ARTHRITIS & RHEUMATISM
Vol. 48, No. 8, August 2003, pp 2246–2255
DOI 10.1002/art.11073

Prevalence, Incidence, Survival, and Disease Characteristics of Systemic Sclerosis in a Large US Population

Maureen D. Mayes,[1] James V. Lacey, Jr.,[2] Jennifer Beebe-Dimmer,[3] Brenda W. Gillespie,[4] Brenda Cooper,[5] Timothy J. Laing,[6] and David Schottenfeld[3]

Objective. **To estimate the prevalence, incidence, survival, and disease characteristics of systemic sclerosis (SSc) in the Detroit tricounty area.**

Methods. **A census of SSc cases for the period 1989–1991 was conducted in the Detroit area, using multiple sources for case identification. Diagnoses were verified by medical record review. Capture-recapture analysis was used to estimate the total SSc population. Cases of localized scleroderma (morphea and linear disease) were excluded.**

Results. **Based on 706 verified cases of SSc, prevalence was initially estimated to be 242.0 cases per million adults (95% confidence interval [95% CI] 213–274), with an annual incidence of 19.3 new cases per million adults per year (95% CI 12.4–30.2). Capture-recapture analysis, based on the degree of overlap of verified cases among multiple sources, resulted in a revised prevalence estimate of 276 cases per million adults (95% CI 245–310). Sex- and race-specific prevalence estimates were significantly higher for women than for men, and for blacks than for whites. The average age at diagnosis was significantly younger for blacks than for whites. Compared with white patients, black patients were almost twice as likely to have diffuse disease (prevalence proportion ratio 1.86, 95% CI 1.48–2.35). Median survival was ~11 years. Factors negatively affecting survival included male sex (hazard ratio 1.81, 95% CI 1.29–2.55) and older age at diagnosis (hazard ratio 1.04, 95% CI 1.03–1.05).**

Conclusion. **This study establishes baseline estimates of SSc occurrence and characteristics in a large US cohort consisting primarily of black adults and white adults. These data should facilitate research regarding the role of geographic, ethnic, racial, and environmental factors for this disease in comparison populations.**

Published estimates of the prevalence and incidence of systemic sclerosis (SSc; scleroderma) vary widely depending on the period of observation, methods of case ascertainment, and the geographic area of study (1–6; for review, see ref. 7). Reported US incidence rates vary between 2.7 cases per million per year for the time period 1947–1968 (2) and 18.7 cases per million per year for the time period 1972–1982 (6). Similarly, US prevalence estimates have varied from 138 cases per million for the period 1950–1979 (4) to 286 cases per million in 1985 (5). These discrepancies may reflect true variation in disease occurrence among different populations or may be related to methodologic differences, such as the degree of scrutiny applied or the classification of disease.

Although survival in patients with SSc has improved in the past several decades, it remains considerably diminished compared with that in age- and sex-matched populations (8–18; for review, see ref. 19). Reliable estimates of overall survival as well as survival by disease subtypes are important in order to gauge effects of new treatment modalities or to document the changing natural history of this disease. Dependable incidence, prevalence, and mortality statistics are neces-

Supported by the NIH (grants N01-AR-5-2217 and N01-AR-0-2251 from the National Institute of Arthritis and Musculoskeletal and Skin Diseases) and by Dow Corning.

[1]Maureen D. Mayes, MD, MPH: University of Texas at Houston Health Science Center; [2]James V. Lacey, Jr., PhD: National Cancer Institute, Rockville, MD; [3]Jennifer Beebe-Dimmer, PhD, David Schottenfeld, MD, MSc: University of Michigan School of Public Health, Ann Arbor; [4]Brenda W. Gillespie, PhD: University of Michigan, Ann Arbor; [5]Brenda Cooper, MS: Port Matilda, Pennsylvania; [6]Timothy J. Laing, MD: University of Michigan Medical Center, Ann Arbor.

Address correspondence and reprint requests to Maureen D. Mayes, MD, MPH, Division of Rheumatology, University of Texas at Houston Health Science Center, 6431 Fannin Street, MSB 5.270, Houston, TX 77030. E-mail: Maureen.D.Mayes@uth.tmc.edu.

Submitted for publication July 10, 2002; accepted in revised form March 24, 2003.

sary to evaluate the health impact of this disease on the population.

The aim of this study was to determine the prevalence and incidence of SSc and survival rates in patients with SSc by conducting a census of all patients with SSc in the Detroit tricounty metropolitan area for the years 1989–1991. This metropolitan area, consisting of the counties of Wayne, Oakland, and Macomb, has a large overall population with a substantial black component (2.917 million adults, of whom 75.0% are white, 21.9% are black, and 3.1% are of other races [according to 1990 US census data]).

PATIENTS AND METHODS

Case ascertainment and identification. Patients were identified from the following sources: 1) the patient population of hospitals and outpatient clinics affiliated with Wayne State University/Detroit Medical Center, 2) the patient population of the University of Michigan Medical Center, 3) patients from the practices of Detroit area rheumatologists, 70% of whom participated in this study, 4) a computer search of medical records from all 55 area hospitals (including those in the counties adjacent to the tricounty area) for patients discharged with the diagnosis of SSc (International Classification of Diseases [ninth version] code 710.1 [20]), and 5) the Southeast Michigan chapter of the Scleroderma Foundation (formerly, the United Scleroderma Foundation), a patient support group.

Verification of diagnosis. All cases were verified by medical record review. Patients with definite SSc were defined as those who met the American College of Rheumatology (formerly, the American Rheumatism Association) criteria for the classification of SSc (21). Patients defined as having probable SSc had a rheumatologist diagnosis of SSc, documented sclerodactyly, and at least 2 other features of the CREST syndrome (calcinosis, Raynaud's phenomenon, esophageal dysfunction, sclerodactyly, and telangiectasia) and did not meet the criteria for systemic lupus erythematosus (SLE) (22). In the course of this study, several patients who initially were classified as having probable SSc were seen at one of the academic centers, and the features necessary for a diagnosis of definite SSc were documented. In most cases of probable SSc, the medical record was lacking documentation of features that indicated a diagnosis of definite SSc. Subjects with either definite or probable SSc were counted as cases. Additional medical records were sought if a diagnosis could not be determined from the first source.

