# **Predictors of Alcoholism in the Lundby Study**

#### I. Material and Methods\*

Olle Hagnell<sup>1</sup>, Jan Lanke<sup>2</sup>, Birgitta Rorsman<sup>1</sup>, and Rolf Öhman<sup>3</sup>

<sup>1</sup>Department of Social and Forensic Psychiatry,

<sup>2</sup>Department of Statistics,

<sup>3</sup>Department of Psychiatry, University of Lund, Sweden

Summary. In this prospective longitudinal study over 15 years (1957 to 1972) the background factors for those who became alcoholics and those who did not were registered before anybody knew what the outcome would be. The population (2,612 inhabitants) lived in 1957 in a delimited area in the South of Sweden, Lundby. In 1957 nearby everyone (98%) was examined by a single psychiatrist, and again in 1972, irrespective of domicile, by two psychiatrists. Among the men who at the outset did not misuse alcohol, 58 became alcoholics. These alcoholics were compared with the non-alcoholics regarding e.g. personality traits, social factors and interactions between factors.

**Key words:** Alcoholism – Epidemiology – Background factors – Prospective longitudinal study – The Lundby study

#### Introduction

It is generally accepted that a multitude of background factors may be operative in the precipitation of alcohol abuse and dependence. As judged from the literature and from drug controlling measures by society such factors may be related to personality, psychopathology, social environment, and patterns of exposure to alcohol. In retrospect, it is surprising how much has been invested by society in medical facilities, social support, and alcohol legislation for the control of alcohol abuse, in particular when one realizes the scarcity of scientific data guiding society in its measures.

The prevalence of alcoholism in Sweden, as in many other Western societies, is for men over 20 years old in the order of 2%-5% for chronic alcoholism, and 7%-11% for alcohol abuse (Hagnell and Tunving 1972; Öjesjö 1983). According to Fox (1967) 1 out of 13 American men is an alcoholic. The current life time probability for a newborn Swedish male, in accordance with present clinical definition of alcoholism, is 20% for alcohol abuse and 8.6% for dependence (Öjesjö et al. 1982). The prospects of these alcoholics are also miserable, as recently illustrated by Öjesjö (1981).

Offprint requests to: O. Hagnell, Department of Social and Forensic Psychiatry, S:t Lars Sjukhus (RPK), Box 638, S-22006 Lund, Sweden \* This study has been supported by the following grants: The Swedish Council for Planning and Coordination of Research, No. 81/1189; The County Council of Malmöhus, 1981

Alcoholism can be seen with reference to medical, social and psychological conditions. It is well known among experienced physicians and psychiatrists that persons with certain personality patterns and psychopathological traits become more and more alike as the alcoholism progresses. However, very little has been written about which personality traits the alcoholic might have had before he started his abuse. Moreover these descriptions have been within a psychoanalytical frame of reference.

The present study was based on all persons registered in the Lundby area on 1 July, 1957. The incidence of alcoholism among men in the Lundby community cohort, 1957–1972, illustrates the development of alcoholism in the population with a 15-year perspective; furthermore we have information about several background factors related to personality, mental health, social factors and exposure to alcohol. In the present paper we describe the population, our ways of collecting the data, our definition of alcoholism and the statistical procedure utilized in the investigation; the results will be communicated in a following sequence of papers (Hagnell et al. 1986a, b).

## **Population**

The investigation was performed within the frame of the Lundby study. The population was derived from a geographically delimited part of Southern Sweden, nationally representative concerning social structure and demography and contained 2,612 inhabitants on the cross section day of 1 July, 1957. Almost every person (98%) in this cohort was visited and examined by one single psychiatrist (Hagnell) in 1957; in 1962 a postal investigation was performed, and in 1972 the population was revisited and personally examined by one of two psychiatrists (Hagnell and Öjesjö) to form a prospective, longitudinal study focused on mental health in the population as related to background factors of personality and psychosocial environment (Hagnell 1966; Hagnell and Öjesjö 1975; Hagnell et al. 1985c). During the 15-year interval, 706 persons had moved out of the area but almost all probands were visited and examined irrespective of domicile. In only 0.6% of the population was information not reliable enough for a psychiatric evaluation.

