2. Materials and Methods

2.1. Geographical and demographical aspects of Norway

Norway is situated between latitude 58–71°N, and longitude 4–31°E. The area amounts to approximately 300,000 square kilometres (Fig. 1).

The southern part of the country is divided by a high mountain plateau into eastern and western regions. In the eastern regions, agriculture and forestry represented the main occupation far into the 20th century. In the western regions, the land is split up by numerous fjords along which the population has been localized. The areas of agriculture and forestry are small but have been intensively exploited. Fishing has represented an important trade.

In the northern part of the country, Trøndelag and North Norway, the conditions along the coast correspond well with the western regions of South Norway. The inland is more similar to eastern regions of South Norway. With some exceptions, all inhabitants have been living near the coast where fishing has been an important trade.

An inland *climate* (Historical Statistics, 1968) is found in eastern regions of South Norway; dry with a relatively low temperature during the winter and high in the summer. The highest monthly 24-h mean air temperature in Oslo (Fig. 1), average 1931–60, is approximately 17°C recorded in July, and the lowest is approximately −5°C recorded in January. The mean amount of rainfall is approximately 800 mm per year.

Western regions have a coastal climate; humid with a relatively higher temperature during the winter and lower in the summer. The highest monthly mean air temperature in Bergen (Fig. 1) is approximately 15°C recorded in July, and the lowest approximately 1°C recorded in January. The mean amount of rainfall is approximately 1900 mm per year.

In Trøndelag and North Norway a combination of coastal and inland climate is found, but more cool than in South Norway. In Tromsø (Fig. 1) the highest mean air temperature is approximately 12°C recorded in July, and the lowest approximately −4°C recorded in January. The mean amount of rainfall is approximately 1000 mm per year.

The climate does not appear to have changed considerably during the last 130 years (Mohn, 1921).

The *total population* of Norway (Historical Statistics, 1968) increased from 1,490,047 in 1855 to 2,649,775 in 1920 by which time all leprosy patients except 14 had been taken ill. Population per square kilometre increased from

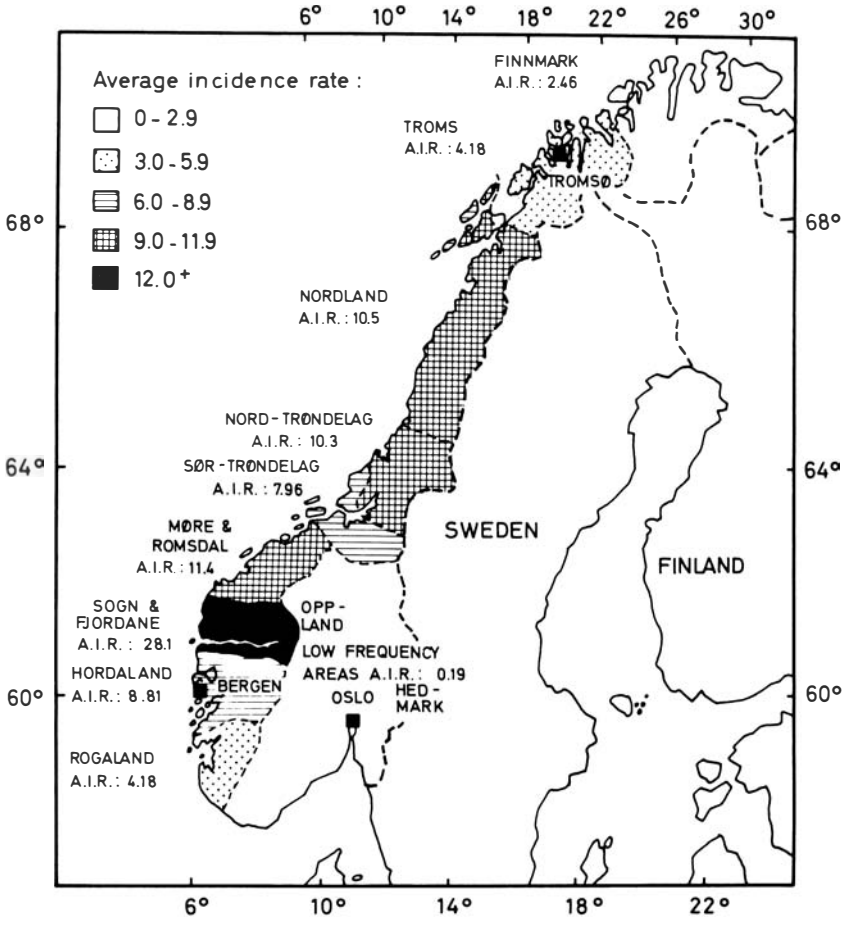


Figure 1. Average incidence rates (A.I.R.) of leprosy in Norway 1851–1920 by county. (The National Leprosy Registry of Norway.)

4·8 in 1855 to 8·6 in 1920. During the same period the number of females to 1,000 males increased from 1,041 to 1,053. In 1855, 36·4% of all males and 33·6% of all females were less than 15 years of age. In 1920, the corresponding figures were 33·6% and 30·6%. In 1855 86·7% of the total population lived in rural municipalities and 83·1% in sparsely populated areas. Corresponding figures in 1920 were 70·4% and 54·7%. In 1855, 53·4% of the population lived in the eastern regions of South Norway and 53·6% in 1920.

Accordingly, climatic and occupational conditions have not changed to any extent during the period 1855–1920. The population increased considerably and most of all in the towns, however, the composition of the population with respect to age and sex was not altered much.

2.2. Computerization of the registry information

According to the official statistics of Norway, 8,218 cases of leprosy were reported from 1856 to 1945 (Melsom, 1948). For each case a considerable amount of data has been collected from primary registration until death or permanent cure. Due to the amount of data, the compilation of the total material of the Leprosy Registry called for the application of computer methods described in the following paragraphs.

2.2.1. SOURCES OF DATA

At the central office (Irgens and Bjerkedal, 1973) the information was kept in two coordinate registers; the district register and the hospital register.

District register

Information was forwarded annually by the District Health Officer, in charge of his health district, to the central office in two separate report forms providing columns for the items of personal data to be recorded.

A *main form* (Fig. 2) was used for the notification of all new cases registered by the District Health Officers each year; either patients taken ill while living in the district or patients who migrated to the district after onset of the disease. The latter had already been notified in another district.

In a *follow up form* (Fig. 3), additional information on patients already registered was reported. The form was used to inform the central office when a patient left the district, either for hospitalization or migration to another district. The form was also used for the notification of deaths and to inform the central office when a case, after observation, turned out to represent another disease and not leprosy.

In the central office most of the information was transferred to the district register which was kept in handwritten books. The pages are structured with separate columns for the different items of data to be transferred (Fig. 4, Table 1). Information from a main form on one patient was entered on one line and is called a *patient report*. Each patient report was identified by a district serial number consisting of 7 digits.

The first digit referred to the number of the register book in which the patient report is found. The next three digits referred to the page number. A full page was reserved for patient reports from one particular health district. The last three digits referred to the number of the line on which the report is entered, starting from no. 001 for the first patient report forwarded from each health district.

Because more than one patient report per person might have been forwarded to the central office and entered into the district register, one patient might have been allotted more than one district serial number. In many cases cross