SSc was designated as limited or diffuse on the basis of the extent of skin involvement as documented in the medical chart, according to the method described by LeRoy et al (23). Limited disease was defined as skin thickness confined to areas of the extremities below the elbows and below the knees. Diffuse disease was defined as skin thickness involving the proximal extremities or the trunk below the clavicles. If the medical record did not include a description of the extent of skin involvement but the above criteria were met, the case was counted as SSc, but a designation of limited or diffuse was not assigned. Cases of localized scleroderma (morphea and linear disease) were excluded from this study.

Chart abstraction and data collection. Medical records were abstracted by study personnel using a coding form containing 74 clinical and laboratory variables (24). Each of these variables was recorded as normal, abnormal, or not performed. Autoantibody assessments were also recorded as normal, abnormal, or not performed, except that the pattern and titer of antinuclear antibodies were noted. All abstractors were trained by study rheumatologists and were given several practice charts to review while under the direct supervision of the rheumatologists. The final diagnosis and eligibility for inclusion were determined by one of the study rheumatologists after reviewing the information on the abstract form. If the diagnosis, eligibility, or other elements of the clinical data were unclear, additional chart sources were sought and/or the patient's physician was contacted by telephone. The date of diagnosis was defined as the date when scleroderma was first mentioned in the medical record or the date when the first non-Raynaud's SSc-related symptom was documented in the medical record.

The data abstracted from the charts were also used to identify internal organ system involvement. Renal involvement was defined as a serum creatinine level $\geq$1.8 mg/dl or a history of hypertensive crisis. Pulmonary involvement consisted of pulmonary fibrosis (defined as bibasilar fibrosis on a chest radiograph or a chest computed tomography scan) or pulmonary hypertension (if labeled as such in the chart or documented by Doppler echocardiography). Gastrointestinal involvement was considered to be present if either esophageal dysmotility or malabsorption was documented. Cardiac involvement was not assessed due to incomplete documentation and difficulty attributing cardiac disease to SSc rather than other causes.

Determination of vital status. For all cases in which the patient was not shown to be deceased according to the hospital discharge database or medical chart review, a search of the National Death Index through 1998 was performed to determine the patient's vital status.

Criteria for inclusion. Patients were included according to the following criteria: 1) a diagnosis of definite or probable SSc according to the above classification system; 2) not deceased prior to January 1, 1989, and SSc diagnosed prior to January 1, 1992; and 3) resident in the Detroit tricounty area for at least part of the time from January 1, 1989, through December 31, 1991, according to their address as listed in the medical record and/or as listed in a commercial guide of household residences with dates of occupancy (Bresser's Guide, Detroit, Michigan). In addition, to be considered an incident case, the patient had to reside in the area at the time of diagnosis.

Statistical analysis. Prevalence and incidence estimates were calculated using sex- and race-specific census data provided by the 1990 US census for the Detroit metropolitan area adult population (ages 18 years and older). Confidence intervals (CIs) were obtained by assuming that the number of cases followed a binomial distribution. Statistical tests for equality of proportions were performed using standard chi-square analysis.

Survival analysis was performed using a left-truncated Kaplan-Meier estimate (25). We estimated the probability of

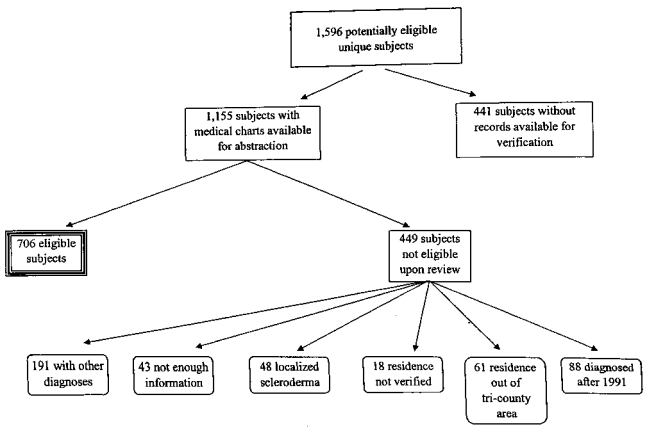

Figure 1. Flow chart of systematic approach to case ascertainment and verification.

survival as a function of time since diagnosis but included only people "at risk" of death while they were under observation, i.e., after January 1, 1989. This method adjusts for a bias inherent in prevalence studies; that is, people with long survival are more likely to be included than are those with short survival. Expected survival probabilities for an age-, race-, and sex-matched Michigan population were calculated using published hazard data for the Michigan population (26). Survival time comparisons, adjusted for age at diagnosis, were performed using left-truncated Cox proportional hazards regression (27,28). The model included age at diagnosis, race, sex, and extent of disease (limited versus diffuse). Age at diagnosis was included as a continuous variable in all models.

Using a method based on that described by Wittes et al (29), capture-recapture analysis was performed to adjust for case underestimation by evaluating the degree of overlap among sources (number of cases captured by all 5 sources, by 4 sources, etc.) to estimate the total number of persons with SSc in the population.

Human subjects approval. This study was approved by the Institutional Review Boards of the University of Michigan and Wayne State University.

RESULTS

Case ascertainment and verification. From our multiple sources, 1,596 potentially eligible, unique patients were identified (Figure 1). This number excludes duplicates, patients who may have been initially identified but later were determined to have died prior to 1989, and Scleroderma Foundation members who reported themselves either as having localized scleroderma or as being friends and family members rather than patients.

Of the 1,596 potentially eligible patients, 1,155 (72.4%) had medical records available for review. Medical records of the other 441 patients (27.6%) were not available for one of the following reasons: permission to review records could not be obtained (365 patients [22.9%]), the record was not provided in spite of multiple requests (61 patients [3.8%]), or the hospital had closed (15 patients [0.9%]).

