U. S. NATIONAL CENTER For HEALTH STATISTICS | Series 2 Number 5

VITAL and HEALTH STATISTICS

DATA EVALUATION AND METHODS RESEARCH

An Index of Health: Mathematical Models

Mathematical models of the distribution of illness episodes, of illness duration, and of mortality in a population. An index of health reflecting both mortality and morbidity is proposed.

Washington, D.C. May 1965

U.S. DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE
Anthony J. Celebrezze
Secretary

Public Health Service
Luther L. Terry
Surgeon General

NATIONAL CENTER FOR HEALTH STATISTICS

FORREST E. LINDER, PH. D., *Director*

THEODORE D. WOOLSEY, *Deputy Director*

O. K. SAGEN, PH. D., *Assistant Director*

WALT R. SIMMONS, M.A., *Statistical Advisor*

ALICE M. WATERHOUSE, M.D., *Medical Advisor*

JAMES E. KELLY, D.D.S., *Dental Advisor*

LOUIS R. STOLCIS, M.A., *Executive Officer*

OFFICE OF HEALTH STATISTICS ANALYSIS

IWAO M. MORIYAMA, PH. D., *Chief*

DIVISION OF VITAL STATISTICS

ROBERT D. GROVE, PH. D., *Chief*

DIVISION OF HEALTH INTERVIEW STATISTICS

PHILIP S. LAWRENCE, SC. D., *Chief*

DIVISION OF HEALTH RECORDS STATISTICS

MONROE G. SIRKEN, PH. D., *Chief*

DIVISION OF HEALTH EXAMINATION STATISTICS

ARTHUR J. MCDOWELL, *Chief*

DIVISION OF DATA PROCESSING

SIDNEY BINDER, *Chief*

Public Health Service Publication No. 1000-Series 2-No. 5

Library of Congress Catalog Card Number 65-60058

PREFACE

This study is presented as a contribution to the methodology of measuring health status. Among the objectives of the National Center for Health Statistics is the development of new techniques in health measurement.

Although it is generally conceded that mortality statistics no longer provide an adequate measure of the health status of a population, no generally accepted method of measuring health in terms of both mortality and morbidity has emerged. Dr. C. L. Chiang of the School of Public Health, University of California, was invited to develop mathematical models which might serve as the basis for a general index which reflects morbidity as well as mortality. The models which have been developed represent one of many possible approaches to the problem. It is hoped that the publication of his work will lead to more intensive investigation of both the conceptual and the mathematical problems involved in constructing such an index.

As pointed out by the author, further testing is needed to determine whether the models presented provide a good description of observed data.

CONTENTS

AN INDEX OF HEALTH: MATHEMATICAL MODELS

C. L. Chiang, Ph.D, *School of Public Health, University of California, Berkeley*

I. INTRODUCTION

The state of health of a nation is one of the most important aspects in the study of a human population; but the lack of quantitative measures to assess health has always been a problem in the field of public health and welfare activities. The purpose of this study is to suggest mathematical models for describing the state of health of a well-defined population over a given period of time, such as a calendar year.

The health of a population is in part a function of such demographic variables as age, sex, and possibly race. People of different ages and sexes have different susceptibilities to diseases, and diseases may act differently upon them. To describe the health status adequately, the population should be divided into subpopulations according to these variables. For convenience of presentation, however, these subpopulations will be assumed homogeneous with respect to all demographic variables except age. From the public viewpoint, a simple and comprehensive index of the current state of health is most desirable. Because of the complexity of the problem, however, a satisfactory approach should begin with detailed investigations of the basic component variables. The state of health is best measured by the frequency and duration of illness, by the severity of illness, and by the number of deaths. These components taken together give a comprehensive picture of health; separately, each describes an aspect of the state of health.

To measure the frequency of illnesses, we need to know the number of illnesses occurring in a calendar year to each individual of a given age group and the distribution of the subpopulation with respect to this variable. A mathematical model will be developed in section II. Although the model is not specifically developed for a particular type of illness, the general line of approach applies equally well for any specific disease. The derived probability distribution characterizes the pattern of proneness and susceptibility of a subpopulation to disease; it also provides an easy means of calculating incidence and prevalence rates. The mean number of illnesses and the corresponding standard deviation will serve to measure the average proneness and its variability for each subpopulation. Furthermore, all these measures can be used for comparing subpopulations or summarized for the entire nation. As a test the suggested model is fitted to actual data from a sickness survey.

The severity of an illness varies with the disease and the individual concerned. It does not lend itself to quantitative measures except as it is related to duration or to termination in death. In section III we present the derivation of a general model for the duration of illness, which again applies either for a particular disease or for all illnesses. Because of lack of data, no attempt is made to find a specific function; however, alternative approaches are described in detail.

Since death must be related to ill health in a population, a study of health is not complete without considering the mortality rate. Mortality is evaluated from the standpoint of health in section IV.

While studies of the component variables give a more detailed picture of the state of health, development of a single measure summarizing the information for the entire population is also essential. Based on the ideas in the preceding sections, a health index is derived in section V. An adjusted index is suggested in section VI.

II. PROBABILITY DISTRIBUTION OF THE NUMBER OF ILLNESSES

Consider the time interval of 1 year $(0,1)$, and for each $t, 0 \le t \le 1$, let the random variable $N(t)$ be the number of illnesses that an individual has during the time interval $(0,t)$, with $N(0)=0$. The purpose of this section is to derive the probability distribution of the random variable $N(t)$,[1]

$$P_n(t) = Pr\{N(t)=n \mid N(0)=0\}. \tag{1}$$

This probability function is an idealization of the proportion of people in the population having n illnesses, for $n=0,1,\ldots$, during the interval $(0,t)$. When an explicit form for the probability function is derived and computed for each value of n, we have a mathematical representation of the state of health of the population in terms of the number of illnesses.