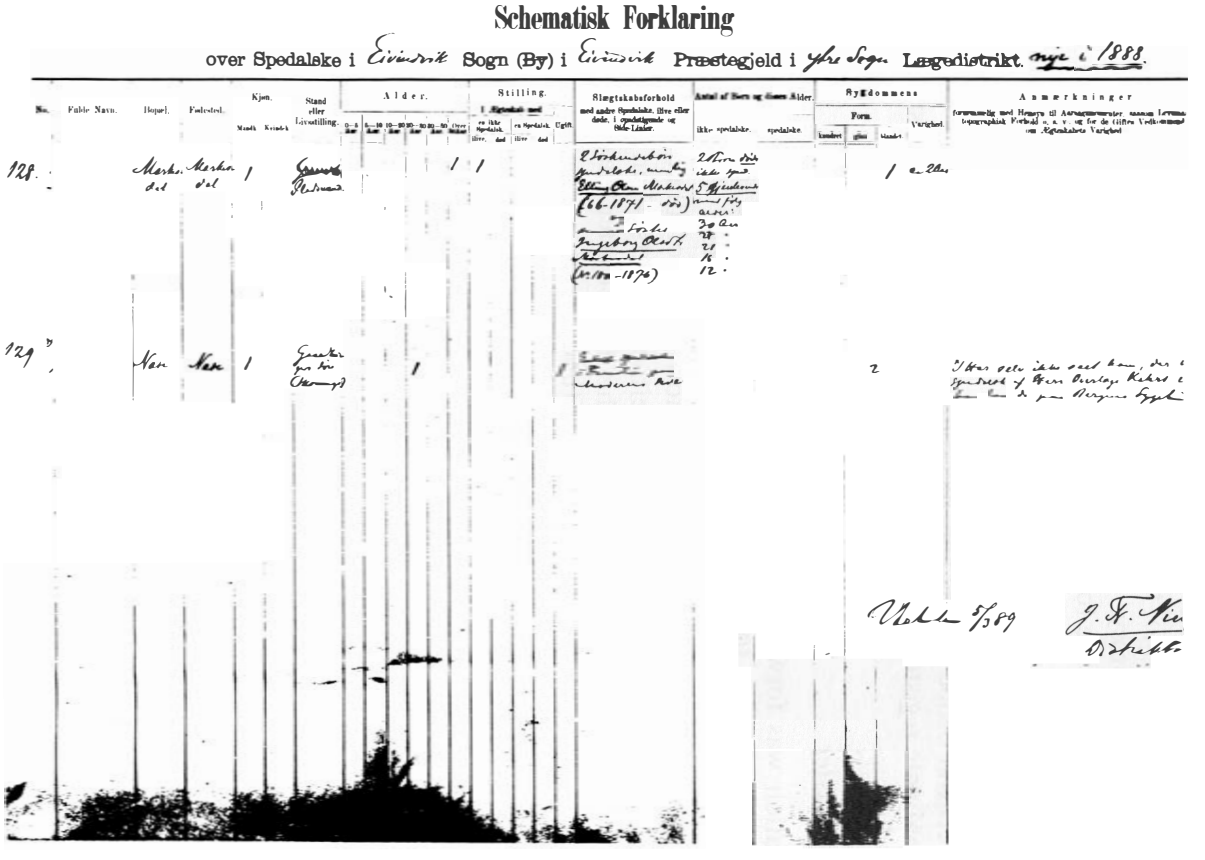
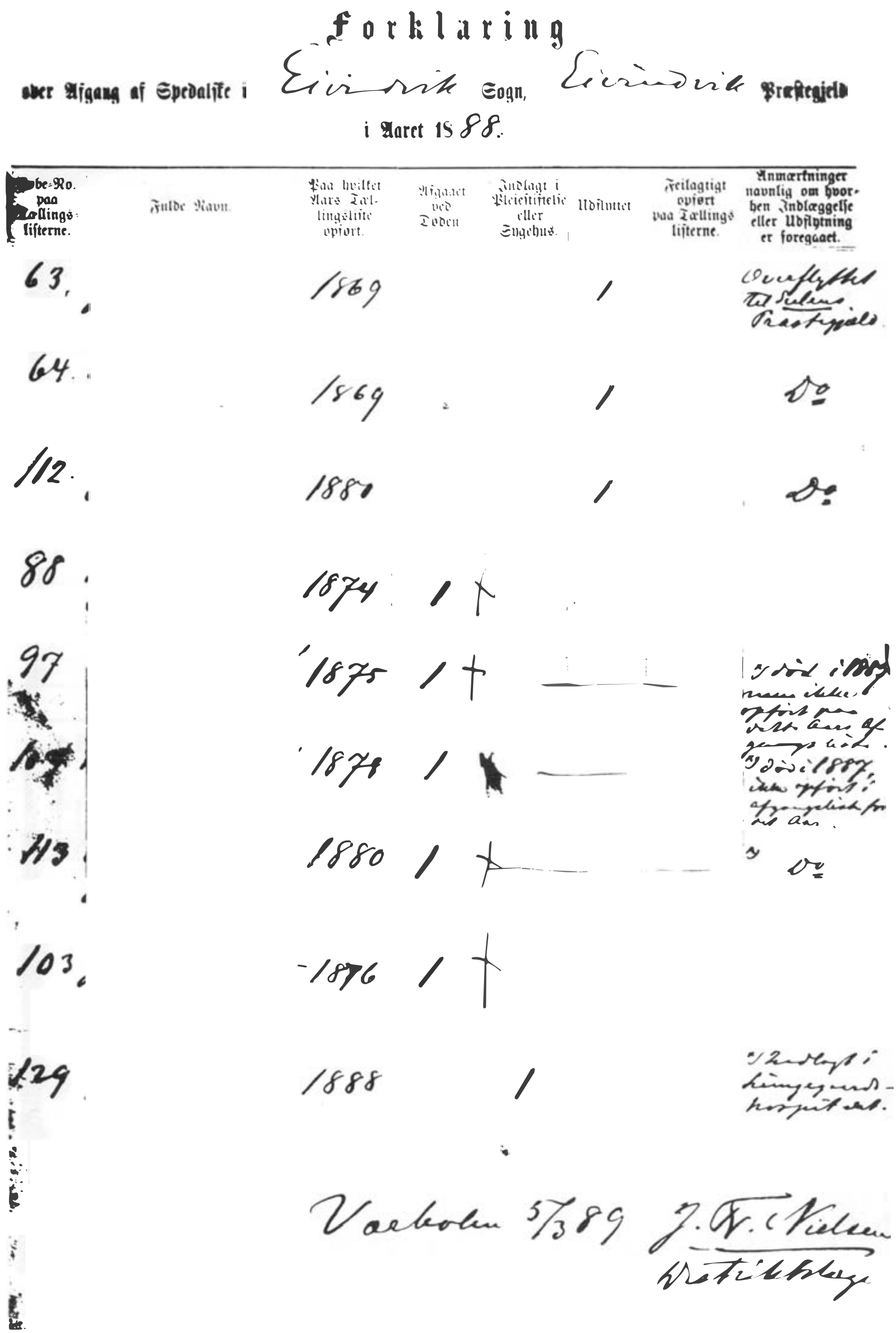


Figure 2. Form used for the notification by the District Health Officers of new leprosy cases to the Chief Medical Officer for Leprosy. (The National Leprosy Registry of Norway.)



Forklaring

over Afgang af Spedalske i Eivindvik Sogn, Eivindvik Præstegjeld

i Aaret 1888.

Løbe-No. paa Tællingslisterne.	Fulde Navn.	Paa hvilket Aars Tællingsliste opført.	Afgaaet ved Døden	Indlagt i Pleiestiftelse eller Sygehus.	Udflyttet	Feilagtigt opført paa Tællingslisterne.	Anmærkninger navnlig om hvorhen Indlæggelse eller Udflytning er foregaaet.
63.		1869			1		Overflyttet til Sulens Præstegjeld
64.		1869			1		Do
112.		1880			1		Do
88.		1874	1 †				
97		1875	1 †				død i 1887 men ikke opført paa dette Aars afgangsliste.
107		1878	1 †				død i 1887, ikke opført i afgangsliste for dette Aar.
113		1880	1 †				Do
103.		1876	1 †				
129		1888		1			Indlagt i Lungegaards-hospitalet.

Vaerholm 5/3 89 J. F. Nielsen

Distriktslæge

Figure 3. Form used for the reporting by the District Health Officers of follow-up information on cases already registered to the Chief Medical Officer for Leprosy. (The National Leprosy Registry of Norway.)

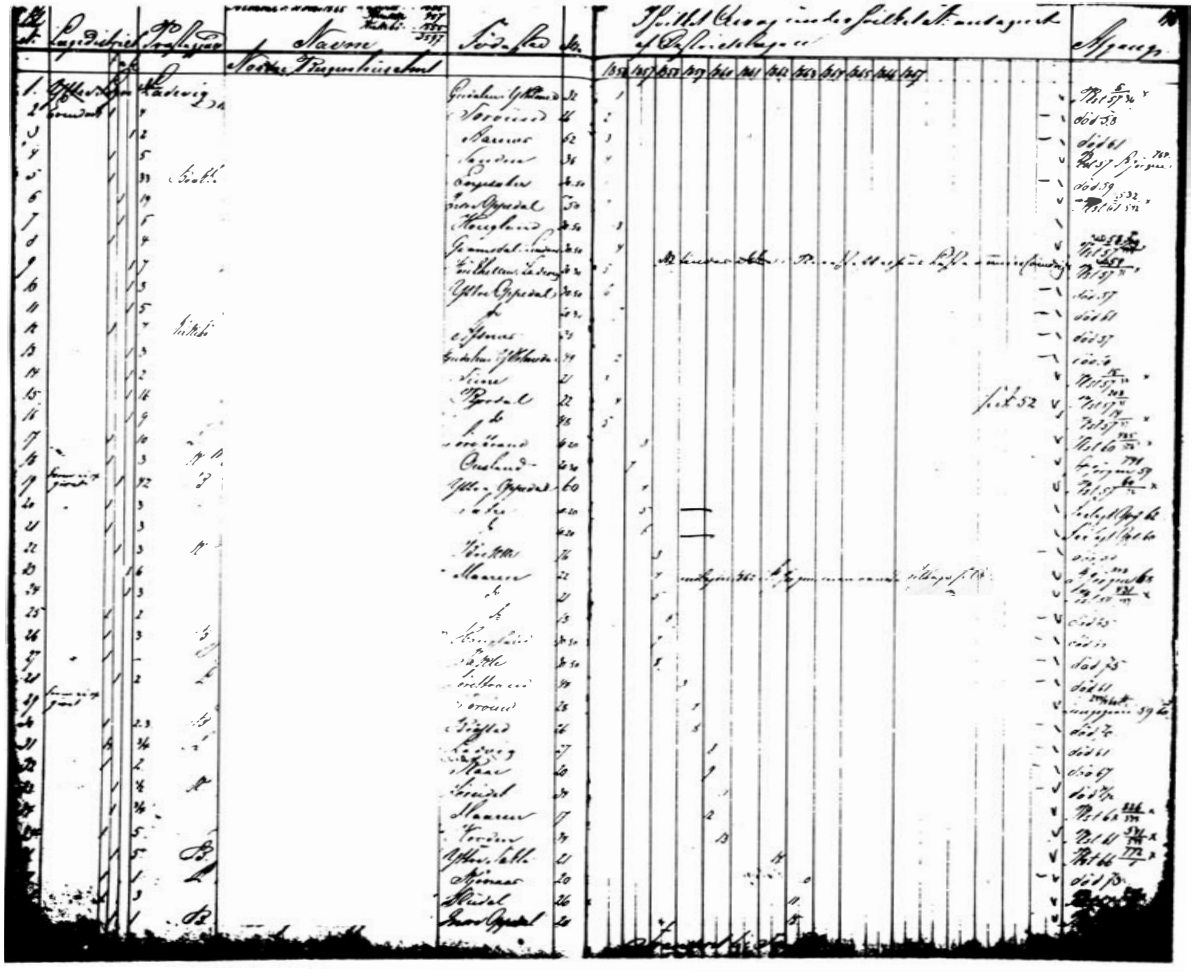


Figure 4. A page from a district register book. (The National Leprosy Registry of Norway.)