Of the 1,155 patients whose records were reviewed by study personnel, 706 (61.1% of those reviewed and 44.2% of the total pool of potential patients) were eligible. Reasons for the ineligibility of the other 449 patients included a date of diagnosis of SSc after 1991 (88 patients [5.5% of total potential pool]), residence outside the tricounty area during the prevalence period (61 patients [3.8%]), residence not verified (18 patients [1.1%]), localized scleroderma (either morphea or the linear form of this disease) (48 patients [3.0%]), available records inadequate to establish the diagnosis (43 patients [2.7%]), or other diagnoses (usually, but not always, related to connective tissue diseases, including SLE, mixed connective tissue disease, or undifferentiated connective tissue disease) (191 patients [12.0%]).

Of the 706 eligible patients identified, 595 (84.3%) met the classification criteria for definite SSc (21), and 111 (15.7%) were considered as having probable SSc. A comparison of definite versus probable cases of SSc indicated no significant differences in the demographic characteristics of sex and race, the frequency of CREST features, the frequency of anticentromere antibodies (ACAs), or overall survival (data not shown).

The contributions of each of the 5 sources to the final patient pool are shown in Table 1. The total for each category represents the number of individual patients identified by that source; however, many patients were identified by more than one source. A total of 328 eligible patients were identified by 1 source only, 255 were identified by 2 sources, 108 were identified by 3 sources, and 15 patients were identified by 4 sources. No case was captured by all 5 sources. Of the 706 unique patients included in this study, 603 (85.4%) were identified in the hospital database. Less than half of the eligible subjects identified were patients from the academic centers (Wayne State University and the University of Michigan), with the remainder coming from the hospital discharge search, private practices, and the Scleroderma Foundation.

Prevalence of SSc. Table 2 shows the overall age-adjusted prevalence of SSc, as well as the prevalence by sex and race, in the study population. Based on the 706 verified cases, the overall prevalence of SSc was determined to be 242.0 cases per million adults (95% CI 213–274). The female-to-male ratio was 4.6:1.0, with a sex-specific prevalence of 389.8 cases per million female

Table 1. Sources of eligible patients with scleroderma in Michigan*

Source	No. of patients potentially eligible	No. of patients eligible
Academic centers (Wayne State University and University of Michigan)	533	310
Hospital discharge search (Health Care Investment Analysts)	1,428	603
Scleroderma Foundation	464	236
Non-university rheumatologists	155	124
Other physicians and self-referrals	22	13

* Within a single source category, the patients are unique. However, there are overlaps among categories: 328 eligible cases were identified by only 1 source, 255 by 2 sources, 108 by 3 sources, and 15 by 4 sources. No case was captured by all 5 sources.

adults (1 case per 2,618 women), compared with 84.1 cases per million male adults (1 case per 11,905 men). SSc prevalence was higher among blacks compared with whites, with an adjusted prevalence ratio of 1.15 (95% CI 1.02–1.30). Differences in prevalence between white women and black women and between white men and black men did not reach statistical significance ($P = 0.089$ and $P = 0.179$, respectively).

Although the mean (±SD) age at diagnosis was not different for female and male patients (46.0 ± 15.8 years versus 46.7 ± 16.9 years; $P = 0.692$), the age at diagnosis for black patients was significantly younger than that for white patients ($P < 0.001$), among both men and women. This correlates with the finding that diffuse disease, which is characterized by both a younger age at onset and a shorter interval between symptom onset and diagnosis, was more common in the black population.

Incidence of SSc. Table 3 presents the characteristics of the incident cases (those newly diagnosed during the period 1989–1991). One hundred sixty-nine patients were identified, for an estimated annual incidence in the US of 19.3 new cases of SSc per million adults. The female-to-male ratio for incident cases (3.2:1.0) is different from that for prevalent cases (4.6:1.0), most likely due to the survival differential between men and women with this disease (see survival curves, below). The overall age-adjusted incidence also differed between the black population (23.7 cases per million) and the white population (18.3 cases per million), but this difference did not reach statistical significance. The adjusted incidence ratio was 1.05 (95% CI 0.79–1.38).

Among incident cases, the mean age of patients at diagnosis did not differ according to sex ($P = 0.334$). However, age at diagnosis did differ by race, with SSc occurring earlier in black patients compared with white patients ($P < 0.001$). This difference was significant for the comparison of age at diagnosis for all white patients compared with all black patients ($P < 0.001$) and for white women compared with black women ($P < 0.001$). Age at diagnosis was not significantly different between black male patients and white male patients, but the number of patients in these subgroups was small.

The age-specific incidence of SSc according to race and sex is illustrated in Figure 2. Among black women, the peak incidence occurred between the ages

Table 2. Prevalence and mean age at diagnosis of SSc in the Detroit tricounty area for the period 1989–1991*

	All subjects			Women			Men		
Group	No. (%)	Prevalence (95% CI)	Age at diagnosis, mean ± SD years	No. (% of all subjects)	Prevalence (95% CI)	Age at diagnosis, mean ± SD years	No. (% of all subjects)	Prevalence (95% CI)	Age at diagnosis, mean ± SD years
Total	706 (100)	242.0 (213–274)	46.1 ± 15.8	591 (83.7)	389.8 (353–430)	46.0 ± 15.8	115 (16.3)	84.1 (68–104)	46.7 ± 16.9
Black	186 (26.3)	315.1 (282–352)	41.0 ± 14.6	157	433.5 (395–476)	40.8 ± 14.5	29	103.9 (86–126)	40.4 ± 15.5
White	505 (71.5)	224.7 (197–256)	48.1 ± 15.9	423	370.6 (335–410)	48.1 ± 15.7	82	78.4 (63–98)	48.3 ± 17.0
Unknown/other race	15 (2.1)	–	–	11	–	–	4	–	–

* Prevalence values are age-adjusted based on the Detroit tricounty area adult population of 2.917 million according to 1990 US census: 75.0% white (2.189 million), 21.9% black (638,700), and 3.1% other (90,400). SSc = systemic sclerosis; 95% CI = 95% confidence interval.

