1

Epidemiology of female breast cancer

Stephen W. Duffy

Introduction

Since the early twentieth century a large evidence and knowledge base of risk factors for breast cancer has developed. The seminal study of breast cancer epidemiology was the Lane–Claypon report in 1926.[1] This study was important not only in the etiology of breast cancer but also in the development of epidemiological methods, being arguably the first systematic case-control study.

Before reviewing the epidemiological research, some important terms are defined. The two most common study designs in epidemiological research are the cohort study and the case-control study. In the cohort study, a group of subjects free of disease are recruited, and their status with respect to potential risk factors is ascertained. The subjects are followed up for the disease in question, to ascertain which factors are predictive of future disease. For example, in a cohort of nurses in the United States, it was found that women with high consumption of red meat were more likely to develop breast cancer.[2] In a case-control study, subjects with the disease in question (cases) are recruited, together with a group of comparable subjects who do not have the disease (controls). Risk factor status of cases and controls is ascertained retrospectively. A notable example is the international series of case-control studies carried out in the 1960s and 1970s by McMahon and colleagues,[3–6] which among other findings firmly established the association of late age at first childbirth with increased risk of breast cancer. It should be remarked that the case-control design is subject to a number of potential biases, such as differential recall of risk behavior between cases and controls, and therefore their results must be interpreted with caution. Findings from the prospective cohort design are regarded as more definitive.

The association between a given factor and risk of disease is usually expressed as the relative risk (RR). If p is the proportion of subjects with the risk factor who go on

Breast Cancer, ed. Michael Michell. Published by Cambridge University Press.

to develop the disease, and q the proportion of subjects without the risk factor who develop the disease, the relative risk is calculated as p/q. This can be estimated directly in a cohort study, and can be approximated by another quantity, the odds ratio, in a case-control study.

The Lane–Claypon report identified several reproductive history factors as important determinants of breast cancer risk, and set the scene for breast cancer epidemiology research for the next half century. Modern research on hormonal and reproductive risk factors (and some non-hormonal factors) has been collated comprehensively by the Collaborative Group for Hormonal Factors in Breast Cancer.[7–9]

In the last three decades, three aspects of breast cancer epidemiology have received particular attention in the research community:

(1) Diet, alcohol, exercise, and body habitus.
(2) Familial and genetic effects on risk.
(3) Radiological breast density.

In this chapter, the classical risk factors for breast cancer are briefly reviewed and the three topics above are discussed in more detail. In addition current practice and future prospects for the potential of epidemiological results to inform intervention for primary and secondary prevention purposes are reviewed.

Known and suspected risk factors

Table 1.1 shows the maturely established risk factors for breast cancer, with approximate estimates of their effect in terms of relative risk,[2–11] stratified by epoch of finding and documentation. The increased risk with late age at first pregnancy, nulliparity or low parity, late menarche, and early menopause all suggest an increased likelihood of breast cancer as a result of proliferative stimulus by cumulative exposure of the breast to ovarian estrogens. The associations of age at menopause, age at first birth (or rather the proxy, age at marriage), and breast feeding with breast cancer risk were documented in 1926 by Lane–Claypon.[1]

Breast cancer rates tend to be highest in Western Europe and North America, where lifetime risk of the disease is slightly higher than 10%. In some African countries, the lifetime risk is only 1%. In East Asian countries, risk has been low in the past but is increasing now toward western rates. The other most important risk factor is age. The disease is almost unheard of in childhood and adolescence, and incidence gradually increases with age. Around the time of the menopause, there is a hiatus in the trend of increasing incidence, and in western populations incidence

Table 1.1 Maturely established risk factors for breast cancer

Epoch of discovery/ documentation	Risk factor	Effect on relative risk (RR)
Pre-1950	Older age	RR = 2–3 per 10 years
Pre-1950?	Western developed countries	RR = 5–10 compared to East Asia/ Africa
Pre-1950	Nulliparity/low parity	RR = 1.3–1.5
Pre-1950	Late first childbirth/pregnancy	RR = 3 for 40–44 vs. 15–19
Pre-1950?	Family history of breast cancer	RR ≥ 2 depending on strength of family history
1950–1979	Ionizing radiation exposure	RR = 2–3 per Gy if exposed before age 40
1950–1979	Early menarche	RR = 3 for menarche before age 11
Pre-1950	Late menopause	RR = 2 for menopause after age 54
Pre-1950	Breast feeding	RR = 0.9 per year of breast feeding
1950–1979	Overweight (postmenopausal disease only)	RR = 2 for BMI > 35
1980 onwards	Sedentary lifestyle	RR = 1.5–2
1980 onwards	Oral contraceptive use	RR = 1.2
1980 onwards	Hormone replacement therapy	RR = 1.35 for use for 5 years or more
1950–1979	Mammographic parenchymal patterns	RR = 2–4 for denser patterns