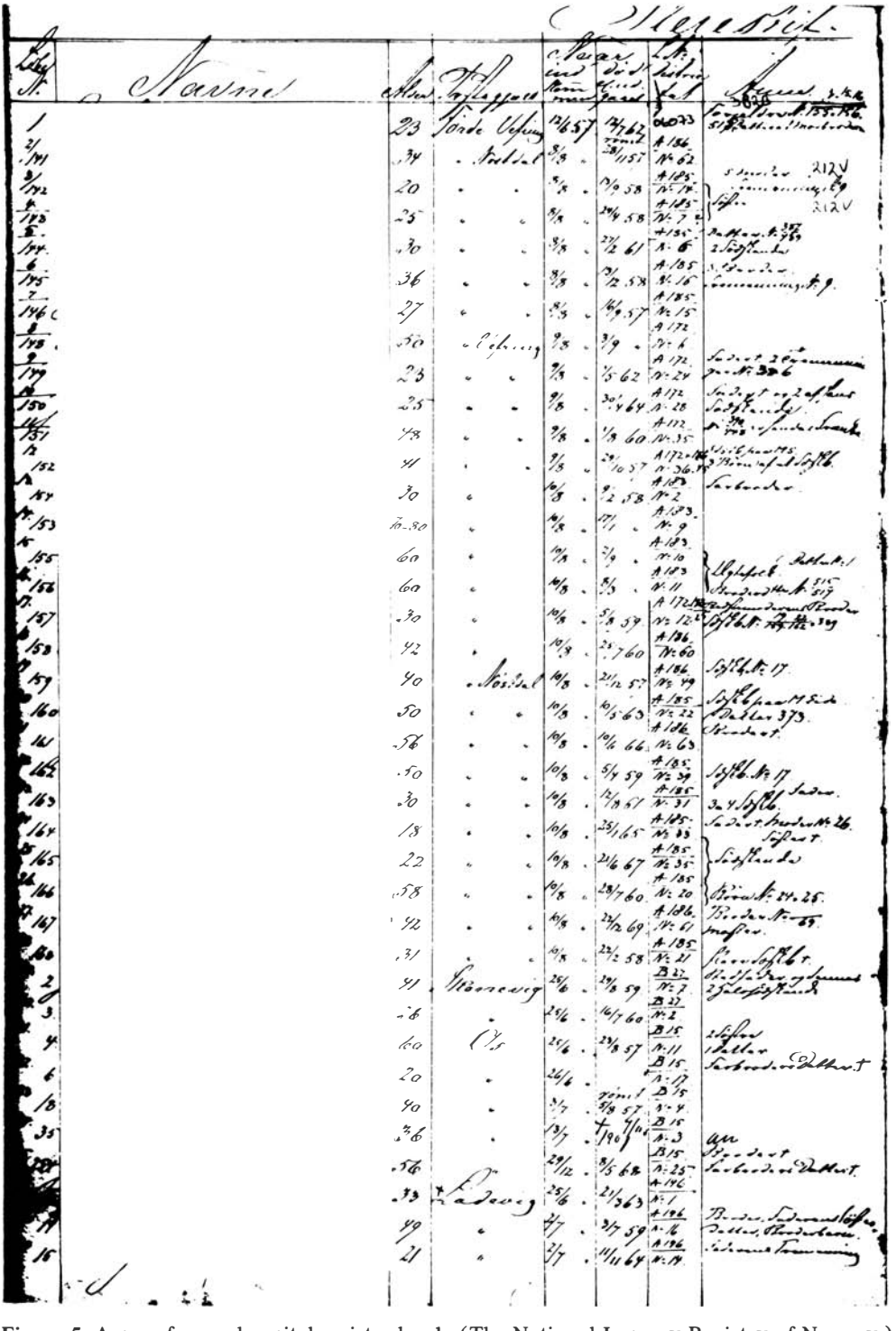


Figure 5. A page from a hospital register book. (The National Leprosy Registry of Norway.)

references were included in a patient report to previous reports on the same patient.

The content of the follow up form was entered in a separate column for additional information on the same line at the end of the patient report (Fig. 4).

Hospital register

The heads of the 5 leprosy hospitals (Irgens, 1973) annually forwarded information to the central office on all admissions and discharges. Apparently, no structured form was used. The information was entered into the handwritten books of the hospital register. The pages were structured with separate columns for the different items of data (Fig. 5, Table 1). Each patient report was identified by a particular *hospital serial number* allotted consecutively according to the date of admission. Each report was entered on a separate line, and each page contained reports from only one hospital. To a considerable extent district serial numbers of the patients were included in the reports.

Hospital patient records

The patient records of the leprosy hospitals represented a source of data in addition to the files of the Leprosy Registry. The records were kept in handwritten books, mostly unstructured, and were identified by the hospital serial number. The books were sent to the central office when the hospitals were gradually closed down.

2.2.2. PROCEDURE OF COMPUTERIZATION

The transfer of data was divided into separate steps; viz. choice of material, organization of patient records, technical transfer, identification of information, checking and completion, and search for doublet records.

Choice of material

Transfer of the total material appeared impossible due to the multitude of the sources. The criteria for the choice of material were related to the content of the different items and the form of their representation in the books.

The demands on *content* were defined without taking into consideration particular working hypotheses which might be of interest in the initial stages of the investigation. The intension was to transfer as much as possible of all information relevant to the epidemiology of leprosy. Thus hypotheses derived in the analytical stages of the investigation also might be tested.

The demands on *form* of representatation had to be carefully considered when a computer file comprising more than 8,000 patients was to be established,

Table 1. Items of personal data used in the district register and the hospital register. (The National Leprosy Registry of Norway)

Items of personal data	
District register	*Hospital register*
Serial number, district	Serial number, hospital
Name	Serial number, district
Residential district	Name
Birthplace	Residential district
Age when registered	Age when admitted
Duration of disease when registered	Duration of disease when admitted
Type of leprosy	Type of leprosy
Relatives with leprosy	Relatives with leprosy
Year of registration	Date of:
Eventual year of:	admission
admission to hospital	discharge
and what hospital	Conclusion:
migration	dead
and to what district	cured
cure	'erroneous' primary diagnosis
correction of diagnosis	deserted
death	expelled

linking together data from different sources to one patient record. To reduce the amount of work, only items which were given a representation easy to interpret and transform, where chosen. Furthermore, items which had not been entered in a standardized way into structured parts of the books with the application of a standardized and precise nomenclature were omitted.

Almost all items in the district register and the hospital register (Table 1) satisfied these demands and were transferred to the computer file. Since most of the information in the report forms had been transferred to the district register, the forms *per se* were not used as a source of data. The hospital patient records, however, were organized in a way which made a comprehensive transfer of the material impossible.

Organization of patient records and technical transfer

All information on each patient was assembled and transferred to one record, the *patient record*. The *content* of this patient record was defined by the choice of material. The *structure* of the patient record is to some extent related to the structure of the patient reports in the two registers (Table 2). The four main parts of the patient record consist of the patients' basic data, data from the district register, data on relatives with leprosy and data from the hospital register.

The data were transferred by punch cards containing one patient report per card. Direct punching was used for all data except for data on relatives affected by leprosy.

Table 2. Layout of the patient record in the computer file. (The National Leprosy Registry of Norway.)

Field No.	Position From	No. of Positions	Contents	Card
01	01	1	Sex	1
02	02	5	Identification no.	1
03	07	31	Name	1, (3)
04	38	4	Occupation	
05	42	1	Diagnosis correct/incorrect	
06	43	5	Birthplace	1
07	48	8	Birthyear	1, 3
08	56	4	Year of onset	1, 3
09	60	4	Review of family	1
10	64	4	Year of death	1, 3
11	68	19	*First registration*	1
1)	68	4	Registration year	1
2)	72	5	Registration district	1
3)	77	1	Type of leprosy	1
4)	78	1	Later information (migrated, cured, etc)	1
5)	79	4	Year	1
6)	83	4	Place	1
12	87	19	Second registration	1
13	106	19	Third registration	1
14	125	19	Fourth registration	1
15	144	18	Family information (9 relatives)	2
1)	144	9	*First relative*	2
(1)	144	2	Characterization (brother, father, etc)	2
(2)	146	7	Identification no.	2
16	225	33	First admission to hospital	3
1)	225	1	Hospital	3
2)	226	9	Hospital serial no.	3
3)	235	5	District admitted from	3
4)	240	1	Hospital admitted from	3
5)	241	7	Date of admission	3
6)	248	7	Date of discharge	3
7)	255	1	State when discharged (cured, dead, etc)	3
8)	256	1	Discharged to another hospital	3
9)	257	1	Type of leprosy	3
17	258	33	Second admission to hospital	3
18	291	33	Third admission to hospital	3
19	324	33	Fourth admission to hospital	3

Sets of codes were prepared for the direct punching of qualitative observations. The sets were applied by the punch operator for coding during the punching process.

Data on relatives with leprosy were transferred by a set of codes difficult to

apply during punching. The data had to be coded into a special form which served as a basis for punching.

Identification of patient reports and patient records

The utilization of the material in a longitudinal epidemiological study called for the introduction of an unambiguous system of identification, important to the technical transfer of data. Accordingly, a superior *identification number* was introduced. Each patient record was identified by its identification number, and a computer *catalogue* was established, linking all district serial numbers of a patient to his identification number.

The new identification number was introduced during punching of data from the patient reports in the district register. All patient reports which lacked a cross reference to a previous report in the district register, were given a new identification number. The remaining patient reports were identified by the identification number already allotted when previous reports were punched. The identification number was used for the linkage of data from different patient reports in the district register into the appropriate patient records.

The patient reports in the hospital register were identified during punching by one of the district serial numbers present in most of the reports. Together with the computer catalogue, these district serial numbers made a correct identification and linkage possible. The correct identification number for the reports in the hospital register which lacked a cross reference to a district serial number, had to be found manually in an alphabetical listing of all patient records established in the computer file. Some patients were reported only to the hospital register and never to the district register. In such cases a patient record was established on the basis of the patient report in the hospital register and a new identification number was allotted.

Information contained in a patient report on a relative with leprosy, was identified during coding by the district serial number of the patient and was entered into his patient record by the catalogue. Also the district serial number of the relative was entered into his patient record (Table 2, pos. 144-).

Checking and completion of data

To find and correct possible errors, all patient reports were listed, and invalid codes were indicated. The lists were proof-read against the original sources up to three times.

The content of the patient records was tested by a programme introducing fixed limits for all quantitative variables. Also relative limits for these variables were used to check the consistency of the information kept in one patient record.

Furthermore, the three relatively independent sources of data were used to compile a computer file as accurate, reliable and complete as possible. To a

considerable extent the patient reports in both registers contain basic personal data about the patients (Table 1). When data were transferred from a patient report to a patient record in the computer file, the consistency between the basic data possibly already present in the record and those in the patient report was checked by the computer. In a case of inconsistency, and provided linkage was correct, the third source of data, the hospital patient records, were considered to give the correct information.

In spite of establishing the patient record on the basis of several patient reports from the two registers, complete records for all cases were not obtained. To complete as far as possible the patient records with respect to basic items of personal data, all incomplete records were listed, and the hospital patient records were used for completion.

The computer was used for the completion of the patient records with respect to information on relatives with leprosy (Table 2). It appeared that patient B was often referred to as a relative in the record of patient A, while information in B's record on A as a relative was lacking. A computer programme was used to establish mutual cross references.

Search for doublet records

Due to lack of complete references within and between the two registers, more than one record per patient might be established in some cases. The deletion of such doublet records was postponed until punching was finished and the computer file was established. Accordingly, the computer could be used in search for doublet records.

A search was first performed using the items: sex, residential district, registration year and name as criteria. The patient records were listed according to the items. This search ensured that patient records resulting from more than one report on a patient from one particular district were listed not far from each other. By looking through the lists, doublet records of this kind were easily detected.

To detect doublet records for a patient based on reports from different districts, another search was carried out on the basis of sex and Christian name[1] only.