Table 3. Annual incidence and mean age at diagnosis of SSc in the Detroit tricounty area, 1989–1991*

Group	All subjects			Women			Men		
	No. (%)	Annual incidence (95% CI)	Age at diagnosis, mean ± SD years	No. (%)	Annual incidence (95% CI)	Age at diagnosis, mean ± SD years	No. (%)	Annual incidence (95% CI)	Age at diagnosis, mean ± SD years
Total	169 (100)	19.3 (12.4–30.2)	52.2 ± 15.2	133 (78.7)	28.5 (19.7–41.1)	53 ± 14.9	36 (21.3)	9.0 (4.7–17.3)	51.8 ± 16.4
Black	38 (22.5)	23.7 (15.8–35.4)	43.8 ± 12.3	33	31.1 (21.9–44.1)	43.4 ± 12.7	5	6.2 (2.8–13.6)	46.8 ± 10.8
White	127 (75.1)	18.3 (11.6–28.9)	55.5 ± 15.2	97	27.0 (18.5–39.4)	56.5 ± 14.4	30	9.7 (5.2–18.2)	52.4 ± 17.4
Unknown/other race	4	–	–	3	–	–	1	–	–

* Incidence values are age-adjusted based on the Detroit tricounty area adult population. SSc = systemic sclerosis; 95% CI = 95% confidence interval.

of 45 and 54 years, whereas the peak incidence among white women occurred in the 65–74-year-old age group. The peak incidence for black men paralleled that for black women. The pattern for white men was less distinct, with a gradually increasing incidence until the ages of 75 to 84 years.

To estimate the extent to which this study was not able to identify all incident and prevalent cases of SSc, a capture-recapture analysis was performed (30,31). According to this analysis, an estimated 103 SSc cases were not identified by our multiple sources. Therefore, the revised prevalence estimate for the Detroit area was 809 cases, and the revised prevalence estimate for the US was 276 cases per million adults (95% CI 245–310). The revised estimate of the annual incidence was 21 new cases per million adults per year (a 12% increase).

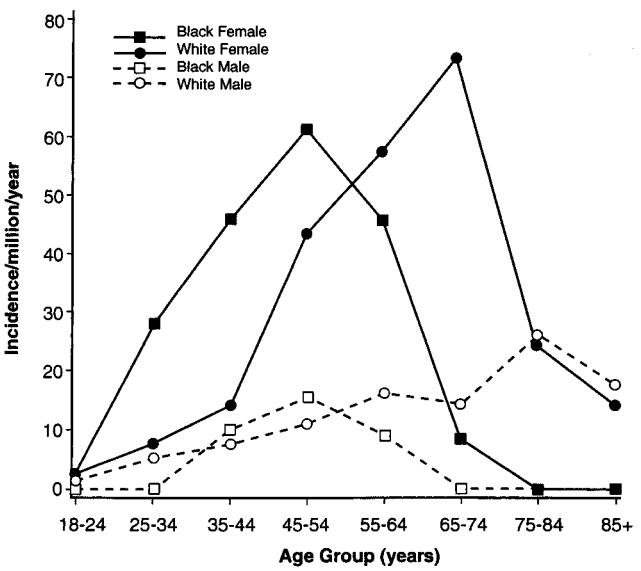

Figure 2. Age-specific incidence of systemic sclerosis by race and sex.

Survival. Two hundred fifteen deaths occurred in this study population; median survival after diagnosis was ~11 years. Absolute survival was 77.9% at 5 years, 55.1% at 10 years, 37.4% at 15 years, and 26.8% at 20 years. As illustrated in Figure 3, the observed survival was considerably less than the expected survival for a population matched for age, sex, and race. Relative survival (the ratio of observed to expected survival) was 62.6% at 10 years and only 35% at 20 years.

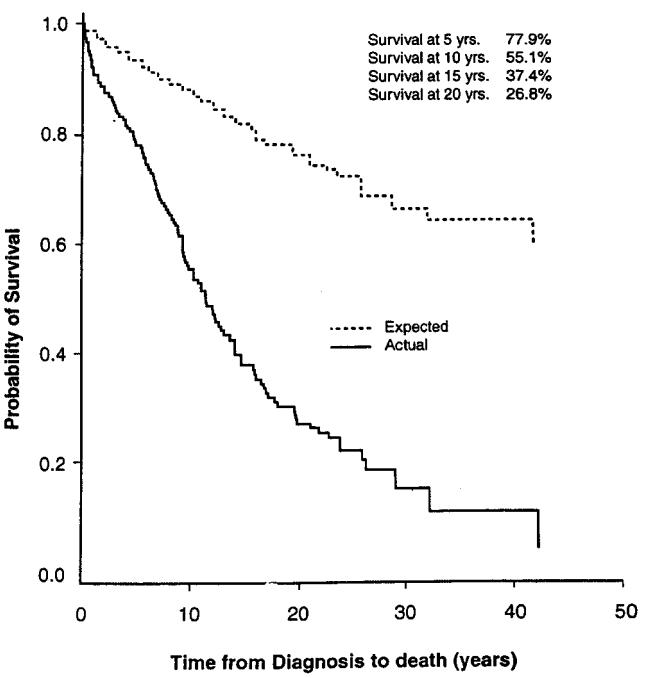

Figure 3. Overall survival from the time of diagnosis of systemic sclerosis. Overall survival was considerably diminished from that expected in a population matched for age and race. The percent survival represents absolute survival at the time points listed.

