Appendix D: Review of Estimates of Progression Probabilities and Duration of States of Cervical Neoplasia

This appendix discusses the major studies yielding estimates of the duration of the various states of cervical neoplasia, the probability of progressing from one state to the next, and the probability of regressing to a previous state. The studies are summarized in chapter 2(tables 4, 5, and 6).

Duration of States of Cervical Neoplasia

The first major study of the natural history of cervical neoplasia in the medical literature is that of Petersen (107), in which a number of women with a diagnosis of cervical intraepithelial neoplasia (CIN) were followed without intervention for a number of years. The researchers observed that the average duration of carcinoma *in situ* (CIS) in these women was 3.7 years, with the time from onset of CIS to onset of invasive cancer ranging from less than 1 year to nearly 9 years in individual women.

Withholding treatment from women diagnosed with CIN rapidly became unacceptable, so subsequent researchers have attempted to estimate the duration of the states of cervical neoplasia in various other ways. One method has been to assume that the process of cervical cancer can be approximated by a Markov process, a type of model that uses a given set of probabilities to relate one state to the next (app. E). Observable variables can be used as inputs to such a model and used to estimate the durations of the various states. A homogeneous model assumes the same transition probabilities for all age groups; a nonhomogeneous model allows the probabilities to vary depending on the group to which they apply.

Barron and Richart used a homogeneous Markov process to assess the duration of CIN using two very different data sets (7,8,9). The first of these was a prospective followup of 557 women in Virginia and New York with known CIN, whose diagnosis was based on 3 successive abnormal Pap smears. Women in the study were followed every 1 to 4 months without intervention if their status was unchanged. The researchers attempted to minimize diagnostic errors by using clearly delineated diagnostic criteria, requiring three smears with CIN for admission to the study, and reviewing smears for reliability and accuracy. Accordingly, the inter-reader reliability in classifying smears was 95 percent, and the Pap smear diagnoses agreed with the colposcopic and ultimate histological diagnoses. Some important information about the sample is not presented in the published accounts: the distribution by age and race of study participants, the comparative characteristics of women who participated and those who refused participation, and the profile of all women with abnormal smears in the institutions studied. The authors found no evidence that transition from one state to another was age-related, but few older women were represented in the sample (the median age was 26, and the maximum age was 65) (8). The median duration of CIN (including severe dysplasia) derived from this population was 3.7 years; the mean duration was 5.7 years (l14).

The second population that Barron and Richart used to estimate duration of CIN was a sample of 11,814 previously unscreened asymptomatic women iiving in Barbados, West Indies, who attended family planning clinics between 1965 and 1968 (9). The sample included 171 women aged 60 to 64.

However, the model only used prevalence data from women aged 20 to 39 to determine the duration of CIN. The resulting estimates were similar to those yielded by the original model. Although this study is based on a population that probably differs from the general U.S. population in risk of cervical neoplasia, the consistency of results between the two studies supports the validity of these researchers' estimates.

In contrast, Coppleson and Brown (32) attempted to demonstrate that the results of a homogeneous Markov model, which assumes that transition probabilities are independent of age, did not fit observed data. They used data collected by Bibbo et al. and age-specific incidence rates for invasive cervical cancer from the Third National Cancer Survey for their estimates, and they used a non-homogeneous Markov model to simulate a process that would yield these real-world results.

The published data used by Coppleson and Brown had some limitations. Bibbo et al.'s series was based on 148,735 women attending the University of Chicago and Planned Parenthood clinics (16). Most of the women were young; only 12 percent of the sample was over age 49. The mean age of this subsample of 17,133 women was given as 65, but no information on the actual age distribution was published. For their model, Coppleson and Brown assumed all of these women to be the average age-- i.e., 65.

Based on their models, Coppleson and Brown estimated the average duration of CIS to be 17 years in women under 65 and 4 years in women aged 65 and over (132). Their findings are valuable because the researchers expressly examined the possibility of differences between age groups, and they concluded that a real difference in the behavior of the disease probably does exist. However, there are several caveats to their findings. Their assumption that all women over 49 were age 65 may invalidate their conclusions for older women, since even if the *average* age of this group were 65, using the actual distribution of ages would give different results from assuming that *all* women were age 65. The effect of this assumption is probably to underestimate the duration of disease states in older women. In addition, several of the assumptions made to fit the model to observed data are at variance with many other researchers' beliefs about the natural history of the disease. For example, Coppleson and Brown assumed that CIN prior to CIS (i.e., dysplasia) was a transient condition, lasting less than one year, and that CIS regressed to normal in a large proportion of cases. In summary, while it may be true that a homogeneous model does not fit the observed natural history of the disease, this analysis does not resolve the issue.