All doublet records detected were deleted, and the data were transferred to the appropriate patient records where space for data from additional reports was provided (Table 2).

[1] In Norway in the past, the Christian name was more firmly attached to a person than the surname, which usually was the name of the farm.

Table 3. Total number of cases and reports registered in the district register and the hospital register, by number of reports per case. (The National Leprosy Registry of Norway.)

Number of reports per case	District register		Hospital register	
	Cases	Reports	Cases	Reports
1	7,140	7,140	4,091	4,091
2	751	1,502	628	1,256
3	65	195	114	342
4	11	44	33	132
Subtotal	7,967	8,881	4,866	5,821
0	530	0	3,631	0
Total	8,497	8,881	8,497	5,821

2.2.3. RESULTS

Data from 14,702 patient reports were transferred; 8,881 from the district register and 5,821 from the hospital register (Table 3).

Of 8,881 patient reports in the district register, 7,967 represented the first report on a patient and formed the basis of 7,967 patient records. The additional 914 reports represented new information on already registered patients, forwarded to the central office by the main form. The reports represented 827 second notifications, 76 third notifications and 11 fourth notifications. Data from the additional reports were transferred to the appropriate records where space was reserved.

Of 5,821 patient reports in the hospital register, 4,866 represented the first admission of a patient. The additional 955 reports represented 775 second admissions, 147 third admissions and 33 fourth admissions. Linkage of these data to the appropriate patient records was performed by the computer for 3,524 reports which contained a district serial number. Correct identification numbers for 2,297 reports which lacked a cross reference to a district serial number, had to be found manually. For 530 hospital patients no patient report could be found in the district register. Apparently, these patients had been sent directly to the hospital without being notified in a main form, and remained in hospital until they died. These reports in the hospital register formed the basis for another 530 patient records.

Thus, a total of 8,497 patient records was established. Of these, 266 records contained information that invalidated the diagnosis. Accordingly, the computer file contains a total of 8,231 patient records (Table 6).

To establish the computer file 1,372·5 h were used for punching and control punching while 287 h were used for programming. Time used for planning and for manual procedures including proof-reading, was not registered.

2.2.4. COMMENTS

During the transfer of data from the original sources to the computer file, it became evident that the practical management of the registry has been well organized and accomplished. This is remarkable, taking into consideration that the registry was established in the middle of the last century in a country where communications were seriously impeded by a rugged landscape. Apparently, this is also the reason why it was possible to establish a complete computer version; the sources available turned out to cover the entire country and the whole period of observation, containing almost complete information on all cases registered.

In the choice of material, the intention, to establish a file on all leprosy cases as reliable, accurate and complete as possible, had to be considered against the practical problems. Almost all data from the two coordinate registers could easily be transferred. However, a complete transfer of the material of the hospital patient records would involve an amount of work far more extensive than for the two registers, in part due to the magnitude of the material, but particularly because of its structure and form. Consequently, the data in the hospital patient records were used only for limited purposes.

Most of the data found in the register books had a form compatible with direct punching. Accordingly, a most time-consuming work was avoided; viz. coding data onto a special form, and thus introducing a new source of error. On the other hand, the punching itself became more time consuming and gave a higher frequency of errors than punching based on special forms. Furthermore, the data had to be proof-read most carefully.

Obtaining an unambiguous system of identification, necessary to bring together data in appropriate patient records, represented the most important, but also the most difficult part of the computerization, and called for the introduction of a new superior identification number. The alternative, to choose one of the district serial numbers as a superior number, was considered inconvenient; in part because no superior number was stated in the original sources, in part because the cross references were far from complete.

Appropriate identification numbers were not allotted to the patient reports before punching. Accordingly, lacking cross references to previous reports would inevitably produce doublet records. The solution of this important problem was postponed, however, with the assumption that the computer might assist in the search for doublet records, which proved correct. The alternative, manual search for and control of all identification numbers before punching would, if accomplishable, require a vast amount of work.

To *evaluate the computerization* and particularly the method of identification, the total number of patient records was compared with The official statistics of Norway (Melsom, 1948). The official statistics are based on the patient reports in the district register and hospital register. It is stated that 8,218 cases have been registered from 1856 until 1945, which, in addition to 4

cases later registered, total 8,222 cases. Compared with the 8,231 patient records in the computer file, the difference amounts to 9 cases or 0·1%.

In this comparison of statistics based on the computerized material with those based on conventional processing of the original sources, possible lack of correspondence might either be due to errors in the manual procedure or in the computerization. Obviously, the manual procedure was associated with risks of simple miscalculations. Furthermore, the lack of complete cross references introduced a risk that one patient might be counted more than once. Still the official statistics are considered reliable. It is of particular importance that the clerks at the central office were closer to the practical problems and perhaps knew the patients individually. They most probably had information exceeding that written in the books. Accordingly, for practical purposes the total number of cases conventionally compiled is considered the correct number.

Errors related to the computerization would be of two kinds. On one hand, some part of the material may not have been found, or possibly some patient reports may have been overlooked during punching. However, the reports were entered into books, and books were found covering the entire country throughout the whole period of observation. Transfer of every report was secured through proof-reading. Consequently, this source of error is considered insignificant.

On the other hand, some patients might still be represented in the file by two or more patient records, possibly due to a failure in the system of identification established. Since the importance of the other possible sources of error is considered insignificant, the comparison of the two totals may serve as an evaluation of the system of identification. On this basis the result, a difference of 9 cases or 0·1%, is considered acceptable.

2.3. Evaluation of the material

The files of the registry, established 120 years prior to the present study and mostly for other purposes (Irgens and Bjerkedal, 1973) had to be examined carefully with respect to their usefulness as a basis for epidemiological studies. Qualitative and quantitative aspects had to be taken into account; the evaluation was focused on *what* has been registered on *whom*.

On one hand the intention was to state whether the pathological condition described as leprosy in Norway, truly represented leprosy, as the disease is known today, and if so, whether there was a correspondence between systems of classification applied at that time and today. Furthermore, the content of each patient record was to be evaluated.

On the other hand the intention was to assess to what extent all persons in Norway affected by leprosy and living after 1856, were represented by a patient record in the computer file, and to what extent all patient records in this file truly represented leprosy patients.

2.3.1. MATERIAL USED IN THE EVALUATION

In the abundant literature on leprosy in Norway some attention has been paid to the registration of cases, but only to its basis, the Royal Decree of 1856, and the outcome of it, The Official Statistics. Thus, the files of the central office have never previously been considered a central patient registry for the entire country.

Accordingly, the evaluation was first of all based on the files of the central office where, in addition to the register books, all patient forms were found together with circulars and instructions for the District Health Officers. Furthermore, an extensive correspondence on the patients was filed. The official regulations concerning leprosy control in general and the registration work in particular, were also considered.

The evaluation was also based on relevant scientific literature. Of particular importance was a monograph published by D C Danielssen & C W Boeck (1847); the result of a research programme initiated by the health authorities in the 1830s. The initiative was related to a census of leprosy patients in 1836 which demonstrated a great uncertainty in distinguishing leprosy from other diseases. The monograph soon obtained a high reputation internationally and dominated the small medical profession in Norway.

The tradition established by Danielssen & Boeck was continued by G H Armauer Hansen & C Looft (1895) who published a textbook: '*Leprosy in its Clinical and Pathological Aspects*'.

A comparison with the view on leprosy of today was based on the textbook by R G Cochrane & T F Davey (1964) and the system of classification introduced by D S Ridley & W H Jopling (1966).

Information from the nominative population censuses in 1865 and 1875 was used as an independent source of data to assess the reliability of the information reported from the districts on patients living in their homes by 1865 and 1875 respectively. Furthermore, genealogies of high frequency districts, sometimes stating that a member of a family was affected by leprosy, were used to assess certain aspects of the validity.

2.3.2. PROCEDURE OF EVALUATION

The evaluation was divided into two steps. First, the files of the central office were studied in the original form together with the additional material mentioned, focusing on organizational and administrative aspects of the registry. These aspects, it was assumed, were closely related to the possible usefulness of the material in epidemiological studies. From the same point of view, it was necessary to clarify the criteria for the diagnosis of leprosy applied in registration work. A basis of information had to be obtained, making it possible to decide whether the criteria defined a nosological entity equal to leprosy as it is known today and, if possible, whether there is correspondence between systems

of classification used at that time and today. In addition to the evaluation of the material itself, the first step was performed to decide whether a computerization seemed justified.

Secondly, the computer file made an evaluation possible based on the data registered. This step concentrated on the validity of the registration work, assessing on one hand the extent to which all patients registered truly represented leprosy, and on the other, the extent to which all patients affected by leprosy were registered. Furthermore, the computer file was used to assess completeness and reliability of the data registered.

2.3.3. RESULTS

The foundation and general management of the Leprosy Registry have been related previously (Irgens and Bjerkedal, 1973). Here, special attention will be given to aspects pertaining to epidemiological use of the material.

Organization

By the Royal Decree of 1856 the registration work became the responsibility of the *public health authorities*. At the same time definite instructions were given on how registration should be carried out in the districts and how the information should be forwarded to the central office. However, the Royal Decree gave no instructions on how the work should be organized at the central office. Neither has the literature dealing with official leprosy statistics given any information on central registration work. Accordingly, the existence of a central national patient registry for leprosy was first evident after the examination of the instructions together with forms and register books found in the files of the central office.

For the hospital register, neither an administrative basis of registration nor a system for forwarding information to the central office has been mentioned in the literature. The hospital register appeared to be complete with respect to admissions and discharges from 1861. It is assumed that the register was established the same year, together with the opening of the last hospital built during the anti-leprosy programme of the Government (Irgens, 1973). Still, much information on earlier admissions is kept in the hospital register, apparently based on information from hospital patient records.

Administration of the registration work was entrusted to the *Chief Medical Officer for Leprosy* (Irgens, 1973) who was a leader of the central registration work and a supervisor of the work in the districts.

Forms, instructions and circulars filed at the central office add to the documentation that the management of the registry was well organized, and so do numerous annual reports from the central office, accounting for the registration work in the districts and its aims (Irgens and Bjerkedal, 1973).

Case finding

Compulsory notification of all cases of leprosy was the responsibility of the *District Health Officers*, each in charge of a population of up to 10,000 inhabitants. Case finding involved more than notifying patients who consulted the doctor. The doctor was to give information as accurate as possible on all leprosy patients living in his district.

Therefore, the District Health Officers were assisted by local *Boards of Health*, established in all districts where leprosy was found, and instrumental in the notification of all cases. The annual reports from the central office show that the cooperation with the local Boards of Health was effective.