The Lundby project was initiated in 1947 by Essen-Möller et al. (1956). About 1,600 of the inhabitants appearing in this study had been examined by Essen-Möller et al. in 1947. In

Table 1. Survey of male inhabitants in Lundby on July 1, 1957

Age 1957	1947 cohort	Newcomers in 1957		
0- 9	0			
10-19	147	69		
20-29	105	65		
30-39	86	84		
40-49	138	55 43 11		
50-59	158			
60-69	88			
70+	102	15		
0+	824	511		
	1,335			

the 1957 study Hagnell revisited and examined these persons and also examined the 1,013 persons who were newcomers to the area, and who had either moved there or been born there. The 1957 cohort consisted of the persons living in the Lundby area on 1 July, 1957, describing essentially personality, health and social adaptation. For the age distribution of the 1957 male cohort see Table 1.

#### **Sources of Information**

Information was collected through extensive observation and interviews by experienced psychiatrists (Hagnell 1966). Additional information was obtained through relatives or other key informants, and from official registers as listed below:

Parish and central population registration
The National Central Bureau of Statistics
Social Insurance Office
Inland Revenue Office
National Police Board (Criminal register)
County Temperance Boards
Hospital case notes, psychiatric
Hospital case notes, non-psychiatric
Official death certificates
Autopsy reports
Regional archives
Key informants
Postal investigation in 1962

On the basis of the information collected an evaluation was performed by psychiatrists concerning mental health.

### **Principles for Scoring Background Factors**

When doing epidemiological research of total populations one of the difficulties is to track down and get sufficient information about all the probands, especially when it comes to collecting information about mental illness since this is a subject which people are still reluctant to talk about.

In most epidemiological investigations, in particular incidence studies, it has been very hard, even impossible sometimes, to determine whether the persons were healthy or ill. The investigator had to be content with a conception close to illness like e.g. 'admitted to hospital', 'consulted a doctor about a certain disease', or when alcoholism was concerned 'having a Temperance Board's registration'.

Also, the investigators reached only persons admitted to hospital or known to the authorities, and for these had to diagnose any disease by means of documents from doctors or institutions. Persons who had been ill, but who had not been admitted to a hospital, or who were not known to the authorities, were consequently not included in the investigation. The drop-out from such a study then becomes large, sometimes so large that the purpose of the study may be questioned. However, most epidemiological investigations have been performed accordingly and still are, mostly because of all the hard and time-consuming work that is connected to an investigation of a total population and because of practical obstacles such as economy.

In the Lundby study we have also used 'admitted to a mental hospital' and 'consulted a psychiatrist' as a kind of diagnosis of mental illness alongside our own diagnostic system used at the evaluation.

In Sweden up to the mid-1970s we had central as well as local Temperance Boards. They were links in an attempt to prevent and control the disease of alcoholism. In each community there was a Temperance Board which consisted of local representatives who knew the population well. The local Temperance Board contacted persons who obviously abused alcohol and tried to help. The abusers were also registered. These registers contained advanced alcoholics as well as persons who had perhaps been registered only once for drunken driving or for disorderly conduct in connection with alcohol consumption. Sometimes it could almost be too easy to become registered with the Temperance Boards: a drinking offence, or a derangement of behaviour such as being loudvoiced, or shouting and rowing as a consequence of alcohol consumption, could cause registration with the local Temperance Board. This applied particularly to young people and to people from lower social class.

In an earlier study of the Lundby population (Hagnell and Sandahl 1972) we showed that one-third of the Lundby probands in spite of being registered with the Temperance Boards were not diagnosed as alcoholics according to our criteria. We also found that one-third of the probands, whom we diagnosed as alcoholics, could not be found in the registers of the Temperance Boards. This third was thus not known by the authorities and was regarded as "hidden alcoholics".

Hence in the present study information on psychiatric treatment such as admittance to a hospital and/or intervention by local Temperance Boards was included as background factors. From an epidemiological point of view this is a conventional way of expressing intensity and character of symptomatology related to mental health. As 'registered by Temperance Boards' was a somewhat controversial background factor that could also be regarded as a kind of diagnosis of alcoholism, calculations were performed both with and without this factor.

# The Definition of Alcoholism

The definition of alcoholism in this study was the same as used in earlier Lundby studies, see Hagnell and Tunving (1972).

Alcohol addiction and chronic alcoholism are together called dependence (DSM-III 303.0). Alcoholism is in this study defined as alcohol dependence plus alcohol abuse (alcohol abuse DSM-III 305.0).