An assumption underlying the probability distribution (1) is that the probability of occurrence of an illness during the infinitesimal interval $(t, t+h)$ equals $\lambda_t h + o(h)$, where λ_t is a function of time t and $o(h)$ is a negligible quantity when h tends to zero. In essence, this means that the probability of an illness occurring within an infinitesimal time interval is a function of time and is independent of the number of previous illnesses. This assumption leads to a system of differential difference equations for $P_n(t)$. Consider two contiguous time intervals, $(0,t)$ and $(t, t+h)$. Exactly n illnesses can occur in the interval $(0, t+h)$ in three mutually exclusive ways: (a) n illnesses will occur in $(0,t)$ and none in $(t,t+h)$ with a probability $P_n(t)[1-\lambda_t h - o(h)]$; (b) $n-1$ illnesses will occur in $(0,t)$ and one in $(t,t+h)$ with a probability $P_{n-1}(t)[\lambda_t h + o(h)]$; and (c) $n-2$ illnesses or less in $(0,t)$ and two or more in $(t, t+h)$, with a probability of $o(h)$. Taking these possibilities together we have the formula:

$$P_n(t+h) = P_n(t)[1-\lambda_t h - o(h)] + P_{n-1}(t)[\lambda_t h + o(h)] + o(h). \tag{2}$$

Transposing $P_n(t)$, dividing through by h, and taking the limit as h tends to zero, yield a system of differential difference equations.

$$\frac{d}{dt} P_0(t) = -\lambda_t P_0(t) \tag{3}$$

$$\frac{d}{dt} P_n(t) = -\lambda_t P_n(t) + \lambda_t P_{n-1}(t), \qquad n=1,2,\ldots$$

The first equation has the solution

$$P_0(t) = e^{-\int_0^t \lambda_\tau d\tau}; \tag{4}$$

the remaining equations are solved successively to give the probabilities

$$P_n(t) = \frac{e^{-\int_0^t \lambda_\tau d\tau}\left[\int_0^t \lambda_\tau d\tau\right]^n}{n!}, \quad n=1,2,\ldots. \tag{5}$$

For a period of 1 year i.e., for $t=1$ the random variable N has the distribution

$$P\{N=n\} = \frac{e^{-\int_0^1 \lambda_\tau d\tau}\left[\int_0^1 \lambda_\tau d\tau\right]^n}{n!}, \quad n=0,1,\ldots \tag{6}$$

Within a period of 1 year, the instantaneous probability $\lambda_t h + o(h)$ of occurring illness need not be dependent upon time t, and λ_t may be

[1]To indicate explicitly that the models are developed for a subpopulation, say age group x, a subscript x should be added in the appropriate places. For the sake of simplicity of presentation, however, such a subscript will not be used in this section or in section III.

assumed to be constant. Under this assumption, we have the ordinary Poisson distribution

$$P\{N=n\}=\frac{e^{-\lambda}\lambda^{n}}{n!}, \qquad n=0,1,2, \ldots \tag{7}$$

The constant λ in formula (7) signifies an individual's susceptibility to diseases and, as such, is a measure of the degree of his health. In fact, λ is the expected number of illnesses occurring to an individual during a period of 1 year. The larger the value of λ, the more illnesses the individual may be expected to have.

The value of λ varies from one individual to another. To describe mathematically the health status of a population, we shall study the probability distribution of λ. The distribution of λ will be denoted by $g(\lambda)\, d\lambda$, the theoretical proportion of people having the specified value λ. Since the sum of the proportions of individuals is unity, the function g satisfies the condition

$$\int g(\lambda)\, d\lambda = 1,$$

where the integral extends over all possible values of λ. The probability distribution of illnesses will be a weighted average of the probability function (7) with the density function $g(\lambda)\, d\lambda$ employed as weights; that is,

$$P\{N=n\}=\int \frac{e^{-\lambda}\lambda^{n}}{n!}\, g(\lambda)\, d\lambda, \qquad n=0,1, \ldots \tag{8}$$

Roughly, formula (8) may be interpreted as follows:

> The expected proportion of individuals who will have n illnesses during the year is equal to the sum of the products of (a) the proportion of individuals having a specific value of λ, and (b) the probability that an individual with the specific value of λ will have n illnesses during the year, where the sum is taken over all possible values of λ.

Choice of function $g(x)\, d\lambda$ is dependent upon the health condition of the particular group of people in question. It appears, however, that the following function may describe the distribution in general:

$$g(\lambda)\, d\lambda = \frac{\beta^{\alpha}\lambda^{\alpha-1} e^{-\beta\lambda}}{\Gamma(\alpha)} \tag{9}$$

where the gamma function $\Gamma(\alpha)$ is defined by

$$\Gamma(\alpha) = \int_0^{\infty} y^{\alpha-1} e^{-y}\, dy \tag{10}$$

The ranges of the constants for which (9) is defined are $\alpha \geq 0$ and $\beta > 0$.

The function $g(\lambda)$ starts at $\lambda = 0$, increases as λ increases at a rate of

$$\frac{d}{d\lambda} g(\lambda) = g(\lambda)\left[(\alpha-1)\lambda^{-1} - \beta\right], \tag{11}$$

and reaches a maximum of

$$g(\lambda) = \frac{(\alpha-1)^{\alpha-1}\beta e^{-(\alpha-1)}}{\Gamma(\alpha)} \tag{12}$$

at

$$\lambda = \frac{\alpha-1}{\beta} \tag{13}$$

After reaching the maximum value, $g(\lambda)$ decreases as λ increases and assumes a value of zero as λ tends to infinity.