According to the Royal Decree the doctors were also assisted by the *ministers of the church* who served parishes which were smaller than the districts of the health officers. The ministers had, for a long time, been entrusted to register all births and deaths in their parishes and were, accordingly, suitable assistants in the registration of leprosy cases.

On his travels through the districts, The Chief Medical Officer examined a number of patients himself to verify the diagnosis with respect to disease and classification and to supervise the health officers in the registration work.

Follow up

One of the aims of the registration work, viz. *patient care*, rendered follow-up information necessary. Instructions were given that all major changes in the patients should be reported by the local doctors. Important follow-up information was also forwarded from the leprosy hospitals, and this type of data amounts to a considerable part of all data registered.

The follow-up information formed the basis for a special classification of the cases.

Definite cases were cases in which the diagnosis was certain. These patients were all followed until they died.

Evanescent cases were patients who were cured after some time without relapse. Also the course of these patients has been followed carefully until the final conclusion of cure could be made.

Erroneous cases were patients who, after careful follow-up examinations, proved not to be affected by leprosy. Like evanescent cases these cases were followed until the final statement could be given.

Routines of notification, including methods of case finding and follow up, underwent remarkably few changes from 1856 until 1957 when the post of the Chief Medical Officer was discontinued. During the whole period, the field work was the responsibility of the District Health Officers directed and aided by the Chief Medical Officer.

Diagnostic criteria

Based on comprehensive clinical and pathological studies, Danielssen and Boeck (1847) introduced definite criteria to distinguish leprosy from other diseases. Differential diagnoses were broadly discussed.

The disease was divided into two forms, the *tuberculous* and the *anaesthetic*. It was also stated that type of leprosy might change in a patient over time (p. 214). Furthermore, many cases could not be classified as one of the polar forms. They were considered to represent a continuum between the two polar forms (p. 247). Thus the dynamics found in the classification of today were introduced. The authors had also observed acute relapses, apparently related to cursive *nodosum leprosum*, which were described in detail (p. 152).

Without great changes, the diagnostic principles of Danielssen & Boeck were applied by Hansen & Looft (1895), who introduced the terms *tuberous* and *maculo-anaesthetic* for the two polar forms. Also intermediate cases were described. However, the dichotomy in the classification was stressed. The intermediate manifestations were interpreted as a result of a transformation of a case from one type to another (p. 3). This transformation, argued Hansen and Looft, justified retaining only two classes (p. 3). However, a third *mixed* form was still used to some extent for classification of cases in the Leprosy Registry as well as in clinical work (*vide* 2.5.3.)

The morphological descriptions of the maculo-anaesthetic form, given by Hansen & Looft (1895, p. 55) and other Norwegian leprologists (Looft, 1891; Lie, 1912; Lie, 1923; Lie, 1927) leave no doubt that pure *tuberculoid* cases occurred in Norway, classified as maculo-anaesthetic.

The criteria of diagnosis and classification were applied by a small and homogeneous profession, considering leprosy an important public health problem. Almost all doctors graduated from the medical school at the only university in the country. The students were, until the 1940's, sent to the main leprosy hospital to obtain the best teaching and training arranged by the Chief Medical Officer and based on the national scientific tradition.

On the basis of clinical and pathological criteria, the classification used in Norway was compared with a system of classification of widespread and increasing use today introduced by Ridley & Jopling (1966) (Table 4). It seems to be generally recognized that the essence of the tuberculoid-lepromatous classification is the 'resistance' of the patient to his infection. Accordingly, to classify a case, the resistance of the patient should be assessed in some way. This was difficult in the present material. However, the clinical and pathological description of the diagnostic groups corresponds well with the criteria used for classification in Norway. To the extent that an association exists between resistance on one hand and sign and symptoms on the other, *E. tuberculosa* or *L. tuberosa* would have apparently been classified today as LL or BL, *E. tuberculo-anaesthetica* or *L. mixta* as BL, BB or BT while *E. anaesthetica* or *L. maculo-anaesthetica* would have been classified as BT or TT (Table 4).

Table 4. Equivalent terms used in diagnosis of leprosy by The National Leprosy Registry of Norway and today.

Equivalent terms used		
by The National Leprosy Registry of Norway		today
Danielsson & Boeck (1847)	Hansen & Looft (1895)	Ridley & Jopling (1966)
Elephantiasis (E.)	Lepra (L)	Leprosy
E. tuberculosa	*L. tuberosa*	LL BL
E. tuberculo-anaesthetica	*L. mixta*	BL BB BT
B. anaesthetica	*L. maculo-anaesthetica*	BT TT

Validity

Obviously, no complete independent source of data was available to evaluate the sensitivity of the Leprosy Registry as a diagnostic procedure applied to the population. Such data were necessary to assess *under-registration*, due to overlooked or false negative cases. However, all patients detected by chance in general nominative censuses or in local genealogies and all persons with leprosy inquired after by descendants today, are represented by patient records in the computer file.

On the basis of the computer file the extent of under-registration was assessed indirectly. Possible delays in registration work with long periods between onset of the disease and registration, might cause insufficient case finding and under-registration, e.g. due to death before registration. However, the duration of the period between onset and registration was short; the median was 1·4 years for all patients with year of onset between 1856 and 1970 (Table 5).

Spontaneous cure before registration might represent a similar cause of under-registration. In the computer file 183 patients were reported cured without relapse, but the material gave no indication that all cured cases were notified.

However, the data made an appraisal possible of whether the administration of the registry itself considered the registration work satisfactory. The demonstration by Hansen of the leprosy bacillus in 1873, gave strong support to the view that leprosy was a contagious disease. Accordingly, immediate isolation of the cases was considered an effective control measure and was introduced in the legislation against the disease (Irgens, 1973). The demands on accurate and efficient registration work were strengthened. However, greater efforts were apparently not required; the duration of the period between onset and registration did not decrease after 1873 (Table 5).

The extent of *over-registration*, i.e. diagnoses of leprosy in persons not affected by the disease, could be assessed more thoroughly. A total of 8,497 persons were notified as leprosy patients during the period 1856–1970.

Table 5. New cases of leprosy in Norway with median and semi-interquartile range of the delay period between onset and registration, by year of onset. (The National Leprosy Registry of Norway)

Year of onset	Number of patients	Delay: onset-registration	
		Median (years)	Semi-interquartile range (years)
1856–1860	1,154	1·3	1·1
1861–1870	2,005	1·2	1·1
1871–1880	1,241	1·6	1·5
1881–1890	615	1·6	1·5
1891–1900	276	2·0	1·9
1901–1910	117	1·7	1·7
1911–1920	43	1·6	1·7
1921–1970	14	2·0	1·7
1856–1970	5,465	1·4	1·7
–1855	2,289		
Unknown	477		
Total	8,231		

Supplementary information revealed that 266 persons had never been affected. Provided these persons can be interpreted as representing the false positive cases, produced by the diagnostic procedure applied on the population by the Leprosy Registry, the predictive value of a primary diagnosis can be calculated. Predictive value, expressed as the percentage of correct diagnoses, was for the whole period 96·9%, increasing from 95·9% in 1856 up to 100% between 1911 and 1970 (Table 6).

The demonstration of false diagnoses in 266 persons did not preclude the existence of further false diagnoses among the remaining 8,231 patients. The magnitude of this number was assessed from information in the hospital register. A diagnosis made in hospital was considered more reliable than a diagnosis made by a local doctor. Of all 8,497 persons notified, 4,866 (57·3%) were admitted to hospital once or more, and the diagnosis was confirmed in 4,807 cases, corresponding to a predictive value of 98·8% (Table 7). Among the 4,807 patients with a confirmed diagnosis, the number of false positive cases is considered insignificant. For the patients not admitted to hospital the diagnosis was confirmed in 3,424 of 3,631 cases; i.e. a predictive value of 94·3%. This lower value reflects effort in follow up and detection of false diagnoses. For this reason the remaining number of false diagnoses among patients not admitted to hospital may be considered low.

Completeness

Of the 8,231 records representing patients with leprosy, 7,515 (91·3%) were complete with respect to name, sex, residential district, years of birth, onset

Table 6. New cases by year of registration and correctness of diagnosis with respect to disease. (The National Leprosy Registry of Norway)

Year of registration	Cases primarily diagnosed as leprosy			
	Total Number (1)	Correct diagnosis Number (2)	Incorrect diagnosis Number (3)	Predictive value of diagnosis ((2)/(1) x 100)
1856	2,095	2,009	86	95·9
1857–1860	1,080	1,037	43	96·0
1861–1870	2,246	2,180	66	97·0
1871–1880	1,531	1,489	42	97·2
1881–1890	801	786	15	98·1
1891–1900	422	411	11	97·4
1901–1910	223	220	3	98·7
1911–1920	72	72	0	100·0
1921–1970	27	27	0	100·0
Total	8,497	8,231	266	96·9

Table 7. Cases registered by correctness of diagnosis with respect to disease, according to whether admitted to hospital or not. (The National Leprosy Registry of Norway)

Admission to hospital	Cases primarily diagnosed as leprosy			
	Total Number (1)	Correct diagnosis Number (2)	Incorrect diagnosis Number (3)	Predictive value of diagnosis ((2)/(1) x 100)
Admitted	4,866	4,807	59	98·8
Not admitted	3,631	3,424	207	94·3
Total	8,497	8,231	266	96·9

Table 8. Completeness of data with respect to year of birth, year of onset, register conclusion and type, of all leprosy patients registered. (The National Leprosy Registry of Norway)

Completeness†	Total		Type of the disease			
			Stated		Omitted	
	No.	%	No.	%	No.	%
B + O + C +	7,560	91·9	7,515	91·3	45	0·6
B + O + C −	152	1·8	151	1·8	1	0
B + O − C +	381	4·6	287	3·5	94	1·1
B − O + C +	40	0·5	39	0·5	1	0
B − O − C +	86	1·0	28	0·3	58	0·7
Other combinations	12	0·2	7	0·1	5	0·1
Total	8,231	100·0	8,027	97·5	204	2·5

†B: Year of birth, O: Year of onset, C: Conclusion, +: Stated, −: Omitted

and primary registration, type of leprosy and registry conclusion specified as dead with leprosy, permanent cure, permanent emigration or still alive in 1970 (Table 8). Year of birth is calculated on the basis of information on year of primary registration and age when registered. In 990 (12·0%) of the 7,515 records no exact information on year of birth is held. The exact age of these patients was not reported to the central office; however, an age interval of 10 years covering the exact age was stated. In compiling the data, the middle year of the interval was used as the basis for the tabulations of patients by age.