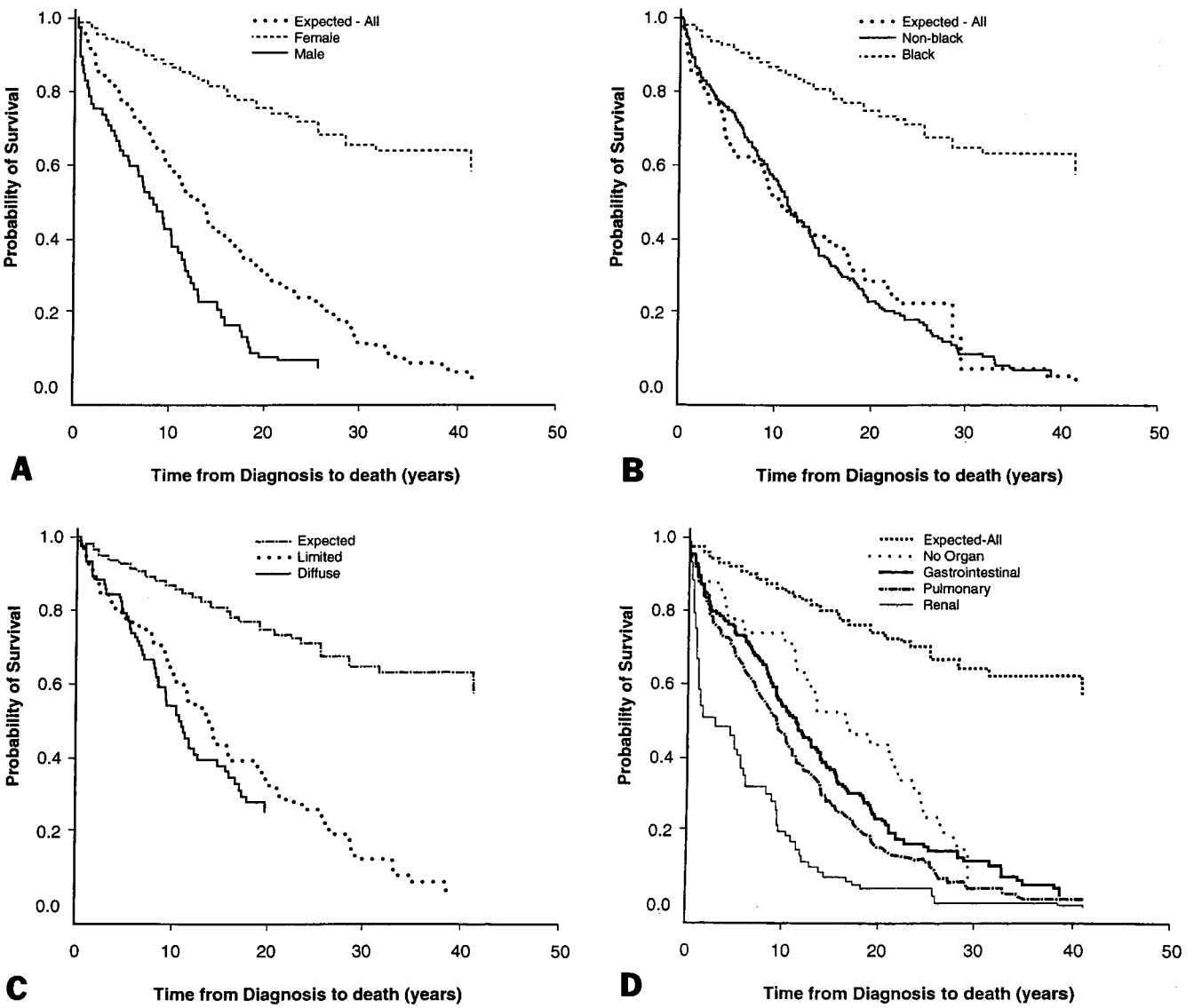

Figure 4. Survival plots, comparing patients with systemic sclerosis according to sex (**A**), race (**B**), skin involvement (**C**), and organ involvement (patients with involvement of multiple organ systems were considered redundantly in these calculations) (**D**).

Age at diagnosis significantly influenced survival ($P < 0.001$). The risk of death increased 5% for each 1-year increase in age at diagnosis (hazard ratio [HR] = 1.04, 95% CI 1.03–1.05).

As illustrated in Figure 4A, survival for female patients was better than that for male patients (for male sex, HR = 1.81, 95% CI 1.29–2.55, $P < 0.001$), and this difference became apparent early in the course of the disease. Survival for black patients versus non-black patients (Figure 4B) was marginally worse during the first 12 years after diagnosis, but in general, survival for both groups was comparable (for black race, HR = 1.15, 95% CI 0.82–1.61, $P = 0.419$). As illustrated in Figure 4C, diffuse disease appeared to be associated with a poorer outcome than did limited skin involvement, but this comparison did not reach statistical significance (for diffuse disease, HR = 1.23, 95% CI 0.85–1.78, $P = 0.28$). Figure 4D shows survival according to organ system involvement. Renal involvement had a considerable impact on both short-term (<5 years) and longer-term (>10 years) survival. Pulmonary disease also contributed substantially to mortality, as did gastrointestinal involvement. Even those individuals without involvement of these specific organs had reduced survival. As

noted above, cardiac disease was not included in our record review.

Extent of disease. Among the 706 prevalent cases, 473 patients (67.0%) could be characterized as having either limited or diffuse disease. As illustrated in Figure 5, diffuse disease occurred less commonly than did limited disease among incident cases (33.8% versus 66.2%) as well as among prevalent cases (34.9% versus 65.1%), suggesting that this difference is not attributable to the survival advantage of patients with limited disease. In addition, we found a difference between racial groups in the extent of disease. Overall, 60.3% of black patients were designated as having diffuse disease, and 39.7% were designated as having limited skin involvement. In contrast, 26.6% of non-black patients had diffuse disease, compared with 73.4% with limited disease. The proportion of black patients for whom the extent of disease was not recorded was not different from that of non-black patients with unknown disease extent (data not shown).

Expressed as a prevalence proportion ratio, the prevalence of diffuse disease among black patients was 1.86 (95% CI 1.48–2.35) compared with that among white patients. However, the prevalence odds ratio (POR) based on these data indicated that black patients were 4.30 times more likely than white patients to have diffuse disease (95% CI 2.77–6.69). There was no difference between the sexes with regard to disease extent, and the proportions of disease type were roughly equal in men and women (POR = 0.98, 95% CI 0.56–1.70).

Serology. Table 4 lists the serologic characteristics of the prevalent cases according to race and sex. Anticentromere antibodies (ACAs) were observed in 22.1% of the patients tested, and anti–Scl-70 (anti–topoisomerase I) antibodies were found in 19.6%. No individual had documentation of the presence of both autoantibodies. ACA positivity was less common among black patients than among white patients ($P < 0.001$), as noted by other investigators (32,33), and was less common among male patients than among female patients ($P < 0.001$). As expected with this autoantibody pattern and as noted above, diffuse disease was more common in the black population compared with the non-black population.