continues to increase thereafter at a slower rate. For some far eastern and other populations, incidence falls after the menopause. It is not clear to what extent this is because of biological differences or cohort-specific risk factors. Figure 1.1 shows breast cancer incidence by age for England, for Chinese women in Singapore, and for Indian women in Singapore. The English incidence follows the pattern of increase to the menopause and a slower increase thereafter as described above, the Singapore Chinese incidence plateaus or falls after the menopause, and the Singapore Indian incidence, while low, continues to rise after menopause.

The period between 1950 and 1979 saw a dramatic increase in research activity in breast cancer etiology, notably a number of studies by Lilienfeld and colleagues in the United States,[12,13] and the international series of case-control studies by McMahon and colleagues.[3–6] These tended to confirm the hormonal and

Table 1.2 More recently identified risk factors

Risk factor	Status	Estimated effect on relative risk (RR)
BRCA1/2 mutation	Well established	RR = 10–20 depending on age
Alcohol consumption	Well established	RR = 1.07 per 10g alcohol per day
Diet rich in meat	Uncertain due to confounding factors	RR = 1.8
Quantitative breast density	Well established	RR = 4.6 for density ≥75% vs. <5%

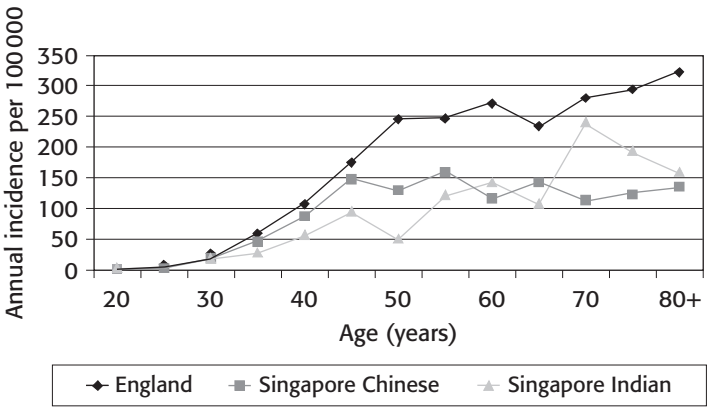

Figure 1.1 Incidence of breast cancer by age in England, in Chinese women in Singapore, and in Indian women in Singapore.

reproductive etiology of the disease. In addition, this period saw the identification of family history and obesity (at least in postmenopausal women) as risk factors.

Table 1.2 shows the more recently identified risk factors for breast cancer. Some of these, such as the *BRCA1* and *BRCA2* mutations are firmly established, with well-documented and validated effects on risk.[14–16] Other factors, such as dietary protein sources, still remain uncertain.[17–19] Breast density is a well-established risk factor, but its practical implications are not yet clear.[20,21]

A rather neglected issue in epidemiology is the fact that, histologically and biologically, breast cancer is not a homogeneous entity. There are a variety of different types of breast cancer, with varying degrees of aggressive potential. Arguably the most obvious biological division is into hormone-dependent, estrogen receptor-positive disease, which usually has a good prognosis and estrogen

receptor-negative breast cancer, which tends to be more aggressive. Type-specific epidemiology has not been well researched to date.