The expectation and variance of λ may be directly computed from (9):

$$E(\lambda) = \int_0^{\infty} \lambda \frac{\beta^{\alpha}}{\Gamma(\alpha)} \lambda^{\alpha-1} e^{-\beta\lambda}\, d\lambda = \frac{\alpha}{\beta} \tag{14}$$

and

$$\sigma_{\lambda}^{2} = \int_0^{\infty} \left(\lambda - \frac{\alpha}{\beta}\right)^{2} \frac{\beta^{\alpha}}{\Gamma(\alpha)} \lambda^{\alpha-1} e^{-\beta\lambda}\, d\lambda.$$

$$= \frac{\alpha}{\beta^{2}} \tag{15}$$

Thus the ratio α/β measures the average health of a population, and the reciprocal of α is the relative variance,

$$\frac{1}{\alpha}=\frac{\sigma_\lambda^2}{[E(\lambda)]^2}\quad, \tag{16}$$

which is a measure of variation of health among individuals in the subpopulation relative to the mean health.

Assuming (9) as the function underlying the distribution of the population with respect to health condition, we have from (8) the probability function of the number of illnesses during the year:

$$P\{N=n\}=\int_0^\infty \frac{e^{-\lambda}\lambda^n}{n!}\frac{\beta^\alpha}{\Gamma(\alpha)}\lambda^{\alpha-1}e^{-\beta\lambda}d\lambda \tag{17}$$

$$=\frac{\Gamma(n+\alpha)}{n!\,\Gamma(\alpha)}\beta^\alpha(1+\beta)^{-(n+\alpha)},\quad n=0,1,\ldots$$

This probability is the expected proportion of individuals in the population having n illnesses during the year, taking into account the variability among individuals in the population as described by formula (9). The expected number of illnesses occurring to an individual in the subpopulation is given by

$$E(N)=\frac{\alpha}{\beta}\quad, \tag{18}$$

and the variance by

$$\sigma_N^2=\frac{\alpha(1+\beta)}{\beta^2}\quad. \tag{19}$$

Formula (17) represents a family of infinitely many probability distributions, depending upon the constants α and β. The health status of a subpopulation may best be described as a member of the probability distribution family for which α and β assume particular values. In order to estimate these values, it is necessary to know the observed frequency distribution of the number of illnesses occurring to the individuals of the subpopulation from which the mean $\bar{N}$ and variance S_N^2 of the number of illnesses are computed. Substituting $\bar{N}$ and S_N^2 for $E(N)$ and σ_N^2, respectively, in (18) and (19) and solving the resulting equations for α and β give the estimates

$$\hat{\alpha}=\frac{\bar{N}^2}{(S_N^2-\bar{N})} \tag{20}$$

and

$$\hat{\beta}=\frac{\bar{N}}{(S_N^2-\bar{N})} \tag{21}$$

Using the estimated values $\hat{\alpha}$ and $\hat{\beta}$ in (17), we have

$$P\{N=n\}=\frac{\Gamma(n+\hat{\alpha})}{n!\,\Gamma(\hat{\alpha})}\hat{\beta}^{\hat{\alpha}}(1+\hat{\beta})^{-(n+\hat{\alpha})},$$

$$n=0,1,\ldots \tag{22}$$

The probability (22) multiplied by the total number of individuals in the subpopulation is the expected number of individuals having n illnesses during the year, for $n=0,1,2,\ldots$. This expected frequency distribution may be compared with the observed frequency distribution by means of the chi-square test to determine whether the model described by (17) is an adequate measure of the state of health. Material collected by the Canadian Sickness Survey, 1950-1951, is used for this purpose (see references 3 and 4).

The data in the Survey were based on a sample of approximately 10,000 households[2] inflated to give the national figures as appeared in the publications. Thus the published figures are much greater than the actual counts in the sample. Not knowing the exact number, we take 13,538 as the sample size and each thousand in the published data as a single count (the total population size is 13,538,000, see table 3). Since the actual sample size is probably larger than 10,000 the exact chi-square values in our test should be somewhat greater.

Two indirect measures of illness were used—the number of doctors' calls and clinic visits and the number of complaint periods that an individual had during the year. For the number of doctors' calls and clinic visits the model is

[2]See page 17 of Reference 3.

fitted for six age groups—under 15, 15-24, 25-44, 45-64, 65 and over, and all ages. The results for the first two age groups and for all ages are presented in tables 1, 2, and 3, and figures 1, 2, and 3, respectively. In each of the first two cases, the fit is quite good. For all ages, however, the chi-square value exceeds the critical value at the 1 percent level of significance.

Data on the number of complaint periods were divided into only four age groups—under 15, 15-64, 65 and over, and all ages. Only the age group under 15 is well described by the present model as shown in table 4 and figure 4.

Although neither of the underlying random variables is that of our model, the chi-square tests show promising prospects when the age intervals are not too large. It is hoped that more appropriate material will be made available for further testing.

III. PROBABILITY DISTRIBUTION OF THE DURATION OF ILLNESS

Let random variable T be the duration of an illness so that

$$Pr\,(T \leq t) \tag{23}$$

is the probability that a person will recover from an illness within a period of time t. We are interested in an explicit function of the probability in (23) and of the corresponding probability density function of T.