In the records of the remaining 716 (8·7%) patients, some information is lacking. The item of greatest significance is perhaps registry conclusion, lacking in 164 (2·0%) records. Apparently, many of these patients emigrated to the USA and this follow-up information was not entered into the register books. The official statistics on leprosy report 288 emigrated patients up to 1890 without stating whether the emigration was permanent (Hansen and Looft, 1895, p. 145), while in the computer file only 84 patients are registered as permanently emigrated.

Out of the 164 patients without information on conclusion, 56 were reported to have deserted from hospital, and remarks were made in the books for some patients that they could not be localized in spite of all efforts.

Information on year of onset is lacking in 477 (5·8%) records. Out of these, 243 (50·9%) patients were registered in 1856, the first year of registration. Type of leprosy is lacking in 204 (2·5%) records, 136 (66·6%) of these being registered in 1856. In 134 (1·6%) records year of birth is lacking.

Data gathered on members of the patients' family affected by leprosy were not complete. In many cases more detailed information was found in hospital records, and this information was transferred to the computer file.

Completeness of data varied from time to time (Table 9) being particularly low in 1856 and improving greatly to a high level until 1980 when a temporary tendency towards incomplete data occurred. At this time, year of onset was permanently dropped as an item in the primary registration. For all patients registered after 1905, information on year of onset is derived from hospital patient records.

Reliability

Evaluation of reliability called for comparison between independent sources. The district register and the hospital register were not strictly independent. Nevertheless, the systems of reporting were so separate that a comparison seemed justified. This comparison was performed during the establishment of the patient record, linking together data from several patient reports in the two registers. All discrepancies concerning the patients' basic data, e.g. years of birth, onset and death, found in both registers, were detected by the computer. In only a few cases were real discrepancies documented.

In a sample, a comparison, with information from general population

Table 9. Completeness of data, expressed as number of patients for whom information on an item was lacking, in percent of all patients registered, by year of registration. (The National Leprosy Registry of Norway)

Year of registration	Lack in registration of					
	Year of birth		Year of onset		Type of leprosy	
	No.	%	No.	%	No.	%
1856	47	2·3	243	12·1	136	6·8
1857–1860	11	1·1	34	3·3	33	3·2
1861–1870	11	0·5	44	2·0	24	1·1
1871–1880	7	0·5	34	2·3	2	0·1
1881–1890	29	3·7	34	4·3	4	0·5
1891–1900	25	6·1	39	9·5	4	1·0
1901–1910	4	1·8	31	14·1	1	0·4
1911–1920	0	0·0	13	18·1	0	0·0
1921–1970	0	0·0	5	18·5	0	0·0
Total	134	1·6	477	5·8	204	2·5

Table 10. Patients admitted to hospital the year of registration or the subsequent year, by year of registration and correctness of diagnosis with respect to the type of disease. (The National Leprosy Registry of Norway)

Year of registration	Leprosy patients admitted to hospital the year of registration or the subsequent year			
	Total Number (1)	Correct diagnosis Number (2)	Incorrect diagnosis Number (3)	Predictive value of diagnosis ((2)/(1) x 100)
1856	17	13	4	76·5
1857–1860	26	20	6	76·9
1861–1870	341	308	33	90·3
1871–1880	555	503	52	90·6
1881–1890	274	243	31	88·6
1891–1900	104	96	8	92·2
1901–1910	82	68	14	82·9
1911–1920	32	32	0	100·0
1921–1970	16	15	1	93·7
Total	1,447	1,298	149	89·7

censuses referring to 374 patients, revealed no discrepancies with respect to basic personal data.

Considering the classification in hospital with respect to type of leprosy as correct, the reliability of the classification in the district was assessed. Evidently, type of leprosy in a patient may change with time. However, many patients were admitted to hospital a short time after primary diagnosis and classification were made. It appeared that in this group of patients, apart from those registered in the first five years after 1856, there was a high and stable agreement

between classification in districts and classification in hospitals (Table 10), indicating a high reliability in field work.

2.3.4. COMMENTS

The material made it possible to state that the disease, described and notified as leprosy in Norway, in fact was leprosy as the disease is known today. Well defined criteria were established to distinguish leprosy from other diseases. Furthermore, the classification into types appeared to be closely related to the classification used today. The monographs, even assessed by today's knowledge, represent dissertations of remarkable value (Crawford, 1973), stressing signs and symptoms of great significance in the diagnostics and classification of today. The Chief Medical Officers were closely related to this scientific tradition, even before Hansen, and a scientific basis for the registration work was secured.

Standardization was obtained through the compulsory notification entrusted upon the District Health Officers and regulated by detailed instructions. The doctors had a common professional background and graduated in a period when leprosy was considered an interesting field of research and a challenging public health problem, offering important national implications. The efforts of the Chief Medical Officer, as a central coordinator as well as a practical supervisor in the field, have no doubt contributed to standardization of the registration work. Obviously, the lack of changes in the routines of registration for more than 100 years, had the same effect.

Documentation of a standardized registration was also obtained through the establishment of the computer file. The classification of a case with respect to type of leprosy, performed in districts and hospitals appeared to be most accordant, and so were the other data reported from both districts and hospitals.

The leprosy hospitals were, by the standards of that time, of a high medical quality. The hospitals had their own research centre (Irgens, 1973) and were, through personal contacts, connected with other scientific centres. Accordingly, it was assumed that the data of the hospital registers were most reliable. Since discrepancies between the two registers were detected only in an insignificant number of cases, the reliability of the data in the district register, for patients not admitted to hospital, was also considered satisfactory.

The validity of the registration scheme was apparently greatly influenced by one of the purposes of establishing the registry; to obtain an epidemiological description of the disease. For this reason registration was made compulsory. The subsequent demonstration of leprosy as an infectious disease gave support to the decree. This task was attempted with assistance from the ministers of the church and the local Boards of Health. Presumably, the local Boards of Health played an important part, searching for patients who hid away from the stigmatizing attitudes of the local society. Unfortunate social effects of these activities will not be discussed here.

Moreover, individual follow up of the patients contributed to the validity of the registration scheme. On one hand, false positive cases were detected and deleted in the books. On the other hand, follow up work in the district led to a closer contact between the doctor and the population. The contact was also valuable in the search for new cases, and of great importance to avoid under-registration. In particular, the routines of current follow up made the calculation of incidence rates far more reliable than rates based on mere patient censuses repeated at intervals.

However, data directly related to validity were only available to a minor extent. In particular, the magnitude of under-registration was difficult to assess. Still, the length of time between onset of the disease and registration gave some indication, and a median no longer than 1·4 years was considered the result of effective efforts in case finding.

Apparently, over-registration, giving a low predictive value of the primary diagnosis, only occurred infrequently. However, information on all erroneous cases might not have been forwarded to the central office. If a doctor felt uncertain, he might have observed the patient for some time, and if he proved not to exhibit leprosy, the patient was not notified. In the opposite case the patient might have been notified as an ordinary case after a period of observation. Today, prevailing practice with respect to such observation cannot be ascertained. However, if there was a tendency to observe a case for a considerable time before notification, a median of the period longer than 1·4 years would have been expected (Table 5). Accordingly, the high predictive values are considered true.

The relatively lower predictive value of diagnoses made in the district for patients subsequently not admitted to hospital, also indicated a desire to avoid under-registration. During the period in which under-registration might have represented a particular problem, i.e. the first years after the establishment of the registry, efforts in case-finding, expressed by a low predictive value, appeared to have been considerable (Table 6). This finding supported the assumption that under-registration represented no serious problem.

Besides, the relatively higher predictive value of the primary diagnosis in patients later admitted to hospital, should be related to the fact that particularly malignant cases with conspicuous manifestations were hospitalized. The patients not admitted to hospital had a milder clinical course, often without characteristic signs and symptoms, implying a lower predictive value of the diagnosis. However, this lower value should also be considered the result of accurate control and a desire to remove all false positive cases from the register. Accordingly, the number of undetected false positive cases among the remaining 3,424 patients in this group is considered low.

2.4. Additional sources of data

2.4.1. LEPROSY CASES

The material from the *leprosy census of 1836* (Irgens, 1973) is at hand in the National Archives, and the forms were accessible for the compilation of statistics. The results of the *leprosy census of 1845* have been published together with the results of the general population census the same year (NOS, 1845).

Apparently, accuracy in case finding varied from census to census, and the results obtained are not directly comparable with statistics based on the registry material. However, assuming approximately the same accuracy in different regions within one census, the results were used for comparisons of time trends in different regions.

2.4.2. TOTAL POPULATION

Data on the total population were derived from *general population censuses* conducted in 1835, 1845, 1855, 1865, 1875, 1890, 1900, 1910 and 1920. The data were published by The Official Statistics of Norway (NOS, 1835; NOS, 1845; NOS, 1855, NOS, 1865; NOS, 1875; NOS, 1890; NOS, 1900; NOS, 1910; NOS, 1920). From 1865 and onwards the censuses were nominative.

Detailed information on *emigration* from Norway to the USA was obtained from official statistics (NOS, 1921), and was used in an attempt to associate the rapid fall in incidence with a demographic phenomenon particularly found in the high frequency regions.

2.4.3. ENVIRONMENTAL FACTORS

In the health district with the highest morbidity rates, Naustdal, data from the *census* of 1865, pertaining to the farm section (*vide* 2.5.2) were processed to obtain detailed information on *production* and *housing* (*vide* 2.5.5). Information on *assets* was derived from the tax rolls (Matrikkel, 1890). Moreover, *ecological data* were collected during an excursion to Naustdal, August 1977 (*vide* 2.5.5). Meteorological observations, organized from 1874, provided *climatic* data for the entire country (Mohn, 1921).

2.5. Methods

2.5.1. DEFINITION OF OBSERVATION PERIOD

For 2,289 (27·8%) of the total of 8,231 patients, year of onset was reported as before 1856 (Table 12). Accordingly, a considerable amount of the total material could not be used as a basis for calculation of rates. In an attempt to extend the observation period, the case fatality rate of the patients with year of

onset in the period 1856–60 was studied. Only 5·8% of these patients were dead by the end of 1860. For 1,201 patients, year of onset was reported in the period 1851–55. Provided the case fatality rate in this period was the same, only 74 patients (5·8%) were not registered because they died before 1856. These patients might represent a biased sample of all patients with year of onset between 1851 and 1855. However, the small number of patients omitted justified the incorporation of the interval 1851–55 in the observation period.