**Table 2.** The occurrence of alcoholism in the Lundby male 1957 cohort in 1957 and 1972<sup>a</sup>

Age 1957	Alcoh in 195		Total male cohort	New alcohe 1957–		Non- alcoholics in 1957
0- 4	0	1	87	2	1	87
5- 9	0	/	82	4	1	82
10-14	0	/	130	8	1	130
15-19	0	/	86	9	1	86
20-24	5	/	85	6	1	80
25-29	8	/	85	6	1	77
30-34	7	/	83	3	1	76
35-39	12	/	87	6	/	75
40-44	15	1	92	7	/	77
45-49	10	1	101	4	1	91
50-54	12	/	125	1	1	113
55-59	11	1	76	0	1	65
60-64	9	/	51	1	1	42
65-69	9	/	48	1	1	39
70–74	4	1	43	0	1	39
75–79	4	1	27	0	1	23
80-84	1	/	34	0	/	33
85-89	0	/	10	0	/	10
90–94	0	1	3	0	/	3
0+	107	1	1335	58	/	1228

<sup>&</sup>lt;sup>a</sup> Those who died during the 15-year period are considered to have in 1972 the status they had at death

#### Registration of Alcoholism

The part of the 1957 cohort of the Lundby study used for this investigation included 1,036 males over 15 years of age registered in the community of Lundby on 1 July 1957 (Table 2). The men younger than 15 years of age were excluded because of incomplete personality information. Females were excluded because there were too few women among the alcoholics. All individuals who due to direct or relevant indirect information were diagnosed as alcoholics on 1 July, 1957, were excluded from the study (n = 107). For each of the remaining 929 individuals the files of the 1972 investigation were searched for information indicating a diagnosis of alcoholism. The development of alcoholism among men over a 15-year period could thus be determined (4.7 per hundred), since 44 men met our definition of alcoholism. Among the 299 men younger than 15 years on 1 July 1957, 14 were alcoholics, and will be subjected to a separate study. Beyond this, some cases may possibly have developed but recovered unnotified at the 1972 interview or in available records covering the period. However, the number was probably very small and the individuals concerned would almost surely have been classified as "alcohol abusers".

## **Statistical Procedure**

In the present investigation, the goal was to "explain" one "outcome variable" (viz. the incidence of alcoholism) by means of a number of "explanatory variables" (viz. the so-called background factors). The usual statistical technique for explaining one variable by means of others is linear regression: one tries to express the outcome y as the sum of a linear

function of the background variables and an "error term" with mathematical expectation 0, i.e. one postulates the model

$$E(y) = c_0 + c_1 x_1 + \dots + c_k x_k$$

However, when, as in our case, the outcome is binary, i.e. takes two values only (conventionally coded as 0 and 1), this technique does not work very well since for a binary variable y it always holds that

$$E(y) = P(y = 1)$$

and hence the above model would amount to

$$P(y = 1) = c_0 + c_1 x_1 + \cdots + c_k x_k;$$

but here the left hand side, being a probability, is confined to the interval from 0 to 1 while the right hand side, being a linear function, can take arbitrary values. One well-established way out of this difficulty is to work with log-odds instead of probabilities. The odds O(A) of an event A is defined as

$$O(A) = \frac{P(A)}{1 - P(A)}$$

i.e. the ratio of the probability that A occurs to the probability that A does not occur. The log-odds of A is simply the logarithm of the odds (usually the natural logarithm, although that is not essential). Since P(A) varies between 0 and 1 the odds O(A) will vary between 0 and infinity and hence the log-odds will vary over the whole real axis. Thus it seems reasonable to try to express the log-odds as a linear function of the background variables:

$$\log O(y = 1) = c_0 + c_1 x_1 + \dots + c_k x_k$$

Transforming this expression back to probabilities one gets the equivalent (but seemingly more complicated) relation

$$P(y=1) = \frac{e^{c_0 + c_1 x_1 + \dots + c_k x_k}}{1 + e^{c_0 + c_1 x_1 + \dots + c_k x_k}}$$

This technique, usually called linear logistic regression, was our main statistical tool in the present investigation. We used the version given as program PLR in the BMDP program package.

After this general description of logistic regression, let us describe in more detail the course followed in the present investigation, using the personality part of the investigation (Hagnell et al. 1986a) as an example, the other parts being in principle managed in the same way. In our cohort, personality and mental health were originally described by over 400 variables. Among these we made a selection, partly based on the frequency of the condition described by the variable (very low frequency meaning exclusion), partly based on knowledge of the subject matter. Each of the remaining 49 variables was then cross-tabulated against the outcome variable 'alcoholism'. These tabulations showed that a number of variables were without importance for the prognosis, at least when considered alone, i.e. not in conjunction with other variables. Such seemingly irrelevant variables were excluded from further analysis although we are aware that some of them might have a prognostic value when considered together with other variables in a way that allows interaction between them.