Consider an infinitesimal time interval $(t, t+\Delta t)$ for $t \geqq 0$. The conditional probability that an individual who is ill for the period of time t will recover from the illness during the interval $(t, t+\Delta t)$ is certainly a function of time t and length Δt. Let this function be denoted by $\psi(t, \Delta t)$. We shall assume that the function ψ is continuous with respect to Δt and has the first derivative, say v_t, at $\Delta t = 0$ for each possible value of t. It follows from Taylor's theorem that

$$\psi(t, \Delta t) = v_t \Delta t + o(\Delta t)\ , \tag{24}$$

where v_t is a function of t and $o(\Delta t)$ is a quantity of a smaller order of magnitude than Δt.

To derive an explicit function for (23), consider

$$Q(t) = 1 - Pr\,(T \leq t) = Pr\,(T > t) \tag{25}$$

the probability that an individual will be ill for a period longer than t, and the time interval $(t, t+\Delta t)$. In order for an individual to be ill for a period longer than $(t+\Delta t)$ he must be ill for a period longer than t *and* not recover within the period $(t, t+\Delta t)$. Accordingly, the corresponding probabilities have the following relationship:

$$Q(t+\Delta t) = Q(t)\ \left[1 - v_t \Delta t - o(\Delta t)\right]\ . \tag{26}$$

Subtracting $Q(t)$ from both sides of (26), dividing (24) through by Δt, and letting Δt tend to zero, we have the differential equation

$$\frac{d}{dt}\ Q(t) = -Q(t)\, v_t\ . \tag{27}$$

The general solution of (27) is easily found to be

$$Q(t) = e^{-\int_0^t v_\tau\, d\tau + c}\ , \tag{28}$$

where c is the constant of integration. Since $Q(0) = Pr(T>0) = 1$ for $t=0$ the constant of integration must be 0, and we have the desired solution

$$Q(t) = e^{-\int_0^t v_\tau\, d\tau}\ , \tag{29}$$

and the probability

$$Pr\,(T \leq t) = 1 - e^{-\int_0^t v_\tau\, d\tau}\ . \tag{30}$$

The derivative of (30) gives the probability density function of the general model of the duration of an illness

$$f(t)\,dt = d\left[1 - e^{-\int_0^t v_\tau\,d\tau}\right]$$

$$= v_t\, e^{-\int_0^t v_\tau\,d\tau}\,dt, \qquad t \geq 0. \tag{31}$$

It should be noted that the random variable T in the above model is the complete duration of an illness rather than the duration of an illness observed within a calendar year as needed in our formulation. In the study of the health of a current population, we are considering a truncated case, where the duration of an illness is defined as the interval extending from the beginning of the year (if it is an illness continuing from the preceding year) or from the date of onset to the date of recovery, death, or the end of the year, whichever comes first. Therefore, we shall consider a truncated distribution of T. For illnesses occurring during the calendar year, truncation could be made at random. But this would result in truncation at a point approximately 6 months after onset and would misrepresent those illnesses having durations of more than 6 months within a calendar year. Random truncation would further distort the picture of the entire distribution which includes chronic illnesses carried over from the preceding year. To avoid distortions, we use one year as the point of truncation.

Let the random variable T^* be the truncated duration of illness expressed in unit of years. The probability distribution, say $f^*(t)\,dt$ of T^*, is equal to the truncated distribution of the original random variable T, for $T \leq 1$; symbolically,

$$f^*(t)\,dt = \frac{f(t)\,dt}{P(T \leq 1)}, \qquad \text{for } 0 \leq t \leq 1,$$

$$= 0, \qquad \textit{elsewhere}. \tag{32}$$

Substituting (30) and (31) in (32) gives the required model

$$f^*(t)\,dt = \frac{v_t\, e^{-\int_0^t v_\tau\,d\tau}}{1 - e^{-\int_0^1 v_\tau\,d\tau}}\,dt. \tag{33}$$

The sum of $f^*(t)\,dt$ over all possible values of t for $0 \leq t \leq 1$ is unity or

$$\int_0^1 \frac{v_t\, e^{-\int_0^t v_\tau\,d\tau}}{1 - e^{-\int_0^1 v_\tau\,d\tau}}\,dt = 1, \tag{34}$$

as can be proven by direct integration. Both the probability distribution (33) and the expectation of T^*,

$$E(T^*) = \int_0^1 t\, f^*(t)\,dt \tag{35}$$

$$= \int_0^1 t\,\frac{v_t\, e^{-\int_0^t v_\tau\,d\tau}}{1 - e^{-\int_0^1 v_\tau\,d\tau}}\,dt$$

$$= \frac{\int_0^1 e^{-\int_0^t v_\tau\,d\tau}\,dt - e^{-\int_0^1 v_\tau\,d\tau}}{1 - e^{-\int_0^1 v_\tau\,d\tau}}$$

are dependent upon the function v_t.

The general model (33) may be applied to a practical situation by specifying the function

v_t. Without appropriate data for testing a specific model, we consider only the following simple case to illustrate a few points.