Dunn (40) and Kashgarian and Dunn (69) drew upon the relationship between prevalence and incidence to estimate the duration of a given state of cervical neoplasia. In his paper, Dunn (4 O) divided the sum of all age-specific prevalence rates for a given state by the sum of all age-specific incidence rates for that state to derive the duration of the state. Kashgarian and Dunn (69) used an equivalent method that did not depend on pre-determined age ranges over which age-specific rates were calculated. They estimated duration by first graphing the incidence of each state (CIS, preclinical invasive cancer, and clinical invasive cancer), with age along the bottom axis of the graph. They then estimated the area under the graph between given ages for CIS. Next, they calculated the age at which the graph of the incidence of preclinical invasive cancer had an area under it equivalent to the area under the defined CIS age interval. The duration of CIS was then presumed to be the difference between this age and the upper limit of the specified CIS age range.

The analyses of these researchers used data for 110,000 white women screened in Memphis (40) and 106,000 person-years of observation of white women and an unspecified number of black women from

Memphis and Shelby counties, Tennessee (69). In both of these samples women aged 20 to 39 were over-represented and women aged 55 to 74 were under-represented. Incidence rates were determined from the result of the third screening test. The literature reports did not specify diagnostic criteria for cytology, although most smears were read at one university laboratory. The authors concluded that older women and black women had shorter durations of CIS than other women. Their estimates for black and white women were 8.5 and 10.7 years, respectively (69). For white women, they estimated the duration of CIS to be between 5 and 16 years for young women and to be 1 year for women aged 65 and over (69).

Difficulties in true ascertainment of a particular disease state due to cytology and misclassification may affect the accuracy of estimates of duration derived in this manner. In addition, estimates derived from this formulation will only be correct if the incidence and prevalence rates over the time of measurement are constant and if both are measured from the same state in the disease process. If prevalence rates are increasing or if screening rates differ, these conditions may not hold true. Also, these types of estimates assume similar population mortality rates for women at all ages. For elderly women, who have higher mortality rates than younger women, this assumption may underestimate the true duration of cervical neoplasia.

However, other authors, using similar methods applied to different data sets, have derived estimates of duration that are similar to those of Kashgarian and Dunn, arguing for the validity of this method. Barron et al. applied this method to two data sets: 1) the incidence and prevalence of CIS in British Columbia, Canada, and 2) the prevalence of CIS in Barbados, West Indies. They examined these two sets of data in two ways: first by using the simple relationship between prevalence and incidence described above, and second by examining the equivalent areas under the curves on graphs of the incidence of each stage at each age. They concluded that the duration of CIS is an age-independent variable with upper and lower limits of 3 to 10 years, respectively(7).

Fidler et al. likewise estimated the duration of CIS in two ways: 1) from the relationship between observed prevalence and incidence among women participating in the British Columbia screening program, and 2) from the difference in mean ages of incidence of CIS and preclinical invasive cancer in this population (46). Age-specific rates were presented for this series, but the number of elderly women was small, yielding estimates with a wide range of error. In 1966, 22 percent of the female population in British Columbia was age 60 or over, but this age group represented only 8.5 percent of women screened (46). The estimate of the duration of CIS using the first method is between 6 and 9.5 years, compared to 12 years using the second method.

Another method of estimating the duration of neoplasia is to determine the modal age-specific incidence rates (the modal number of cases per age group). The Canadian Task Force presented such estimates using data from the British Columbian population (22). As with the data used by most other researchers, these data are cross-sectional and may obscure differences among cohorts. Also, estimates obtained as a result of subtracting modal, or even mean, ages of incidence only yield a correct estimate if the durations of all states being considered are equal. This is not likely to be true since the probability of ascertainment is a function of the duration of the lesion (7). The estimates derived in this manner agree the least with estimates from other methodologies, and they are most likely overestimates.