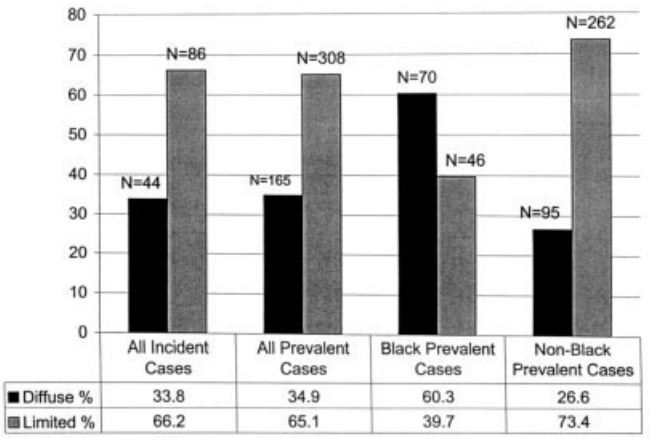

Figure 5. Distribution of diffuse and limited disease in patients with systemic sclerosis.

Table 4. Serologic findings*

Group	% ANA positive (no. positive/no. tested)	% ACA positive (no. positive/no. tested)	% Scl-70 positive (no. positive/no. tested)
Total	89.3 (482/540)	22.1 (112/506)	19.6 (39/199)
Race			
White	90.1 (345/383)	27.0 (97/359)	18.1 (26/144)
Black	86.8 (125/144)	9.7 (13/134)†	22.0 (11/50)
Other/unspecified	92.3 (12/13)		
Sex			
Female	89.7 (410/457)	25.1 (107/427)	19.4 (32/165)
Male	86.7 (72/83)	6.3 (5/79)‡	20.6 (7/34)

* ANA = antinuclear antibody; ACA = anticentromere antibody.
† $P < 0.001$ versus white race.
‡ $P < 0.001$ versus female sex.

DISCUSSION

This study of prevalence and incidence is based on the largest SSc population in the US assembled to date. Multiple methods of case ascertainment were used, and cases included both inpatients and outpatients identified from university referral sources, community physicians, and patient support groups. The observed prevalence (based on the 706 identified and verified cases) is 242.0 cases per million adults. The estimated prevalence, using capture-recapture analysis of 809 presumed cases, is 276 cases per million adults. A major assumption of this method is that the sources of case ascertainment are independent; that is, that identification by one source (e.g., hospital discharge database) is not related to identification by another source (e.g., physician office). Although this assumption frequently is not valid, correction for dependency results in a larger estimate of total cases was not done for this study (30,31).

Large population studies are necessary to provide reliable estimates of the prevalence and incidence of rare diseases. This approach, however, relies on available medical records, which frequently lack documenta-

tion of key clinical, laboratory, or other diagnostic testing. In clinical practice, a description of SSc features such as digital pitting scars or scars of old ulcers may not be documented in the physical findings but contribute to the diagnostic impression. As noted above, this study included probable as well as definite cases: probable cases included patients in whom SSc was diagnosed by a rheumatologist and who had documented sclerodactyly plus at least 2 additional features of the CREST syndrome.

Even using this approach, an additional source of potential underestimation was involved in case ascertainment. Among the 1,596 unique individuals identified as potentially eligible, medical records for 441 of these patients (27.6%) were unavailable for review. Assuming that the proportion of eligible individuals in this nonreviewed group would be the same as that in the group whose medical records were reviewed (61.1% of patients whose records were reviewed were eligible), an additional 265 cases would be added to our total. The 103 additional cases predicted by the capture-recapture analysis may be included in this group but would not account for the entire number of potentially eligible nonreviewed cases. Had the capture-recapture analysis used less stringent assumptions, the predicted number of cases would be greater but unlikely to increase by more than double. The prevalence estimate reported here should be interpreted as a conservative value.

Our estimate of an annual incidence of 19.3 new cases per million adults per year (95% CI 12.4–30.2) is similar to the incidence estimate of 18.2 per million reported by Steen et al (6) for the period 1978–1982 in Allegheny County, Pennsylvania. For the earliest study period (1963–1967) in the same area of Pennsylvania, these investigators reported a lower incidence of only 9.7 cases per million but speculated that the apparent increase in SSc may have been attributable to improved case detection rather than a true increase in incidence. Our data would suggest that the annual incidence of SSc has not increased but has been relatively stable, at least during the period 1978–1991.

Our predicted prevalence of 242.0 cases per million adults based on our 706 verified cases is similar to the prevalence estimate of 286 cases per million reported by Maricq et al (5), based on 2 cases of definite SSc in a South Carolina population. However, the small number of cases in the Maricq study precludes much confidence in this estimate. Our capture-recapture analysis suggests that the true population prevalence is at least 276 cases per million adults (95% CI 245–310), with an annual incidence of 21 new cases per million adults per year.

In contrast to this relative agreement among recent US reports, similar epidemiology studies in other countries suggest far lower SSc prevalence and incidence estimates. In England, the prevalence of SSc has been reported as 30.8 cases per million (based on 156 cases) (34), in Japan the prevalence estimate is 38 cases per million (based on 357 cases) (35), and in Iceland the prevalence estimate is 71 cases per million (based on 37 cases) (36). These studies included both limited and diffuse forms of SSc. Although methods of case ascertainment and verification and the length of the study period vary among these reports, it is unlikely that methodologic differences alone account for this considerable disparity. For example, the difference in point prevalence between the British study and the current report is 8-fold (30 cases per million versus 242 cases per million). It would therefore appear that SSc occurs more commonly in the US than in other countries in which it has been systematically investigated.

Estimates of SSc occurrence similar to those in our study were reported among residents of the Australian state of South Australia, by Roberts-Thomson et al (37). Those investigators estimated the point prevalence of SSc to be 233 cases per million (based on 348 cases), with an annual incidence of 16 cases per million. In contrast, a lower prevalence of SSc was reported in the Australian city of Sydney (state of New South Wales) (86 cases per million for 1988) (38). Reasons for the difference between these 2 adjacent regions are not clear. Similar female-to-male ratios and similar proportions of limited and diffuse disease were reported in both studies. The racial and ethnic backgrounds of patients were not described in either report.