Supposing that $v_t = v$ is independent of t, we have from (33) a truncated negative exponential distribution

$$f^*(t)\,dt = \frac{v\,e^{-vt}}{1-e^{-v}}\,dt\,, \qquad \text{for } 0 \le t \le 1\,,$$

$$= 0\,, \qquad \textit{elsewhere}. \tag{36}$$

Under this model the function $f^*(t)$ assumes a maximum value of $v(1-e^{-v})^{-1}$ at $t=0$ and decreases as t increases at a rate of

$$\frac{v^2}{1-e^{-v}}\,e^{-vt}\,. \tag{37}$$

The expected value and the second moment of T^* can be obtained from (36).

$$E(T^*) = \int_0^1 t\,\frac{ve^{-vt}}{1-e^{-v}}\,dt$$

$$= v^{-1} - \frac{e^{-v}}{1-e^{-v}} \tag{38}$$

and

$$E(T^{*2}) = \int_0^1 t^2\,\frac{ve^{-vt}}{1-e^{-v}}\,dt \tag{39}$$

$$= 2\,v^{-2} + (2\,v^{-1}+1)\left[1-(1-e^{-v})^{-1}\right].$$

Therefore the variance of T^* is

$$\sigma^2_{T^*} = E(T^{*2}) - \left[E(T^*)\right]^2$$

$$= v^{-2} - \frac{e^{-v}}{(1-e^{-v})^2}\,. \tag{40}$$

In applying the model in formula (33) to actual data, we need to estimate the parameter v. A typical statistical approach is to use the method of maximum likelihood. Consider a total of M illnesses and let t_i^* be the truncated duration of the i-th illness, for $i = 1, 2, \ldots, M$. The joint density function of the M random variables is

$$\Phi = \prod_{i=1}^{M} \frac{ve^{-vt_i^*}}{1-e^{-v}} = \left(\frac{v}{1-e^{-v}}\right)^M e^{-v\sum_{i=1}^{M} t_i^*}\,. \tag{41}$$

Maximizing the likelihood function (41) with respect to v leads to an equation

$$\bar{T}^* = \hat{v}^{-1} - \frac{e^{-\hat{v}}}{1-e^{-\hat{v}}}\,, \tag{42}$$

where

$$\bar{T}^* = \frac{1}{M}\sum_{i=1}^{M} t_i^* \tag{43}$$

is the mean truncated duration of an illness and $\hat{v}$ is the likelihood estimate of v. Formula (42) can be solved for $\hat{v}$ by a reiterative procedure. By comparing (42) with (38) we see that the maximum likelihood estimate $\hat{v}$ may be obtained by estimating the expectation $E(T^*)$ with the corresponding observed mean duration $\bar{T}^*$.

An alternative method of estimating v is to consider the probability distribution $f(t)dt$ in (31) of the complete duration of an illness. When $w_t = v$ is independent of t,

$$f(t)\,dt = ve^{-vt}\,dt\,, \qquad \text{for } t \ge 0, \tag{44}$$

and we have the expected duration of an illness

$$E(T) = \int_0^\infty tve^{-vt}\,dt = \frac{1}{v}\,. \tag{45}$$

Now we take a sample of M illnesses and record the *complete* duration of each illness. The observed mean duration, say $\bar{T}$, so determined is an estimate of $E(T)$, and its inverse is an estimate of v,

$$\hat{v} = \frac{1}{\bar{T}}\,. \tag{46}$$

For application later, we also compute the second and the third moments about the origin,

$$E(T^2) = \int_0^\infty t^2 \nu e^{-\nu t}\, dt = \frac{2}{\nu^2}$$

and (47)

$$E(T^3) = \int_0^\infty t^3 \nu e^{-\nu t}\, dt = \frac{6}{\nu^3}.$$

Using the estimated value $\hat{\nu}$ from either method, the expected relative frequency of illnesses with a truncated duration between, say t_1 and t_2 may be computed from

$$\int_{t_1}^{t_2} \frac{\hat{\nu} e^{-\hat{\nu} t}}{1 - e^{-\hat{\nu}}}\, dt = \frac{e^{-\hat{\nu} t_1} - e^{-\hat{\nu} t_2}}{1 - e^{-\hat{\nu}}} \tag{48}$$

for any interval (t_1, t_2). Multiplying (48) by M gives the expected number of illnesses. The discrepancy between the observed and expected number of illnesses for each interval (t_1, t_2) may be evaluated by the chi-square test to determine the sufficiency of the model in formula (36).

Equation (36) may be applied either to all diseases as a whole or separately to individual diseases. When all diseases are taken as a single group, the computations involved are quite simple. When individual diseases are studied separately, they can be summarized by the compound distribution

$$f^*(t)\, dt = \sum_j \pi_j f_j^*(t)\, dt, \tag{49}$$

where π_j is the known proportion of illnesses of the jth disease and $f_j^*(t)$ is the corresponding probability density function as given in (36). In this case, the overall expected duration of illness is given by

$$E(T^*) = \sum_j \pi_j \left[\nu_j^{-1} - \frac{e^{-\nu_j}}{1 - e^{\nu_j}} \right]. \tag{50}$$

Since π_j is known and ν_j can be estimated for each disease, the estimate of $\hat{E}(T^*)$ can be computed from (50).

Illnesses are often classified as acute or chronic. A mathematical model to describe illness from this viewpoint is essential but difficult to formulate, because the demarcation line between acute and chronic illness is not always well defined, and the exact proportions of illnesses in the two categories are never known. In the discussion to follow, we shall use the duration of the illness as a criterion of classification and make an attempt to solve the problem.

According to our formulation, a general mathematical model in this case may be represented by the probability density function

$$f^*(t)\, dt = \tag{51}$$

$$\left[\pi \frac{\nu_{1t} e^{-\int_0^t \nu_{1\tau} d\tau}}{1 - e^{-\int_0^1 \nu_{1\tau} d\tau}} + (1-\pi) \frac{\nu_{2t} e^{-\int_0^t \nu_{2\tau} d\tau}}{1 - e^{-\int_0^1 \nu_{2\tau} d\tau}} \right] dt$$

When $\nu_{1t} = \nu_1$ and $\nu_{2t} = \nu_2$ are assumed to be independent of time t, the probability function becomes

$$f^*(t)\, dt = \left[\pi \frac{\nu_1 e^{-\nu_1 t}}{1 - e^{-\nu_1}} + (1-\pi) \frac{\nu_2 e^{-\nu_2 t}}{1 - e^{-\nu_2}} \right] dt. \tag{52}$$

Here π is an *unknown* proportion and may be interpreted as the probability that an illness will be acute with the severity signified by ν_1. A similar interpretation holds for $(1-\pi)$ and ν_2. Consequent to our arbitrary classification of illnesses as acute or chronic, a graphic representation of the model will show a bimodal curve.