Thereafter each variable still under consideration was transformed into a binary variable (if not already so) or, in some cases, into two such variables. The latter case occurred with the Sjöbring variables where each Sjöbring dimension (such as validity), originally described by a variable taking five values (coded as 3, 4, 5, 6, 7) was split up into two binary variables: one called subvalidity (corresponding to the values 3 and 4 in the original description) and one called supervalidity (corresponding to the original values 6 and 7). After this elimination of variables and splitting up of some of the remaining ones, we were left with 27 binary explanatory variables.

For each of these 27 variables its odds ratio (also called cross-product)

$$w = \frac{ad}{bc}$$

was computed; here a, b, c and d were the frequencies in that two-by-two table which described the simultaneous distribution of the variable in question and the outcome variable:

$$\begin{array}{c|c} \text{Var} = 0 & 1 \\ \text{Not alcoholic} & \begin{array}{c|c} a & b \\ c & d \end{array}$$

Such an odds ratio can be thought of in the following way. Suppose we have two probands,  $I_0$  and  $I_1$ , where  $I_0$  is known to have 0 as value of the prognostic variable in question, while  $I_1$  is known to have the value 1. If we agree to measure "risk of becoming an alcoholic" by means of the odds

$$O(alcoholic) = \frac{P(alcoholic)}{1 - P(alcoholic)}$$

rather than by the more customary measure P(alcoholic), then the odds ratio describes how many times greater the risk of turning alcoholic is for  $I_1$  than it is for  $I_0$ .

What has been described so far can be said to constitute 27 different attempts at predicting alcoholism by means of 1 single prognostic variable at a time. The next step was to construct a model for predicting alcoholism by means of one single model encompassing all the 27 variables. This was done by means of logistic regression, as described above. Thereby we obtained coefficients  $c_1, c_2, \ldots c_{27}$ ; these were immediately transformed into  $w_1, w_2, \ldots, w_{27}$  where

$$w_{i_1}=e^{c_i}$$

the reason being that these latter numbers are easier to understand intuitively. For suppose we have, as before, two probands  $I_0$  and  $I_1$  where  $I_0$  is known to have 0 and  $I_1$  to have 1 as value of the prognostic variable number i and suppose furthermore (and this is what is new in the present set-up) that, for each of the remaining variables,  $I_0$  and  $I_1$  are known to have the same value. Then  $w_i$  describes how many times greater the risk of becoming alcoholic is for  $I_1$  than it is for  $I_0$ . In other words:  $w_i$  describes the influence of variable number i when all the other variables are kept constant. Also these numbers  $w_1$ ,  $w_2$ , ...,  $w_{27}$  are called odds ratios, although they are so in a somewhat more sophisticated sense that the previously introduced cross-products.

In both situations described above, i.e. in the single-variable models and in the simultaneous model, we computed standard errors for the odds ratios. These showed that quite a number of the odds ratio were not statistically significant (i.e. not significantly different from 1). Thus we were faced with the problem of picking out a set of explanatory variables which would give as good a description of the outcome as the whole set, at the same time containing no unnecessary explan-

atory variables. There are two different techniques available for doing that: the forward technique and the backward one.

The forward selection technique starts by finding that explanatory variable, say  $x_{i_1}$ , which best predicts y without the help of the other variables. Then it finds that  $x_{i_2}$  which, together with the  $x_{i_1}$  already selected, best explains y. In this way more variables  $x_{i_3}, x_{i_4}, \ldots$  are entered into the model, one by one, until no significant improvement is achieved. The backward elimination technique goes to other way round: it starts with a model containing all the explanatory variables, then finds that variable whose omission will cause the smallest loss in explanatory power, then that among the remaining ones whose omission will cause the smallest loss, and so on.

In our case we applied both techniques and got the same answer: a model containing 5 of the 27 variables (see Table 2a, Hagnell et al. 1986a), the same 5 with the forward procedure as with the backward one.