The hormonal risk factors, together with evidence from animal studies, led to the development of hormonal therapies for breast cancer. In turn, the possibility of hormonal chemoprevention has been demonstrated in a number of trials.[22] Although tamoxifen does prevent estrogen receptor-positive breast cancer, its side-effect profile is such that delineating a suitable risk group is difficult.[22,23] A trial of chemoprevention with aromatase inhibitors, which have fewer serious side-effects, is under way.[24]

Two points are worth making with respect to the established risk factors for breast cancer. First, for the most part they are unmodifiable at an individual level (for example, family history) or only modifiable through complex social engineering, the results of which are uncertain (age at menarche, age at first childbirth). Second, most individual risk factors account for only a small proportion of cases. Even those factors with very high relative risks, such as a high-risk *BRCA1* mutation, account for only a small fraction of the incident cases, because of the low prevalence of the risk factors.

The more recently documented risk factors such as breast density, familial and genetic effects, and factors related to diet and obesity are of considerable current interest not only because they help our understanding of the genesis of the disease but also because they have the potential to inform prophylactic or surveillance measures for primary or secondary prevention. Some also have the unusual advantage of having a high prevalence and therefore providing the potential for prevention of a significant proportion of cases (see following section on Radiological breast density). These will be dealt with in more detail in the following sections.

Diet, exercise, alcohol, and obesity

The association of obesity with increased breast cancer risk in postmenopausal women has been known for some time, and an eloquent description of the possible mechanisms of effect was given by de Waard in 1975.[25] The effect of obesity on risk raises the question of weight reduction as a preventive measure. There is retrospective evidence that weight loss both before and after the menopause confers a reduction in future postmenopausal breast cancer risk.[26] If this is confirmed in prospective studies, weight reduction could feature as a preventive measure in

breast cancer as it does in many non-malignant chronic diseases. It is also worth noting that tall stature has been associated with breast cancer risk, at least after the menopause.[27]

It follows from the establishment of obesity as a risk factor that exercise or diet or both may be risk factors, and importantly, modifiable risk factors. Considerable research effort was expended in the 1980s on case-control studies of diet and breast cancer. The major tendency of results was to suggest that diets high in fat and animal protein, and low in antioxidant vitamin sources, mainly fresh fruit and vegetables, were associated with increased risk of breast cancer. In general, these findings were not confirmed by subsequent prospective research,[28] with the exception of some results on protein sources. In a cohort study in Norway, Vatten and colleagues found that women with higher fish intakes and lower meat consumption were at reduced risk of breast cancer,[17] and the US Nurses' Health Study found increased risks with higher intakes of red meat, as noted previously.[2] Case-control findings of a protective effect of high intakes of soy protein suggested a protective effect of isoflavonic phytoestrogen consumption.[18] In general these have not been borne out by prospective cohort studies,[29] although at least one cohort study did find a reduced risk with high intakes of soy isoflavones,[30] and another found a reduced risk with high plasma levels of isoflavones.[31] One possible reason for the varied findings with respect to isoflavone intake is that the studies with negative results tended to be conducted in western countries where typical intakes may be too low to observe the effect.[32]

A comprehensive overview of 10 cohort studies and 43 case-control studies of alcohol and breast cancer found a 7% increase in relative risk of breast cancer per 10 g of alcohol per day, approximately, per drink per day.[7] However, in the prospective cohort studies only, the overall result was the slightly lower figure of 5% increase per 10 g of alcohol per day.

In recent years there have been a number of studies investigating physical activity and breast cancer risk. Physical activity in early life is inversely related to breast cancer risk.[33] This is consistent with the hormonal etiology, as such activity may delay puberty or reduce ovulation. Exercise later in life has also been found to have a protective effect,[34] with a 6% reduction in relative risk per additional hour of exercise per week.[35] As with the predisposing effect of obesity, the influence appears to be stronger in postmenopausal women.

Thus it seems that there is a potential to prevent a substantial number of breast cancers by lifestyle changes. These include increased physical activity, weight loss, a diet lower in animal products, and reductions in alcohol consumption.