The uncertainties about the duration of cervical neoplasia in elderly women arise primarily from the lack of data on women in this age group. The critical question in assessing the duration of each state of cervical neoplasia is whether the duration of disease

is, or is not, dependent on age. Although Coppleson and Brown's analysis suggested that duration of different states was indeed different in the elderly than in the younger population, there is still little direct evidence to support or refute this hypothesis. The hypothesis is biologically plausible based on current knowledge of the interactions between age and hormonal and immune factors (see ch.2}

Probability of Progression and Regression of Each State

The probabilities of remaining in a given state, progressing to the next state, or regressing to the prior state are difficult to determine for cervical cancer. The best estimates of early disease are based on groups of women with CIN and CIS[1] who were followed without intervention. In one such series, Barron and Richart followed 557 women with CIN (8) and collected data on the distributions of grades of CIN after the initial and two followup exams. Using a Markov model with these data as inputs, the authors estimated that after 10 years 66 percent of all CIN lesions will progress to CIS, 28 percent will remain in CIN, and 6 percent will revert to normal. Regression to normal only occurred from very mild or mild CIN, and the overall probability of progression to CIS increased with the severity of dysplasia (8}

In contrast, Fox (48) followed 278 women with CIN and noted that 31 percent regressed to normal, 9 percent remained in CIN, and 60 percent progressed to CIS. This high regression rate may be due in part to misclassification, as only one smear interpreted as CIN was necessary for inclusion in the study; several women 'regressed" after termination of pregnancy or completion of anti-infection treatment. In addition, 13 percent of the women whose smears originally returned to normal subsequently developed CIN.

The "re-development" of CIN after regression to normal has been noted in other series as well (147). In fact, in one series, among women over age 45 whose initial CIN lesions 'regressed," 40 percent recurred (147). All these factors suggest that estimates of the regression rate of CIN have often been overstated, due to misclassification biases. In contrast, three smears interpreted as CIN were necessary for inclusion in Barron and Richart's series, which should minimize this type of bias.

A Swedish study of 894 women age 15 to 72 with CIN, who were followed for an average of over 4 years, found that 54 percent of lesions regressed, 16 percent persisted and 30 percent progressed (lOO). A number of patients with "persisting CIN" in this study had periods of normal smears for more than 12 months before being rediagnosed as having CIN.

Evidence on the relationship between regression rates and age is mixed. One study noted a lower regression rate for older than for younger women. In this study, women under age 45 with CIN had a regression rate of 38 percent per year, compared to 29 percent per year for women age 45 and over (147). Another study, however, found that fewer lesions progressed, and more regressed, in older women than in young women (100).

The difference in rates of regression noted in nonbiopsy studies compared to biopsy studies suggests that the act of establishing the diagnosis can produce a cure, and that the act of measurement often alters the results (8). Nasiell and colleagues (100)

1 As noted in ch. 2, CIN as used in this report include mild and moderate dysplasia; CIS generally includes CIS and severe dysplasia, because these latter two conditions are hard to distinguish. Barren and Richart, however, specifically attempted to separate different levels of dysplasia and CIS. In the discussion of their studies, CIN includes severe dysplasia; CIS includes only carcinoma in situ.

found that significantly more biopsied than nonbiopsied lesions disappeared during followup (50 v. 57 percent), and fewer biopsied lesions progressed (25 v. 27 percent). Thus, different biopsy rates may be one reason studies report varying results. In addition, nonuniformity of diagnostic criteria and observer variability contribute to the wide range of reported probabilities.

Confounding these difficulties in determining the "true" course of cervical neoplasia in elderly women is the lack of age-specific observational data. The main source of findings regarding age-specific information is the research of Coppleson and Brown, which found that an age-dependent model of disease progression fit actual incidence and prevalence data best (32). The researchers concluded that there is no regression from the state of CIS and that the probability of disease progression from CIS to invasive cancer increases with age. For CIN, the probability of progression did not vary with age in their model. As discussed above, however, their conclusions have some uncertainties due to limitations of the underlying data. Also, the older women in their data set had lower screening rates than young women, which may have resulted in older women being detected at the end of a given state more often than younger women. This could bias the model to predict a higher probability of disease progression in the elderly. Still, as with the duration of disease states, a higher probability of disease progression is biologically plausible in older women.