The 1,088 patients with year of onset before 1851, obviously constituted a most biased sample of all patients taken ill during the decades prior to the establishment of the registry, representing the patients with the longest survival.

Year of onset between 1921 and 1970 was reported only for 14 patients. Accordingly, an observation period of 70 years from 1851 until 1920 was considered useful for studies of time trends. Information on year of onset was lacking in 477 records. Hence, a total of 6,652 patients were reported as taken ill during the observation period (Table 12).

2.5.2. DEFINITION OF GEOGRAPHICAL AREAS

Mean population of the whole country through the observation period was 1,984,791.

According to the morbidity rates, the country was divided into high and low frequency areas (Fig. 1). The *high frequency areas* consisted of the counties of West and North Norway, and were divided into the *southern region*. (Rogaland, Hordaland, Sogn & Fjordane) with a mean population through the observation period of 380,864, the *middle region* (Møre & Romsdal, Sør Trøndelag, Nord Trøndelag) with a mean population of 328,619, and the *northern region* (Nordland, Troms, Finnmark) with a mean population of 212,036. The *low frequency areas* consisted of the remaining 10 counties with a mean population of 1,058,175.

From 1851 to 1920 the proportions of the total population living in the southern region decreased from 20·3% to 19·0%, in the middle region the proportion decreased from 17·5% to 15·7% while in the northern region the proportion increased from 8·9% to 11·7%. The proportions of the population living in the low frequency areas by 1851 and 1920 were 53·4% and 53·6% respectively. Thus, the proportions of the total population living in the different areas were remarkably stable through the observation period.

Sogn & Fjordane, with a mean population of 87,074, was the *county* in which the highest morbidity rates were registered. The health districts of this county were studied in more detail and Naustdal, with a mean population of 2,609, was the *health district* with the highest rates (Fig. 31).

In a special study of the epidemiological situation in Naustdal, the *farm* (Norwegian: gaard) was considered the smallest geographical unit for calculation or rates. In some analyses, the farms were divided into the basic units for housing and production, the *farm sections* (Norwegian: bruk) (*vide* 2.5.5).

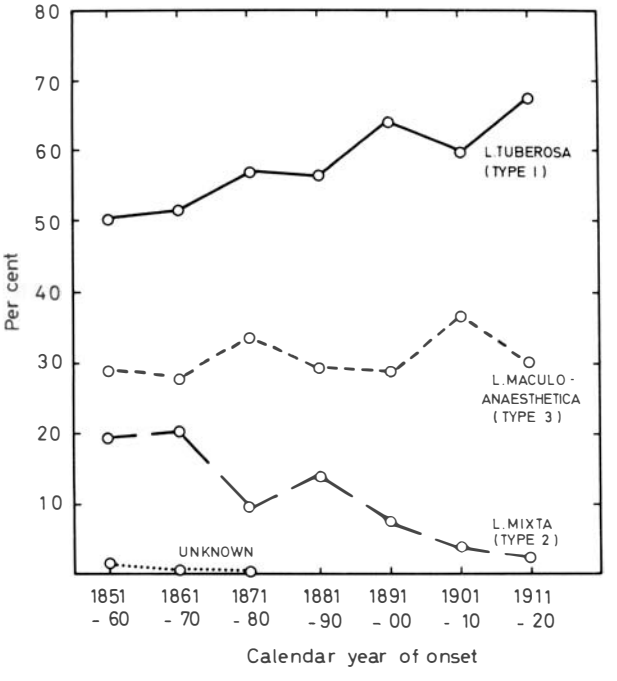


Figure 6. Relative frequencies of types of leprosy in Norway 1851–1920 by year of onset. (The National Leprosy Registry of Norway.)

2.5.3. EPIDEMIOLOGICAL MEASURES

In this study, *prevalence rate* was defined as number of patients at a specified time per 10,000 population. The measure referred to the situation at the end of a year. In some analyses patients in *hospital* were included in the total number of patients for each district, in others the rate referred only to patients present in the district at the specified time. All patients were included from year of onset, or year of registration, until year of death.

Incidence rate was defined as annual number of new patients, according to year of onset, per 100,000 population and was calculated for intervals of 5 or 10 years. Analyses of time trends in different geographical areas throughout the observation period were based on crude rates. In analyses of time trends in the distribution of cases according to age and sex, age- and sex-adjusted incidence rates were applied.

Average incidence rate was an attack rate referring to the complete observation period. The measure was defined as total number of patients whose year of onset was reported in the observation period, multiplied by 100,000 and divided by the average number of inhabitants in the observation period and by the length of the observation period, i.e. 70 years. Average number of inhabitants was calculated as the sum of the mean number of inhabitants for each of the 7 decades in the observation period, divided by 7.

Due to several minor alterations of the borders between the districts, average incidence rates on a district level had to be calculated on the basis of inhabitants present in 1865. At this time approximately one half of the patients had been registered. Since the total number of inhabitants and the composition of the population according to age and sex, varied within narrow limits through the observation period, average incidence rates in these districts were comparable to rates of other areas.

Average incidence rate was used to compare total case load in different geographical areas and to obtain age- and sex-specific incidence rates for the observation period as a total.

In the analysis of the epidemiological situation in the health district of Naustdal, Sogn & Fjordane, an attack rate called *total farm rate* was used. The rate was defined as total number of patients registered at a farm divided by number of inhabitants living at the farm in 1865, according to the general population census, and multiplied by 1,000.

Due to a low number of patients at each farm, it was considered necessary to utilize information on all patients and not only those with year of onset in the observation period, and the measure could not be defined as an annual rate. Accordingly, total farm rates are not comparable to average incidence rates, the latter measure being defined as an annual rate including only patients with year of onset between 1851 and 1920.

Mortality rate was defined as annual number of deaths among the leprosy patients per 100,000 population, calculated for periods of 5 years. Only crude rates were used.

Sex ratio was calculated as the ratio between male and female age-adjusted incidence rates multiplied by 100.

Mean age at onset of patients taken ill in a period was calculated on the basis of the number of cases derived by the application of the age- and sex-specific incidence rates of the period on a standard population of 1885. Mean age of the groups 0–14 years and 50 + years was calculated as 10·7 years and 60·3 years respectively. The number of cases so derived were also used for the calculation of the proportions of patients in different age groups taken ill in a decade.

At farm level, calculation of sex ratio and mean age at onset could not be based on age- and sex-specific incidence rates, because the numbers of inhabitants in the different age and sex groups present at a farm in 1865 were not considered sufficient as a basis for calculation of age- and sex-specific rates. Accordingly, measures had to be based on crude number of patients.

The most malignant type of leprosy, *lepra tuberosa*, is called type 1 in the present study; the intermediate type, *lepra mixta*, is called type 2; and the most benign type, *lepra maculoanaesthetica*, is called type 3. Type of leprosy was registered up to 8 times in a patient record. When nothing else is indicated, type of leprosy in the present study refers to the chronologically last report in which type was stated.

Type index was calculated as number of type 1 patients divided by the sum of type 1 and type 3 patients multiplied by 100. Thus the type 2 patients were proportionally distributed into type 1 and 3. This was necessary to enable comparisons between observations made at different times. Due to instructions from the Chief Medical Officer, classification of a case as type 2 was avoided more and more (Fig. 6) and an index defined as the type-1-proportion of all cases would gradually increase with time, due to a purely administrative decision. Obviously, the type index used in the present study is not suitable as a basis for direct comparisons with a lepromatous index conventionally calculated in other studies, type index tending to be too high. However, in this study, comparisons to other materials of absolute levels on indices rather than trends, were considered precarious and of restricted interest.

2.5.4. STATISTICAL TESTS

Statistical tests were based on computer 'packages' (BMDP, 1975, SPSS, 1975).

In general, χ^2 *tests* were used in comparisons of frequencies and *t-tests* in comparisons of arithmetic means. *Yate's corrected* χ^2 was used for fourfold tables (Hamilton, 1979), except for tables in which one or more expected cell values were 20 or smaller. For these tables *Fisher's exact test* was used.

Simple and multiple regression techniques were used in analyses of possible associations between incidence rates and mean age at onset, sex ratio, and type index, and in analyses of the effects of isolating infectious patients (*vide* 2.5.5.) In simple and multiple analyses, p-values referred to, indicate the probability based on analyses of variance (F-test), that the total variance of the dependent variable is not reduced by the regression function. In simple analyses the p-values may also be interpreted as the probability of non-zero regression coefficients.

Discriminant analysis was used to characterize, by a set of environmental variables, small geographical areas, i.e. farms, respectively with and without leprosy cases.

As a rule, p-values are not given in text or tables; it was considered more important to focus on and discuss trends and gradients in a total view. However, if nothing else is stated, the differences mentioned and discussed are significant, p being less than 0·05.

2.5.5. SPECIAL PROCEDURES

Distribution according to sex, age and type

A modified Lexis' table (Lexis, 1875) was used as a basis for *cohort analyses*, describing the occurrence of the disease in consecutive birth cohorts (Fig. 35). Calculation of age- and sex-specific incidence rates in birth cohorts was based on members of each cohort, i.e. all persons born during the period defined by

the cohort. As numerator, numbers of male and female cohort members for whom onset of leprosy was registered in the specified age group were used. As denominator, estimated numbers of cohort members alive when the cohort reached the specified age groups were used. The rates were calculated per 100,000 population per year. Calculation of mean age at onset in consecutive cohorts was based on age-specific incidence rates and a standard cohort of inhabitants born 1851–60.

Prediction of incidence rates

Construction of indices, which might be used today as a substitute for incidence rates in areas where such rates are lacking or inaccurate, was based on analyses of associations within an area between characteristics of patients registered in a period and incidence rates in the same period.

The county was chosen as the geographical unit. Material for the analyses was derived from the 7 central counties in the high frequency areas, from Hordaland along the coast to Troms (Fig. 1).

Variables for each county were calculated per decade. The first decade in a county after which a decrease in the incidence rate was observed, and all subsequent decades in the county until 1920, were utilized in the analyses. Decades in which number of patients was less than 4, were omitted.

The characteristics of patients to be used in the analyses were *age at onset*, *sex* and *type of leprosy*. To avoid errors due to the long observation period, measures based on age and sex-specific incidence rates were preferred to measures based on crude number of patients (*vide*: 2.5.3). Accordingly, *mean age at onset* and *sex ratio* were used, in addition to *type index*, as independent variables.

Incidence rates, used as dependent variables, were age- and sex-adjusted rates.