The statistical problem is to fit such a model to empirical data and to estimate the parameters π, ν_1, and ν_2 in the formula. In this case the simplest approach is to consider the complete duration of a sample of M illnesses and to use the method of moments, which allows the parameters to be estimated from the first three sample moments u_1, u_2, and u_3 (see equations (45) and (47))

$$
\begin{aligned}
u_1 &= \frac{\hat{\pi}}{\hat{v}_1} + \frac{1-\hat{\pi}}{\hat{v}_2} \\
u_2 &= \frac{2\hat{\pi}}{\hat{v}_1^2} + \frac{2(1-\hat{\pi})}{\hat{v}_2^2} \\
u_3 &= \frac{6\hat{\pi}}{\hat{v}_1^3} + \frac{6(1-\hat{\pi})}{\hat{v}_2^3}
\end{aligned}
\tag{53}
$$

These equations may now be solved for $\hat{\pi}$, $\hat{v}_1$ and $\hat{v}_2$. Eliminating $\hat{\pi}$ from the first two equations in (53) gives

$$2 - 2u_1(\hat{v}_1 + \hat{v}_2) + u_2\hat{v}_1\hat{v}_2 = 0 \tag{54}$$

and from all three equations gives

$$6u_1 - 3u_2(\hat{v}_1 + \hat{v}_2) + u_3\hat{v}_1\hat{v}_2 = 0 \tag{55}$$

Solving (54) and (55) simultaneously, we have

$$\hat{v}_1 + \hat{v}_2 = \frac{6u_1u_2 - 2u_3}{3u_2^2 - 2u_1u_3} \tag{56}$$

and

$$\hat{v}_1\hat{v}_2 = \frac{12u_1^2 - 6u_2}{3u_2^2 - 2u_1u_3} \tag{57}$$

These can be computed from the sample moments. Now let

$$\hat{v}_1 + \hat{v}_2 = a_1 \tag{58}$$

and

$$\hat{v}_1\hat{v}_2 = a_2\,, \tag{59}$$

and formulate a quadratic equation and

$$Y^2 - a_1Y + a_2 = 0\,. \tag{60}$$

The two roots of (60) are $\hat{v}_1$ and $\hat{v}_2$. Substituting these values into the first equation in (53), we obtain the estimate of π:

$$\hat{\pi} = \frac{\hat{v}_1(1 - u_1\hat{v}_2)}{\hat{v}_1 - \hat{v}_2} \tag{61}$$

Using the estimated values of π, v_1, and v_2, we can compute the expected relative frequency from

$$\int_{t_1}^{t_2} f(t)\,dt = \int_{t_1}^{t_2}\left[\hat{\pi}\frac{\hat{v}_1 e^{-\hat{v}_1 t}}{1-e^{-\hat{v}_1}} + (1-\hat{\pi})\frac{\hat{v}_2 e^{-\hat{v}_2 t}}{1-e^{-\hat{v}_2}}\right]dt \tag{62}$$

for each interval (t_1, t_2). The expected number of illnesses of durations between t_1 and t_2 is obtained by multiplying (62) by the total number of illnesses.

IV. TIME LOST DUE TO DEATH

In general, a high frequency of mortality indicates ill health within a population and, consequently, the death rate is one basic measure of the state of health. When the amount of illness in a calendar year is studied, not only the length of the illness before death must be considered, but also the period from the time of death to the end of the year. This period will be referred to as "the time lost due to death" and will be denoted by ξ. We want to determine the probability distribution of ξ, and the expected value of ξ among the deaths occurring during the year.

The time lost due to death is determined by the time of death. Although deaths may be subject to seasonal variation, as an approximation we assume that they take place uniformly throughout the year. Consequently the random variable ξ

also has a uniform distribution within the interval (0, 1), and its distribution function is given by

$$P\{\xi \le \theta\} = \int_0^\theta d\tau = \theta \text{, for } 0 \le \theta \le 1 . \tag{63}$$

Direct computation gives the expected value and variance of ξ,

$$E(\xi) = \frac{1}{2} \tag{64}$$

and

$$\sigma_\xi^2 = \int_0^1 \left[\xi - E(\xi)\right]^2 \quad d\xi = \frac{1}{12} . \tag{65}$$

This formulation indicates that the average time lost for each death is one-half year.

Let m_x denote the age-specific death rate, which is the approximate proportion of people of age x who die during the year. Since, on the average, one-half year is lost by each death, the average time lost to each individual in the entire subpopulation is $\frac{1}{2}\, m_x$.

The problem now is to determine how much weight to apply to the time lost due to death as compared with the duration of illness. In general, we may say that 1 day lost due to death is equivalent to w days of illness. The problem is to determine w. From the point of view of health, both illness and death are states of ill health, and time lost due to death and time lost due to illness should be weighed equally. It may then be assumed that w is unity. Thus $\frac{1}{2}\, m_x$, the average time lost due to death for an individual in the entire subpopulation, is directly comparable to the average duration of illness and will be used in the formulation of a measure of the state of health in the following section.