Familial and genetic influences

The increased risk of breast cancer associated with a family history of the disease has been observed for decades.[36,37] There is an approximate doubling of risk with a first-degree relative affected. The additional increase in risk associated with early onset in both the proband and the affected relative is characteristic of a genetic effect rather than (or in addition to) shared environmental risk factors within families. The 1990s saw the discovery of two high penetrance gene mutations, *BRCA1* and *BRCA2*,[14,15] inheritance of which increases lifetime risk of breast cancer to 40–85%, with much of the increased risk seen in early adult life. These mutations also confer an increased risk of ovarian cancer. The prevalence of these is estimated to be between 1 in 250 and 1 in 500 in western countries, with higher prevalence in Ashkenazi Jewish women.[17]

There are other high-risk genetic syndromes, including Cowden syndrome, caused by a mutation in the *PTEN* gene, and Li–Fraumeni syndrome, a manifestation of a mutation in the p53 tumor suppressor gene.[17] These have extremely low prevalence, and therefore only account for a very small fraction of breast cancer cases. Research to discover other genetic factors continues.[38]

The ability to identify subjects at very high genetic risk raises the issue of how to manage that risk. A number of options are available.[17] These include bilateral prophylactic mastectomy, which is a fairly extreme measure, but which reduces risk by more than 90%. Oophorectomy has a doubly protective effect for both ovarian and breast cancer, but is necessarily appropriate only in women who have no future plans for children. Another option, which has received considerable attention in recent years, is intensive surveillance, so that breast cancer is not prevented but is detected at a very early and treatable stage. Women at high genetic risk often have radiologically dense breast tissue, rendering X-ray mammography less effective; however, magnetic resonance imaging (MRI) has proved very sensitive in such cases.[39]

For women with a strong family history, but not sufficiently strong to raise serious suspicion of a high-risk gene mutation, one option is mammography at an earlier age and more frequently than is offered to the general population. This is under research in the UK at the moment[40] and is being offered to women from 40 years of age.

Radiological breast density

Breast density is most easily conceptualized as bright areas on a mammogram, representing volumes of fibroglandular as opposed to fatty tissue. In 1976, Wolfe

classified the radiological appearance of the breast into four mammographic parenchymal patterns.[41] He observed a higher risk of subsequent breast cancer in the patterns which he named P2 (prominent duct pattern with density occupying up to 75% of the breast) and DY (homogeneous fibroglandular density occupying a large proportion of the breast). This was confirmed qualitatively by subsequent studies, but the magnitude of the increase in risk was lower than that observed by Wolfe. Other methods of classification exist, such as the Tabár and Breast imaging reporting and data system (BI-RADS) classifications.[42,43] The combined evidence suggests that a continuous quantitative assessment of density is a better predictor of risk than a qualitative or semi-quantitative categorization.[20] Computer-assisted methods are available for estimating total dense area of the breast and percent dense area.[21]

A meta-analysis of published studies found a steady increase in risk of breast cancer with percent breast density, with a combined relative risk of 4.64 for density of 75% or more compared with density of less than 5%. Another finding was that the increased risk of future breast cancer with density could not be wholly attributed to the masking phenomenon, whereby the greater incidence of breast cancers after mammograms showing high density is a result of the symptomatic appearance of cancers that were present at the time of the mammogram but missed owing to the density.

It was also concluded that the effect of breast density on risk persisted after adjustment for other factors. However, it should be noted that percent breast density is negatively confounded with two other risk factors: body mass index (BMI) and age. The first is probably because subjects with higher BMI tend to have a greater amount of non-dense (i.e. fatty) tissue in the breast, and therefore a lower percent density.[44] Thus, absolute dense area is unlikely to be confounded with BMI. The second relationship follows the natural replacement of fibroglandular with fatty tissue with aging. One implication of this confounding is that percent breast density should be considered conditional on or controlling for age and BMI, otherwise its effect on risk will be underestimated.

While the exact implications of the role of density in breast cancer control remain to be determined, it is likely that they will be important. Density is unique in terms of breast cancer risk factors in having a relatively high attributable fraction. This means that while most breast cancer risk factors only account for a small number of disease cases, density may account for a substantial proportion. It has been estimated that density of 50% or more could account for around a third of breast cancers.[45] In addition, it is modifiable. In the IBIS-1 randomized trial of tamoxifen for prevention of breast cancer, density was significantly reduced with tamoxifen compared with

placebo.[46] The major question now is whether an alteration in density in middle age will alter future risk of breast cancer. This is under research at the moment.[46] If the answer is yes, then density can be used as a marker of whether a preventive agent is working, thus giving an early endpoint for chemoprevention trials, and providing a signal in the clinical service setting to suggest that a particular preventive strategy for an individual patient may need to be changed.