Of the 27 variables 1 was 'Temperance Board registration', and that turned out to be 1 of the 5 variables selected — in fact the most influential of them all. For reasons explained in Hagnell et al. (1986a), we repeated the whole process when that variable was omitted, i.e. with only 26 permitted explanatory variables. Also in this case the forward and backward techniques gave the same answer: a certain set of 4 explanatory variables (see Table 2b, Hagnell et al. 1986a) was singled out, although not exactly that set which, together with the Temperance Board variable, constituted the 5-variable model previously obtained.

This latter phenomenon may call for an explanation; let us give it in terms of one outcome variable y and three possible explanatory variables  $x_1$ ,  $x_2$ ,  $x_3$ . Assume that, when considered on their own,  $x_1$  is the variable that best explains y, and that  $x_2$  is the second best. Then the forward selection technique will start by picking out  $x_1$ ; in the second step it does not necessarily prefer  $x_2$  since the variable then selected is that which, together with  $x_1$ , works best, and that may well be  $x_3$ , viz. if  $x_2$  is what could be called a diluted version of  $x_1$ . Thus it is conceivable that the model  $(x_1, x_3)$  will be selected if  $x_1$  is permitted while, if  $x_1$  is not permitted, the one-variable model  $(x_2)$  will be preferred to  $(x_3)$ .

Now we come to the concept of interaction. This is what occurs if the influence of one variable, say  $x_2$ , depends on the value of another, say  $x_1$ . Then the log-odds will depend on  $x_1$  and  $x_2$  through a function of the type

$$c_1^{(0)}x_1 + c_2^{(0)}x_2 + c_{12}x_1x_2.$$

Inclusion of interactions thus simply means that pairwise products of explanatory variables are also considered as explanatory variables. (Strictly speaking, this is just two-variable interactions. Higher interactions, which correspond to products of more than two variables, can also be considered; they are, however, almost impossible to grasp intuitively and we have preferred not to take them into account.)

In a set of 27 variables one can form  $27 \cdot 26/2 = 351$  pairwise products. Clearly it is impossible to investigate a model with 27 + 351 = 378 explanatory variables; therefore we had to single out a psychiatrically meaningful subset of the total set of 351 interactions. This was done by defining groups of the 27 variables such that only interactions within groups were reasonable to consider. Three such groups were defined from clinical experience, case studies and ethanol's effect on the CNS especially via the catecholamine-GABA systems. Quite empirically we made a model with three hypothetical groups

of people: (1) hyperactive, the executive type (sedation), (2) impaired, low vitality (activation), (3) depressive states, low mood plus anxiety (anaesthesia). The three groups were not disjoint, but that is of no importance; their sizes were 8, 5 and 8, respectively and hence we had only  $8 \cdot 7/2 + 5 \cdot 4/2 + 8 \cdot 7/2 = 66$  possible interactions to consider.

In order to find an appropriate model including relevant interactions the following strategy was adopted. First we started a forward selection search among the 66 possible interactions; when 4 of these had been included, no significant further improvement was obtained. Then we considered the model consisting of the 5 variables in Table 2a (Hagnell et al. 1986a) together with the 4 interactions just found; on that set of 9 variables we performed backward elimination and ended up with the 5-variable model given in Table 3a (Hagnell et al. 1986a). In the same way we added the 4 interactions to the model in Table 2b (Hagnell et al. 1986a), applied backward elimination, and obtained the 6-variable model in Table 3b (Hagnell et al. 1986a).

When one has constructed a linear logistic regression model it is natural to ask oneself how good the model is. A simple way of answering that question is to consider the model's skill at prediction. (A note for the statistically-oriented reader: perhaps 'classification' would have been a better word than 'prediction', but since the two are performed in the same way there is no harm in conforming to the usage in the PLR program in BMDP.) What the model is primarily intended to describe is the probability P(y = 1) as a function of  $x_1, x_2, \ldots, x_k$ ; let  $P^*(y = 1)$  denote the value that the model attributes to the probability of having y = 1. Such a description can in a natural way be transformed into a rule for predicting the value of y; using  $\hat{y}$  as notation for the predicted value, one simply takes

$$\hat{y} = \begin{cases} 1 & \text{if } P^*(y=1) > K \\ 0 & \text{else} \end{cases}$$

where K is a cut-off point presently to be discussed.

