreover, one study using RT-PCR for cytokeratin (CK) 19 indicated that the presence of BM micrometastases was associated with a high risk of relapse in patients with proven metastatic breast cancer undergoing BM transplantation [253]. The detection of occult carcinoma in BM or PB stem cell (PBSC) collections may be of special concern in the setting of autologous BM transplantation for breast cancer patients, as high dose chemotherapy or granulocyte-colony stimulating factor may mobilize cancer cells into circulation [260, 381]. The possibility of contaminated PBSC collections was significantly higher in patients with CK19 positivity in BM at diagnosis, and there was a trend towards longer relapse free survival in patients transplanted with CK19-negative PBSC collections as compared with the others [260]. Similarly, conventional-dose chemotherapy caused cancer cell mobilization into the circulation, and the presence of circulating maspin positive cells was associated with disease progression [302].

On the other hand, obtaining serial blood samples is much easier and less invasive than the process involved in obtaining BM aspirates or biopsy samples. Currently available IHC methods are not sensitive enough to detect low numbers of CTCs in PB [257], but the introduction of RT-PCR techniques have opened up the way to a more sensitive and meaningful analysis of CTCs in PB [250, 291, 382]. Despite the relatively recent introduction of this methodology, several studies have demonstrated preliminary but promising data. For instance, in patients with gastrointestinal or breast cancer, the detection rates for CEA mRNA in PB increased with advancing stage of disease; moreover, in patients who underwent curative surgery, those with positive for CEA mRNA showed higher relapse rates than those who were negative [383]. Furthermore, another study found a significant association between positivity for CK19 mRNA and the presence