V. INDEX OF HEALTH

The term "illness" discussed in the preceding sections needs clarification before we proceed further. An illness may be defined as a continuous state of ill health over a period of time regardless of the number of diagnoses; it may also be defined as a continuous state of ill health for each disease. Consider, for example, a child who had the first symptom of chickenpox on February 1 and came down with a cold on February 11; he recovered from the chickenpox on February 14 and from the cold on February 18. According to the first definition, the child had one illness with a duration of 18 days; but according to the second, he had two illnesses with durations of 14 and 8 days, respectively. Each of the two definitions has its merit in describing health. The models suggested in the preceding sections should apply equally well in both cases, although the constants involved in the models will take on different values. For the purpose of deriving an index of health, however, we shall use the first definition and consider that the child enjoyed good health 365-18= 347 days during the year, providing, of course, there were no other illnesses for the rest of the year.

In the study of the state of health of a population, the paramount question would seem to be: What is the average fraction of the year in which an individual is healthy? This fraction will be referred to as the mean duration of health and will be used as an index of the health of a population.

The index of health so defined is closely related to the three component variables presented in the preceding sections. From the distribution of the number of illnesses, we can calculate the expected number of illnesses occurring to an individual in a calendar year; and from the study of the duration of illness, the expected duration of an illness. The product of the two quantities is the expected total duration of illness during the year. For age group x let $\bar{N}_x$ be the observed average number of illnesses per person and $\bar{T}_x{}^*$ be the average duration of an illness in a year. The product $\bar{N}_x\, \bar{T}_x{}^*$ is an estimate of the expected duration that an individual is ill. From the standpoint of the index of health, $\bar{N}_x\, \bar{T}_x{}^*$ is simply the average duration of illness per person per year and can be estimated directly from a single sample. Suppose a sample is taken from age group x. For each individual in the sample the fraction (I_x) of the year that he is ill is determined; the average of this fraction, denoted by $\bar{I}_x$, is equal to the product,

$$\bar{I}_x = \bar{N}_x \bar{T}_x{}^* . \tag{66}$$

In the discussion of mortality, we determined that the average time lost due to death for an individual of age group x is one-half the age-specific death rate. Since the age-specific death rate is usually available in vital statistics publications, it need not be computed again from a sample. The average duration of ill health is the sum of the average length of time that an individual is ill and the time lost due to death

$$\bar{I}_x + \frac{1}{2} m_x . \tag{67}$$

Let H_x denote the mean duration of health, or the fraction of the year in which an individual in age group x is living and free from illness. Obviously, H_x is the complement of the average duration of ill health,

$$H_x = 1 - (\bar{I}_x + \frac{1}{2} m_x). \tag{68}$$

In formula (68) the unit of I_x is years; the average duration of ill health cannot exceed one; and H_x is between zero and unity.

To derive the corresponding measure for the entire population, only a weighted average of H_x need be computed. Several principles may be used to determine the weights. The simplest is to use the population proportion in each age group as weights. Let P_x be the age-specific population and

$$P = \sum_x P_x \tag{69}$$

be the total population. Then the weighted average

$$H = \frac{1}{P} \sum_x P_x H_x \tag{70}$$

is the mean duration of health, or *the index of health*, for the entire population. Since P_x is the actual population having experienced the duration of health indicated by the value of H_x, in formula (70) we have a meaningful measure of the state of health.

The values of the index of health H are obviously between zero and one; the healthier a population is, the larger will be the value of H, and vice versa. If no illnesses and no deaths occur to the people of a current population, then $\bar{I}_x = 0$ and $m_x = 0$ for each x; hence $H_x = 1$ and $H = 1$. At the other extreme, if every individual in a population were ill during the entire year, then $\bar{I}_x = 1$, ($m_x = 0$ in this case), and both H_x and the index of health would assume the value of zero.

The quantities H_x and H are, of course, random variables, and they are estimates of the corresponding true unknown expected values. In terms of the models discussed in the preceding sections, the expected values may be derived as follows:

$$E(H) = \frac{1}{P} \sum P_x E(H_x), \tag{71}$$

$$E(H_x) = 1 - [E(\bar{I}_x) + \frac{1}{2} E(m_x)] \tag{72}$$

and

$$E(\bar{I}_x) = E(\bar{N}_x \bar{T}_x^*) \tag{73}$$

Assuming independence between $\bar{N}_x$ and $\bar{T}_x^*$, we have

$$E(\bar{N}_x \bar{T}_x^*) = E(\bar{N}_x)\, E(\bar{T}_x^*)$$

$$= E(N_x)\, E(T_x^*) \quad ; \tag{74}$$

where, in light of formulas (18) and (38),

$$E(N_x) = \frac{\alpha_x}{\beta_x} \tag{75}$$

and

$$E(T_x^*) = v_x^{-1} - e^{-v_x} \left[1 - e^{-v_x}\right]^{-1} \tag{76}$$

The expected value of m_x for the age group of n_x years is approximately equal to

$$\frac{q_x}{n_x\, [1 - (1 - a_x)\, q_x]} , \tag{77}$$

where q_x is the probability that an individual of exact age x will die before reaching exact age $x + n_x$, and a_x is the average fraction of the age interval $(x, x+n_x)$ lived by individuals who die at the age covered by the interval. It follows that

$$E(H_x) = 1 - \frac{\alpha_x}{\beta_x}\left[v_x^{-1} - e^{-v_x}(1-e^{-v_x})^{-1}\right]$$

$$-\frac{1}{2}\frac{q_x}{n_x\left[1-(1-a_x)q_x\right]} \qquad (78)$$

and the expectation of the index of health can be obtained upon substitution of (78) in (71).