Such a prediction role will inevitably misclassify a number of individuals. Thereby one should distinguish between two types of misclassification: some individuals with y = 0 can falsely be given the predictor  $\hat{y} = 1$  and some with y = 1 can be given  $\hat{y} = 0$ . The outcome of the prediction can, when the true results are known, be described by a two-by-two table

$$\begin{array}{c|cccc}
y & \hat{y} & 0 & 1 \\
0 & a & b \\
1 & c & d
\end{array}$$

where a + d is the number of correctly predicted outcomes and where b and c are the numbers of misclassifications of the two types. What this table looks like clearly depends on K: an increase in K brings about an increase in a, an equally large decrease in b, an increase in c and an equally large decrease in d. Borrowing from the terminology of diagnostic tests one can say that the above two-by-two table shows the prediction rule to have a specificity of a/(a+b) and a sensitivity of d/(c+d). Clearly one would like both these numbers to be large; however, there is a trade-off between them: a change in K will cause the sensitivity and the specificity to move in directions opposite to each other. As a measure of the predictive power of a logistic regression model we decided to use the specificity obtained when K is so chosen that the sensitivity has a prescribed value (50%, 60%, 75% or 90%); these specificities will be denoted  $Sp_{50}$ , and so on. In most cases, K cannot be chosen to make the sensitivity exactly equal to the prescribed value. Then interpolation will be necessary; this corresponds to a randomized prediction rule.

The intuitive meaning of  $Sp_{50}$  can be formulated as follows. If we use the model in question for predicting alcoholism and if we calibrate the predictions in such a way that 50% of the alcoholics-to-be are correctly identified, then  $Sp_{50}$  denotes the percentage of the non-alcoholics-to-be that are correctly identified.

In our application there is one point which requires some caution: we have 27 possible explanatory variables, and that is a fairly large number; hence a model singled out by one technique or another is the best (in some sense) among a large number of possible models. Its apparent predictive power should then be considered with some scepticism, the reason being that even if there is no connection whatever between the explanatory variables and the outcome, there is a fair chance that what looks like a relation will be found if one investigates a considerable number of models. In order to assess the size of this effect a special study was undertaken (Lanke 1986). The results indicated that in a data set of the size we have, chance is likely to produce values like  $Sp_{50} = 70\%$ . Since our results (Table 4, Hagnell et al. 1986a) surpassed this level with wide margins, the existence of a true predictive power in our data set was established.

#### References

Essen-Möller E, Larsson H, Uddenberg CE, White G (1956) Individual traits and morbidity in a Swedish rural population. Acta Psychiatry Scand [Suppl] 100

Fox R (1967) Multidisciplinary approach to the treatment of alcoholism. Am J Psychiatry 123:769

Hagnell O (1966) A prospective study of the incidence of mental disorder. Scand Univ Books, Svenska Bokförlaget/Norstedts-Bonniers, Stockholm

Hagnell O, Sandahl A (1972) Secret alcoholism in the hospital. Alkoholfrågan 66:43–49

Hagnell O, Tunving K (1972) Prevalence and nature of alcoholism in a total population. Soc Psychiatry 7:190-201

Hagnell O, Öjesjö L (1975) A prospective study concerning mental disorders of a total population investigated in 1947, 1957, and 1972. The Lundby Study III (preliminary report). Acta Psychiatr Scand [Suppl] 263:1-11

Hagnell O, Lanke J, Rorsman B, Öhman R (1986a) Predictors of alcoholism in the Lundby Study (II). Personality traits as risk factors for alcoholism. Eur Arch Psychiatr Neurol Sci 235: 192–196

Hagnell O, Isberg P-E, Lanke J, Rorsman B, Öhman R (1986b) Predictors of alcoholism in the Lundby Study (III). Social risk factors for alcoholism. Eur Arch Psychiatr Neurol Sci 235:197–199

Hagnell O, Lanke J, Rorsman B, Öjesjö L (1986c) A prospective longitudinal study of the incidence of mental disorders in a total population investigated in 1947, 1957, and 1972. The Lundby Study (to be published)

Lanke J (1986) On nominal and true predictive power for logistic regression models with many explanatory variables. Techn Report, Dept of Statistics, Univ of Lund

Öjesjö L (1981) Long-term outcome in alcohol abuse and alcoholism among males in the Lundby general population, Sweden. Br J Addict 76:391–400

Öjesjö L (1983) An epidemiological investigation of alcoholism in a total population; the Lundby Study. Lund (Thesis)

Öjesjö L, Hagnell O, Lanke J (1982) Incidence of alcoholism among men in the Lundby community cohort, Sweden, 1957–1972. J Stud Alcohol 43:1190–1198