In the meantime, the high relative and attributable risks associated with density mean that it can be used as a selection criterion for primary or secondary preventive interventions. In addition, it may be used to devise personalized screening regimens. For example, in the Gothenburg trial of mammographic screening, two-view mammography was used where the previous mammogram indicated a high degree of density, and single-view otherwise.[47] Nowadays, two-view mammography is the standard of care in screening, but high levels of density could also be an indicator for more frequent screening, referral to an extra X-ray reader, or use of another screening method such as MRI.[39]

Conclusion

Our understanding of the epidemiology of breast cancer together with our ability to influence the natural history of the disease both in individuals and in populations is almost unrecognisable compared with a few decades ago. Screening and early detection together with improved therapy have resulted in a striking improvement in survival.

Contemporaneously, the role of epidemiology has changed from simply finding clues to etiology and mechanisms of disease to informing and triggering intervention. Familial and genetic risk is used to determine prophylactic or surveillance measures. The hormonal epidemiology informed the development of hormonal chemoprevention. Epidemiological results also suggest weight loss as a preventive measure in postmenopausal women. Breast density is already used to recruit subjects to preventive or surveillance studies, and its role may well extend to monitoring of subjects undergoing prevention regimens, and acting as a surrogate for breast cancer in prevention trials.

It could be argued that in the last century, breast cancer epidemiology has supplied a paradigm for chronic disease epidemiology, initially documenting who actually develops the disease, then developing analytic studies to research why they develop it, and applying that knowledge to intervention to prevent the disease or

mitigate its effects. Although this has been successful in contributing to control of the disease, there is still work to do in breast cancer epidemiology. Further clarifying the role of breast density and validation of the effect of weight loss in reducing risk of postmenopausal breast cancer are clear examples. In addition, it is appreciated that breast cancer is not a homogeneous disease. There are different types of breast cancer, which need to be treated (and sometimes diagnosed) in different ways. There is a clear gap in our knowledge of different etiological pathways for different types of breast cancer.[48] There is plenty to keep us busy in the future.

REFERENCES

1. Lane-Claypon JE. *A Further Report on Cancer of the Breast, with Special Reference to its Antecedent Conditions*. Reports on Public Health and Medical Subjects No. 32. London: Ministry of Health, 1926.
2. Cho E, Chen WY, Hunter DJ, et al. Red meat intake and risk of breast cancer among premenopausal women. *Arch Intern Med* 2006; **166**: 2253–9.
3. MacMahon B, Cole P, Lin TM. Age at first birth and breast cancer risk. *Bull WHO* 1970; **43**: 209–21.
4. Lowe CR, MacMahon B. Breast cancer and reproductive history of women in South Wales. *Lancet* 1970; **i**: 153–6.
5. Mirra P, Cole P, MacMahon B. Breast cancer in an area of high parity: São Paulo, Brazil. *Cancer Res* 1971; **31**: 77–83.
6. Ravnihar B, MacMahon, Lindtner J. Epidemiologic features of breast cancer in Slovenia 1965–1967. *Eur J Cancer* 1971; **7**: 295–306.
7. Collaborative Group on Hormonal Factors in Breast Cancer. Alcohol, tobacco and breast cancer–collaborative reanalysis of individual data from 53 epidemiological studies, including 58515 women with breast cancer and 95067 women without the disease. *Br J Cancer* 2002; **87**: 1234–45.
8. Collaborative Group on Hormonal Factors in Breast Cancer. Breast cancer and hormone replacement therapy–collaborative reanalysis of individual data from 51 epidemiological studies of 52705 women with breast cancer and 108411 women without breast cancer. *Lancet* 1997; **350**: 1047–59.
9. Collaborative Group on Hormonal Factors in Breast Cancer. Breast cancer and hormonal contraceptives: further results. *Contraception* 1996; **54** (Suppl 3): S1–106.
10. Dumitrescu RG, Cotarla I. Understanding breast cancer risk–where do we stand in 2005? *J Cell Mol Med* 2005; **9**: 208–21.
11. McPherson K, Steel CN, Dixon JM. Breast cancer–epidemiology, risk factors and genetics. *Br Med J* 2000; **321**: 624–8.
12. Lilienfeld AM. The relationship of cancer of the female breast to artificial menopause and martal status. *Cancer* 1956; **9**: 927–34.