The highest prevalence of SSc has been reported in a Choctaw Native American group in Oklahoma (660 cases per million, based on 14 cases) (39). Genetic influences have been proposed to account for this increase but have not yet been definitively identified (39; for review, see ref. 40). In addition, no specific environmental factors have been recognized (39).

In terms of serology, ACAs were found more frequently in women than in men, a finding that has not been reported previously. As in other studies, ACAs were less frequent in black patients than in non-black patients (32,33). Of the 20 prevalent cases involving black men (5 of whom had limited disease), no individual had ACA positivity. Due to the retrospective nature of this study and the fact that individuals were seen by various physicians, a complete autoantibody profile was

not available for all patients, nor was the designation in the medical record regarding the extent of skin involvement always clear enough to permit classification of the disease as limited or diffuse.

For patients in this cohort, median survival from the time of diagnosis was ~11 years. This represents an improvement in survival compared with the 35% survival at ~10 years from diagnosis reported by Medsger and Masi in 1971 (2). The survival rates reported in different studies vary widely, depending to a large extent on the proportion of limited and diffuse disease in the cohort (for review, see ref. 19). In the current study, survival from the time of diagnosis was 77.9% at 5 years, 55.1% at 10 years, 37.4% at 15 years, and 26.8% at 20 years. Observed survival was substantially less than expected for the case population overall as well as for all subgroups. The deleterious effect of renal involvement on long-term survival likely does not reflect the new onset of renal crisis late in the course of the disease, but rather the fact that individuals who have experienced renal crisis, most of whom are left with impaired renal function, have a poorer prognosis.

It should be noted that the mean age at diagnosis for the prevalent case patients was 46.1 years (Table 2), which is younger than that for the incident case patients (52.2 years) (Table 3). This apparent discrepancy in age at diagnosis between incident cases and prevalent cases may be attributable to the survival advantage of younger individuals in terms of all-cause mortality; that is, patients diagnosed at an earlier age have a greater likelihood of surviving to be counted in the prevalence period. Alternatively, older prevalent case patients might have been systematically excluded if they were no longer being followed up by rheumatologists or were not hospitalized.

Both the prevalence and incidence of SSc were higher in blacks than in whites, although the comparison for incidence did not reach statistical significance. Incident cases in black patients may have been missed more frequently than were incident cases in white patients if there was a longer delay before diagnosis or a delay before specialist referral for black patients compared with patients of other races. Also, and as noted above, because the mean age at diagnosis was younger among black patients than among white patients, more black patients may have survived to be counted in the prevalence group.

We found a rather marked difference in the proportion of diffuse and limited disease between racial groups (black patients were 1.86 times more likely than non-black patients to have diffuse disease), suggesting that disease expression is different between racial groups. Alternatively, we could have missed cases of limited disease among black patients due to a delayed diagnosis or specialist referral compared with that in other racial groups. If this were the case, then the black/white differential in prevalence and incidence is even higher than we report here. Although our data do not permit us to resolve this issue, it is an important consideration, because it affects our interpretation of the role of genetic influences on disease occurrence and expression.

In summary, both sex and race appear to have a strong influence on the incidence (i.e., disease susceptibility) of SSc, whereas race additionally influences disease expression, particularly age at onset and extent of skin involvement. This implies that sex-related risk factors (e.g., hormonal, reproductive, sex-specific environmental exposures and the like) may serve as or amplify the initiating trigger for the disease, whereas genetic factors may exert a stronger influence on disease severity.

This study firmly establishes baseline estimates of SSc disease occurrence and expression in a large US population consisting primarily of black adults and white adults residing in a metropolitan area. These data should facilitate research regarding the role of geographic, ethnic, racial, genetic, and environmental factors in SSc susceptibility and disease expression in comparison populations.

ACKNOWLEDGMENTS

We thank Swan Atty, Scleroderma Registry coordinator, for assisting in this project. We also thank the referring physicians for their cooperation in identifying patients and the Michigan Chapter of the Scleroderma Foundation for promoting the study to their members.

REFERENCES

1. Kurland LT, Hauser WA, Ferguson RH, Holley KE. Epidemiologic features of diffuse connective tissue disorders in Rochester Minn., 1951 through 1967, with special reference to systemic lupus erythematosus. Mayo Clin Proc 1969;44:649–63.
2. Medsger TA Jr, Masi AT. Epidemiology of systemic sclerosis (scleroderma). Ann Intern Med 1971;74:714–21.
3. Medsger TA Jr, Masi AT. The epidemiology of systemic sclerosis (scleroderma) among male U.S. veterans. J Chronic Dis 1978;31: 73–85.
4. Michet CJ Jr, McKenna CH, Elveback LR, Kaslow RA, Kurland LT. Epidemiology of systemic lupus erythematosus and other connective tissue disease in Rochester, Minnesota, 1950 through 1979. Mayo Clin Proc 1985;60:105–13.
5. Maricq HR, Weinrich MC, Keil JE, Smith EA, Harper FE, Nussbaum AI, et al. Prevalence of scleroderma spectrum disorders

