Chapter 3

Materials

In this chapter, the study design, which included the selection of cases and controls, inclusion criteria of the study subjects, the method of data collection and the various factors studied in the present thesis for risk estimation is provided.

3.1. Study design

A case-control study was conducted at the Regional Cancer Center (RCC), Trivandrum, Kerala between 2002-2005 in order to study the risk factors associated with female breast cancer. A total of 1169 cases and 1208 controls were included in the study. The cases were women with histologically confirmed incident (new) primary breast cancer who came for treatment at the RCC. The controls were disease free subjects who never had a history of cancer and accompanied with cancer patients [Relatives of breast cancer patients were excluded from the controls] during the same time period, and frequency matched to cases by age (±5 years) and residence status (urban/rural). The institutional review board of the RCC approved the study. Written informed consent was obtained from all participants before interview.

The participation rates were more than 90% for both cases and controls. The cases were scrutinized to confirm whether the patients attending the hospital were histologically diagnosed as breast cancer and controls were scrutinized to confirm that they are not relatives of breast cancer patients.

In-person interview of each case and control was conducted in the hospital using a pre-tested structured questionnaire. The questionnaire contains information on demographic and socio-economic factors, reproductive factors, household and occupational physical activity levels, residence history, family history of cancer, smoking
and chewing of tobacco, alcohol consumption, prior and current history of other chronic illness, and frequency of consumption of different food items. Anthropometric measurements such as height, weight, waist and hip circumferences were taken at the end of interview.

The present thesis focused on socio-demographic factors, reproductive factors, anthropometric factors and level of physical activities. Lists of these factors and the corresponding categories used in the present thesis were given as follows (Appendix 2).

3.2. Risk factors

3.2.1. Socio-demographic factors

Socio-demographic factors were age, religion, marital status, education, socio-economic status and residence status.

Age was considered as a continuous variable. Religion was categorized into three groups such as Hindus, Muslims and Christians. Marital status was grouped into four categories such as Unmarried, Married, Widowed, and Divorced/ Separated. Level of education was calculated according to the number of years spent for schooling and were classified as ‘illiterate & < 5 years of schooling’, ‘5-8 years of schooling’, ‘9-12 years of schooling’, and ‘>12 years of schooling’.

Socio-economic status was derived by summing up the independent scores given to home ownership, number of the rooms and number of people living in the house, availability of toilet and running water as well as comfort/luxury items such as electrical/gas stove, refrigerator, television, air conditioner, car, motorcycle/scooter, bicycle, and computer owned by the study subject and the summed up score was divided into quintiles.
Residence status (urban/rural) was collected according to the definition of national census. All participants were asked to list their places of residence where they had lived for at least one year, starting with the place of birth. If the subject lived in a ‘Panchayat’ area, residence status is defined as ‘rural’ and all other areas such as ‘municipality’ and ‘corporation’ is defined as ‘urban’. The duration of stay at each place of residence was also collected. If the subject migrated to urban area and lived there during the immediate previous 10 years, residential status was assigned as ‘urban’ and vice-versa.

3.2.2. Reproductive factors

Reproductive factors included were age at menarche, menopausal status, age at menopause, age at marriage, age at each pregnancy, number of live births/still births, parity (number of children), status of breastfeeding with respective durations, hormone use, and methods adopted to prevent pregnancy.

The age at menarche was categorized into five groups (<12 years, 12-13 years, 14-15 years, 16-17 years and >17 years) and age at menopause was categorized into four groups (<41 years, 41-45 years, 46-50 years and >50 years).

Age at marriage was classified into five age groups such as <18 years, 18-19 years, 20-22 years, 23-24 years, and ≥25 years. Age at first and last pregnancy were collected by starting with the age at earliest pregnancy and finishing with the age at most recent pregnancy and were classified into four age groups such as <21 years, 21-25 years, 26-30 years, and >30 years.

Parity was grouped into four categories such as nulliparous, subjects with <3 children, 3 children, and ≥3 children. Information on age at first childbirth was categorized into six groups such as <20 years, 20-21 years, 22-23 years, 24-25 years, 26-
27 years, and ≥28 years. Parity and age at first childbirth were calculated by considering live births only (abortion and miscarriage were excluded).

The lifetime duration of breast-feeding for parous women was calculated by summing up breastfed duration for each child. Total duration of breast feeding was categorized into five groups such as ≤1 year, 2-3 years, 4 - 5 years, 6-7 years, >7 years.

Prevention strategies adopted for preventing pregnancies such as use of condoms, natural methods (calendar method, interruption), intrauterine devices, tubal ligation, vasectomy of the spouse were also included.

3.2.3 Anthropometric factors

The height (in cm) and weight (in kg) of each participant were measured using standard equipments. All subjects were asked to remove their shoes before measurements were taken. In addition to this, weight was measured with light clothing. Waist size (in cm) was measured using a standard tape at the level of navel around the skin and hip size (in cm) was measured with light clothing at the widest part of hip. All measurements were done twice in succession and averaged for a final value.

Body mass index (BMI) was computed as weight in kilogram divided by height in meters squared. Three mutually exclusive BMI groups were created: lean weight (BMI <25 kg/m²), over weight (25 kg/m² < BMI < 30 kg/m²), and obese (BMI ≥ 30 kg/m²) (WHO 1998).

Waist-to-hip ratio (WHR) was computed by taking the ratio of waist size (in cm) and hip size (in cm). Two mutually exclusive WHR groups were created ≤0.85 and > 0.85 (Royal College of Physicians Report 1998).
Body sizes at different ages were also assessed using a pictogram (Figure 3.1). This pictogram was shown to each subject to mark their body sizes in different periods of life (at 10 years, 20 years and the period when the data was collected).

![Figure 3.1: Pictogram used in the questionnaire to measure the body size at different ages.](image)

### 3.2.4. Physical activity

Information on the amount of time spent on a normal day in household activities such as cleaning the house (dusting and sweeping), cooking, washing clothes (without machine), walking to buy things or do errands, carrying water, working in the garden or farm, taking care (actively) of children or grandchildren and ironing clothes was obtained at interview. Total time spent for all these household physical activities on a normal day was also calculated by summing up all time spent by the participant for individual household activities. Total time spent on a normal day for such activities were categorized into five groups (<180 minutes, 180-239 minutes, 240-299 minutes, 300-359 minutes and ≥360 minutes per day).

To assess the sedentary activities, the time spent to watch television during weekdays and weekends were obtained separately. The collected information were categorized into four groups (<60 minutes, 60-119 minutes, 120-179 minutes and ≥180 minutes per day).
minutes per day) for watching television during week days and three groups (<60 minutes, 60-179 minutes and ≥180 minutes per day) for week ends, as people used to spend more time for watching television during weekends.

To assess the occupational physical activity, the main occupation of the subject was used in the present analysis. The occupation was categorized as farmer/ agricultural worker, teacher, industrial worker, maid/casual worker and office worker (secretary, clerk, officer). All other employed jobs were combined together in the present analysis. If the subject had retired at the time of data collection, the main occupation before retirement was collected.

Information on time spent for sports/ exercise at the time of data collection and when the subject was young was also collected. The cases were asked to relate the events one year before diagnosis of cancer assuming that the disease might have changed their level of physical activity. As the available information on sport/exercise are limited, not included in the analysis.