The measures, mean age at onset, sex ratio and type index, referring to a specified decade and county, were used as independent variables in simple and stepwise linear regression analyses (BMDP, 1975). The observations were weighted according to number of patients per observation. To predict *level of incidence*, the logarithm of the observed incidence rate, referring to the same period, was used as the dependent variable. To predict *time trend of incidence*, the logarithm of the difference between observed incidence rates in the same period and the subsequent period was used as the dependent variable. This difference represents the slope of the straight line drawn through two subsequent observations of incidence rates.

According to the aim of the study, the practical implications of the findings were illustrated by evaluating the validity of the predictions. Thus, the ability of the predictions was assessed to classify correctly the observations to above or below an optimal cut-off point, which referred to level of incidence and time trend of incidence respectively.

Isolation of patients

The intentions were to analyse associations, in part between prevalence rates in one period and incidence rates in a subsequent period, in part between isolation in hospitals of infectious patients in one period and relative fall in incidence in a subsequent period.

The county was chosen as the geographical unit, and the material for the analyses was derived from the 7 central counties in the high frequency areas (Fig. 1).

Variables for each county were organized in consecutive double sets of periods from 1856 to 1920. The quinquenniad 1851–55 had to be excluded because information on isolation in hospital was lacking. In the first period of each set, period 1, the magnitude of the factor whose effect was to be studied, was measured. In the second period, period 2, the presumed effect of the actual factor was quantified. The lengths of these periods had to be chosen so that as much of the effects as possible in period 2 were due only to factors acting in period 1. In view of the long and varying incubation period of leprosy, this seemed difficult to attain.

However, in the present material the majority of cases were lepromatous, in which mean incubation period appears to cover approximately 10 years (Feldman, 1973). Accordingly, to choose each period to be of 10 years duration, and to let period 2 follow immediately after period 1, involving a mean incubation period of 10 years, seemed justified. Thus 5 double sets for each county were obtained from 1856 to 1915.

The associations between the variables were assessed by simple linear regression analysis (BMDP, 1975). The p-values indicate the probability that the regression coefficient is equal to zero.

In the first analysis the association between *prevalence* as a factor in period 1 and *incidence* as an effect in period 2 was examined. In part, prevalence rates of a county in a period were calculated on the basis of total patient years in the period, of patients registered in the county, spent either in the households or in leprosy hospitals. In part, prevalence rates were calculated on the basis of patient years spent in households only. Mean population of the county in the same period, multiplied by the length of the period in years, was used as the denominator. Age- and sex-adjusted incidence rates were used.

In the next analysis, the association between *degree of isolation* in period 1 as a factor and *relative fall in incidence rates* between period 1 and period 2 as an effect was examined. Degree of isolation in a county was expressed by the number of patient years spent in hospital in percent of total number of patient years spent by patients registered in the county. Relative fall in incidence was calculated by incidence rate in period 1 minus incidence rate in period 2 divided by incidence rate in period 1.

Sets of periods in which prevalence or incidence rates in period 1 were 0, were omitted in the analyses.

Leprosy in families

A *family* was defined as a group of patients known to be mutually related in some way. Families of different *categories* were considered: *all patients* mutually related, *siblings*, one or two *parents with their children* and *spouses*. In the different categories, the families were divided into *groups* according to number of patients, and each group of families was divided into *subgroups* according to numbers of patients in the family with the three different types of leprosy.

A special test was performed to study whether the distribution of cases in the families with respect to type was random or whether special subgroups of families were particularly common. Based on observed frequencies in each group of the three different types, and on the assumption of independence with respect to the type of disease between members of any one family, expected frequencies in each group of all possible subgroups were calculated. Expected number of subgroups were compared with observed number by a χ^2 statistic, and χ^2 values for all groups were added to characterize a category of families.

To study direction of trend in possible differences between observed and expected number of different subgroups, a special measure was constructed: *mean distance* with respect to type of the disease between patients in a family. Distances within all possible pairs of patients in a family were calculated as the difference between the figures indicating the type of the disease: e.g. the distance between a patient with type 1 and a patient with type 3 was 2. In a family with n patients, $\frac{1}{2}n\ (n-1)$ pairs of distances were obtained. Mean distance between patients in a family was calculated as the sum of all distances divided by number of pairs or distances. In a group of families, mean distance between patients was calculated as the sum of mean distances within the families divided by number of families. To calculate mean distance between patients in a category of families, mean distance for each group was weighted according to total number of patients.

This observed distance between patients was compared with an expected mean distance calculated on the basis of expected number of the different subgroups.

A modified procedure was applied when pairs of parents and their children were analysed, calculating a simple mean on the basis of all pairs.

Leprosy in a high frequency district

This part of the study was based on allocation, within a health district, of all patients registered to their appropriate farms. Registration of the patients' surname, which was the name of the residential farm, made such an allocation possible. The farm was the smallest geographical area for which disease rates could be calculated. However, in Norway the farm, to which the name is attached, is not the basic unit of housing and production. Each farm is divided

into sections with identical names, owned by one peasant and his family. Theoretically of equal rank, these sections may differ considerably with respect to production, housing and assets.

The *leprosy status* of the farm, viz. *leprosy positive* or *leprosy negative*, was related to whether leprosy cases were registered at the farm or not. Leprosy status of the farm and *total farm rate* were used as dependent variables in a series of analyses, and were related to independent *farm variables* on production, housing and assets derived from the general population census of 1865.

The variables on *production* were based on amounts of oats and potatoes sown each year, stated in bushels (= toenner, 1 toenne = 139 litres), and were calculated per person living at the farm in 1865. A *production unit* was defined as one bushel of oats or potatoes. In addition a relative estimate of production of milk was based on number of cows. In the whole district a total of 6·3 production units of oats and potatoes were produced per cow. Accordingly, to introduce a production unit comparable with the units for production of oats and potatoes, 1 cow was defined as representing 6·3 production units for production of milk.

Housing conditions were quantified by the variable: number of persons per house.

Information on *assets* was derived from the tax rolls referring to the basis for taxation prepared in the 1870's stating the tax values in oere (taxation of total land properties in Norway was 50 million oere) for each farm (Matrikkel, 1890). The variable: tax value per farm section was used to characterize a farm.

Still, a variable at farm level, e.g. total production of oats per person, might cover a considerable range with some wealthy and some poor farm sections. However, information in the census on production and housing was available at farm section level. Accordingly, to characterize a farm, a *farm index* was used in addition to the other variables. The farm index was defined as the proportion at each farm of farm sections where the value of the actual variable exceeded a fixed limit. *Farm index for production* was related to farm sections where production per person was lower than the 25-percentile, while *farm index for housing* was related to farm sections where number of persons per house was higher than the 75-percentile. The percentiles were based on information on all farm sections in the health district, and were used to obtain a practical relative scale of production and housing.

Tax value for each farm section was not known, and the proportion of farm sections at each farm with a tax value less than the 25-percentile, could not be calculated.

To test the hypothesis that sphagnum vegetation under special conditions may represent a source of mycobacteria in nature, relevant to the occurrence of leprosy in man (Kazda *et al.*, 1979), leprosy status of the farm and total farm rate were compared with a *sphagnum index* calculated for each farm. This index was based on information derived from the map and observations made during an excursion to Naustdal, August 1977. The 7 discriminating variables,

Table 11. Recoded values of the discriminating variables used in the construction of the sphagnum index

	Variables	Possible observations	Recoded values
Var. 1	Sphagnum vegetation in the neighbourhood	Present	2
		Not present	1
Var. 2	Height of vegetation above sea level (metres)	No vegetation	1
		0–174	5
		175–349	4
		350–524	3
		525 +	2
Var. 3	Orientation of vegetation	No vegetation	1
		North	2
		North East, North West	3
		East, West	4
		South East, South West	5
		South	6
Var. 4	Distance between vegetation and farm (metres)	No vegetation	1
		0–499	2
		500–999	3
		1000–1499	4
		1500 +	5
Var. 5	Height of farm above sea level (metres)	0–124	4
		125–249	3
		250–374	2
		375 +	1
Var. 6	Orientation of farm	North	1
		North East, North West	2
		East, West	3
		South East, South West	4
		South	5
Var. 7	Water supply of the farm	Diffusion through sphagnum vegetation	3
		River through sphagnum vegetation	2
		Well without sphagnum vegetation	1
		Others without sphagnum vegetation	1
		Unknown	2

the *sphagnum variables* (Table 11), used in the construction of the index were: (var. 1) whether sphagnum vegetation was present in the surroundings or not (observed in the field), (var. 2) altitude and (var. 3) orientation of possible vegetation (observed in the field and derived from the map), (var. 4) distance from the vegetation to the farm (observed in the field and derived from the map), (var. 5) altitude and (var. 6) orientation of the farm (derived from the map) and (var. 7) water supply of the farm in the last century (observed in the field).

All possible values of each variable were recoded by the use of consecutive whole numbers. High numbers were assigned to values, which, according to the

hypothesis, were most likely associated with good conditions of mycobacterial growth (Kazda, 1979) (e.g. southern orientation and low altitudes), and with a high risk for the inhabitants of contact with the vegetation (e.g. drinking water supplied to the farm by diffusion through vegetation and short distance to the vegetation).

For each discriminating variable, weighting coefficients were determined by two-group linear discriminant analysis (SPSS, 1975) with all leprosy positive farms in one group, and all leprosy negative farms in the other group.

The discriminant function was defined as the *sphagnum index* of a farm and was calculated as the sum of the products of the weighting coefficients and the recoded values of the discriminating variables.

The relative importance of the individual farm- and sphagnum-variables was assessed in stepwise two-group linear discriminant analyses with leprosy positive and negative farms in the two groups, selecting the variable which would maximize Rao's V. (SPSS, 1975).

The discriminating power of all sphagnum variables, of all farm variables and of all environmental variables pooled together, was compared by an assessment of to what extent discriminant functions, based on the variables, would give a correct classification of the farms as leprosy positive or negative. As cut-off points for each of the functions, the values which would provide the most correct classification, were chosen.

The risk that a farm was a leprosy farm, when its values of the different functions exceeded the optimal cut-off point, was compared with the risk that a farm below the cut-off point was a leprosy farm by a ratio of risks:

$$R = \frac{\frac{a}{a+c}}{\frac{b}{b+d}}$$

Here, a and c denote the frequencies of farms with values above the cut-off point, respectively leprosy positive and negative farms, while b and d denote the frequencies of farms with values below the cut-off point, respectively leprosy positive and negative farms.