in the general population of South Carolina. Arthritis Rheum 1989;32:998–1006.
6. Steen VD, Oddis CV, Conte CG, Janoski J, Casterline GZ, Medsger TA Jr. Incidence of systemic sclerosis in Allegheny County, Pennsylvania: a twenty-year study of hospital-diagnosed cases, 1963–1982. Arthritis Rheum 1997;40:441–5.
7. Mayes MD. Scleroderma epidemiology. Rheum Dis Clin North Am 1996;22:751–64.
8. Medsger TA Jr, Masi AT. Survival with scleroderma. II. A life-table analysis of clinical and demographic factors in 358 male U.S. veteran patients. J Chron Dis 1963;26:647–60.
9. Medsger TA Jr, Masi AT, Rodnan GP, Benedek TG, Robinson H. Survival with systemic sclerosis (scleroderma): a life-table analysis of clinical and demographic factors in 309 patients. Ann Intern Med 1971;75:369–76.
10. Barnett AJ, Miller MH, Littlejohn GO. A survival study of patients with scleroderma diagnosed over 30 years (1953-1983): the value of a simple cutaneous classification in the early stages of the disease. J Rheumatol 1988;15:276–83.
11. Silman AJ. Mortality from scleroderma in England and Wales 1968-1985. Ann Rheum Dis 1991;50:95–6.
12. Altman RD, Medsger TA Jr, Bloch DA, Michel BA. Predictors of survival in systemic sclerosis (scleroderma). Arthritis Rheum 1991;34:403–13.
13. Lee P, Langevitz P, Alderdice CA, Aubrey M, Baer PA, Baron M, et al. Mortality in systemic sclerosis (scleroderma). Q J Med 1992;82:139–48.
14. Kaburaki J, Lee CC, Kuwana M, Tojo T, Ikeda Y, Takano M, et al. Initial predictors of survival in patients with systemic sclerosis (scleroderma). Keio J Med 1992;41:141–5.
15. Bond C, Pile KD, McNeil JD, Ahern MJ, Smith MD, Cleland LG, et al. South Australian Scleroderma Register: analysis of deceased patients. Pathology 1998;30:386–90.
16. Jacobsen S, Ullman S, Shen GQ, Wiik A, Halberg P. Influence of clinical features, serum antinuclear antibodies, and lung function on survival of patients with systemic sclerosis. J Rheumatol 2001;28:2454–9.
17. Scussel-Lonzetti L, Joyal F, Raynauld JP, Roussin A, Rich E, Goulet JR, et al. Predicting mortality in systemic sclerosis: analysis of a cohort of 309 French Canadian patients with emphasis on features at diagnosis as predictive factors for survival. Medicine (Baltimore) 2002;81:154–67.
18. Ferri C, Valentini G, Cozzi F, Sebastiani M, Michelassi C, La Montagna G, et al. Systemic sclerosis: demographic, clinical, and serologic features and survival in 1,012 Italian patients. Medicine (Baltimore) 2002;81:139–53.
19. Silman AJ. Scleroderma: demographics and survival. J Rheumatol Suppl 1997;48:58–61.
20. Jones MD, Castillo LA, Hopkins CA, Aaron WS, editors. St. Anthony's ICD-9-CM code book for physician payment. 5th ed. Reston (VA): St. Anthony Publishing; 1996.
21. Preliminary criteria for the classification of systemic sclerosis (scleroderma): Subcommittee for Scleroderma Criteria of the American Rheumatism Association Diagnostic and Therapeutic Criteria Committee. Arthritis Rheum 1980;23:581–90.
22. Tan EM, Cohen AS, Fries JF, Masi AT, McShane DJ, Rothfield NF, et al. The 1982 revised criteria for the classification of systemic lupus erythematosus. Arthritis Rheum 1982;25:1271–7.
23. LeRoy EC, Black C, Fleischmajer R, Jablonska S, Krieg T, Medsger TA Jr, et al. Scleroderma (systemic sclerosis): classification, subsets and pathogenesis. J Rheumatol 1988;15:202–5.
24. Schottenfeld D, Burns CJ, Gillespie BW, Laing TJ, Mayes MD, Heeringa SG, et al. The design of a population-based case-control study of systemic sclerosis (scleroderma): commentary on the University of Michigan study. J Clin Epidemiol 1995;48:583–6.
25. Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. J Am Stat Assoc 1958;53:457–81.
26. Office of the State Registrar and Center for Health Statistics. Michigan Health Statistics. Lansing: Michigan Department of Public Health; 1993.
27. Cox DR. Regression models and life tables [with discussion]. J Royal Stat Soc B 1972;74:187–220.
28. Lawless JF. Statistical models and methods for lifetime data. New York: John Wiley & Sons; 1982.
29. Wittes JT, Colton T, Sidel VW. Capture-recapture methods for assessing the completeness of case ascertainment when using multiple information sources. J Chron Dis 1974;27:25–36.
30. Hook EB, Regal RR. The value of capture-recapture methods even for apparent exhaustive surveys: the need for adjustment for source of ascertainment intersection in attempted complete prevalence studies. Am J Epidemiol 1992;135:1060–7.
31. Hook EB, Regal RR. Accuracy of alternative approaches to capture-recapture estimates of disease frequency: internal validity analysis of data from five sources. Am J Epidemiol 2000;152: 771–9.
32. Laing TJ, Gillespie BW, Toth MB, Mayes MD, Gallavan RH Jr, Burns CJ, et al. Racial differences in scleroderma among women in Michigan. Arthritis Rheum 1997;40:734–42.
33. Reveille JD, Fischbach M, McNearney T, Friedman AW, Aguilar MB, Lisse J, et al. Systemic sclerosis in 3 US ethnic groups: a comparison of clinical, sociodemographic, serologic, and immunogenetic determinants. Semin Arthritis Rheum 2001;30:332–46.
34. Silman A, Jannini S, Symmons D, Bacon P. An epidemiologic study of scleroderma in the West Midlands. Br J Rheumatol 1988;27:286–90.
35. Tamaki T, Mori S, Takehara K. Epidemiologic study of patients with systemic sclerosis in Tokyo. Arch Dermatol Res 1991;283: 366–71.
36. Geirsson AJ, Steinsson K, Guthmundsson S, Sigurthsson V. Systemic sclerosis in Iceland: a nationwide epidemiological study. Ann Rheum Dis 1994;53:502–5.
37. Roberts-Thomson PJ, Jones M, Hakendorf P, Dharmapatni AA, Walker JG, MacFarlane JG, et al. Scleroderma in South Australia: epidemiological observations of possible pathogenic significance. Intern Med J 2001;31:220–9.
38. Englert H, Small-McMahon J, Davis K, O'Connor H, Chambers P, Brooks P. Systemic sclerosis prevalence and mortality in Sydney 1974-1988. Aust N Z J Med 1999;29:42–50.
39. Arnett FC, Howard RF, Tan F, Moulds JM, Bias WB, Durban E, et al. Increased prevalence of systemic sclerosis in a Native American tribe in Oklahoma: association with an Amerindian HLA haplotype. Arthritis Rheum 1996;39:1362–70.
40. Tan FK, Stivers DN, Arnett FC, Chakraborty R, Howard R, Reveille JD. HLA haplotypes and microsatellite polymorphisms in and around the major histocompatibility complex region in a Native American population with a high prevalence of scleroderma (systemic sclerosis). Tissue Antigens 1999;53:74–80.