For making statistical inferences and for comparing the state of health in different calendar years, it is necessary to know the variance of the index of health. It is not necessary here to derive the complicated formula for the true unknown variance; for practical purposes, it suffices to have the formula for the sample variance, which can be computed directly from the observed data without referring to the models discussed in the preceding sections. The sample variance of H can be written as

$$S_H^2 = \frac{1}{P^2}\sum P_x^2 S_{H_x}^2 , \qquad (79)$$

and for each age group x

$$S_{H_x}^2 = S_{\bar{I}_x}^2 + \frac{1}{4} S_{m_x}^2 , \qquad (80)$$

where the sample variance of the average duration of illness $S_{\bar{I}_x}^2$ can be computed directly from the sample. The sample variance of the age-specific death rate is given in reference 2.

$$S_{m_x}^2 = \frac{m_x(1-a_x n_x m_x)}{P_x\left[1+(1-a_x)n_x m_x\right]} \qquad (81)$$

VI. REMARK—ADJUSTED INDEX OF HEALTH

The index of health H is evidently a meaningful and useful measure of the state of health in a single population. Since it is a weighted average of H_x with the current population proportion P_x/P as the weight, however, the value of H is affected by the current population composition. Such an effect will produce a distortion when two populations with different age compositions are compared. To adjust for the difference, we may use a standard population and compute the weighted average of H_x,

$$H^* = \frac{1}{P_s}\sum P_{sx} H_x , \qquad (82)$$

where P_{sx}/P_s is the population proportion of age group x in the standard population. This weighted average may be called the age-adjusted index of health. The sample variance of H^* is given by

$$S_{H^*}^2 = \sum \frac{P_{sx}^2}{P_s^2} S_{H_x}^2 , \qquad (83)$$

where the sample variance of H_x may be computed from formula (80).

REFERENCES

1. Feller, W.: *An Introduction to Probability Theory and Its Applications.* 2d edition. New York. John Wiley and Sons, Inc., 1957.
2. National Vital Statistics Division: Standard error of the age-adjusted death rate, by C. L. Chiang. *Vital Statistics--Special Reports,* Vol. 47, No. 9. Public Health Service. Washington, D.C., 1961.
3. The Department of National Health and Welfare and The Dominion Bureau of Statistics: *Illness and Health Care in Canada, Canadian Sickness Survey, 1950-51.* Ottawa. The Queen's Printer and Controller of Stationery, 1960.
4. The Dominion Bureau of Statistics and The Department of National Health and Welfare; *Canadian Sickness Survey, 1950-51.* No. 8. Volume of Health Care (National Estimates); D.B.S. Reference Paper No. 51. Ottawa The Queen's Printer and Controller of Stationery, 1955.

———ooo———

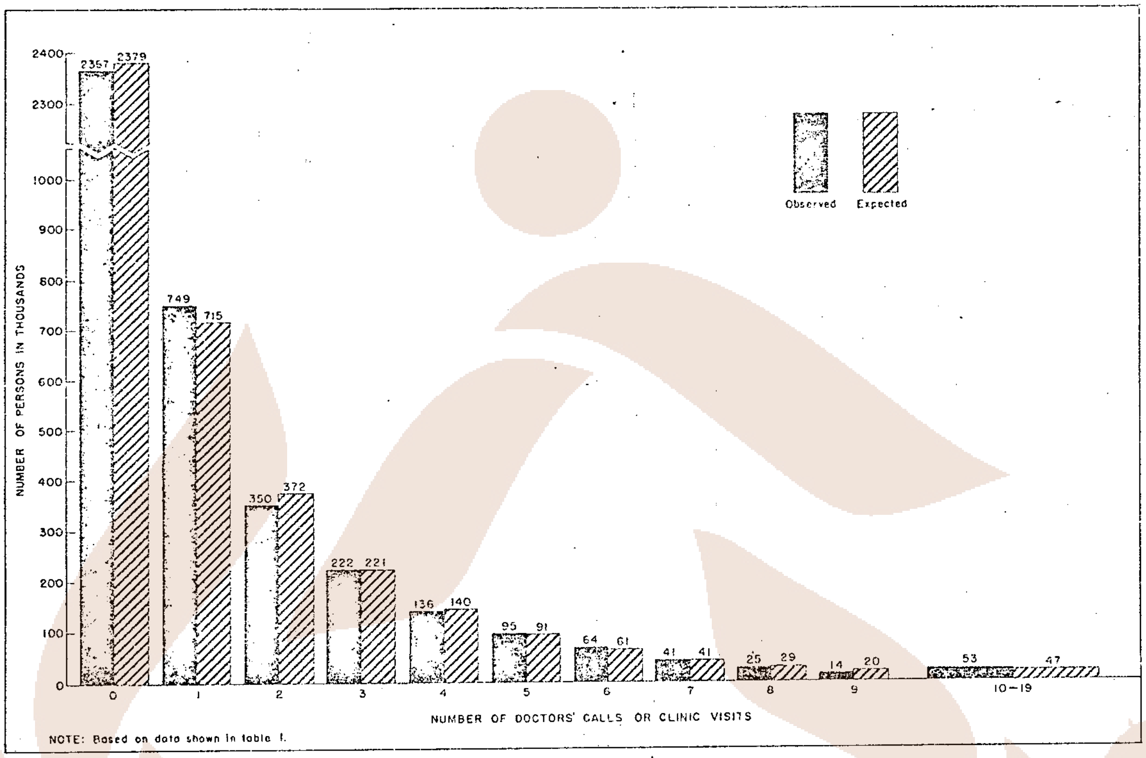

Figure 1. Observed and expected number of persons under 15 years of age by number of doctors' calls or clinic visits in a year.

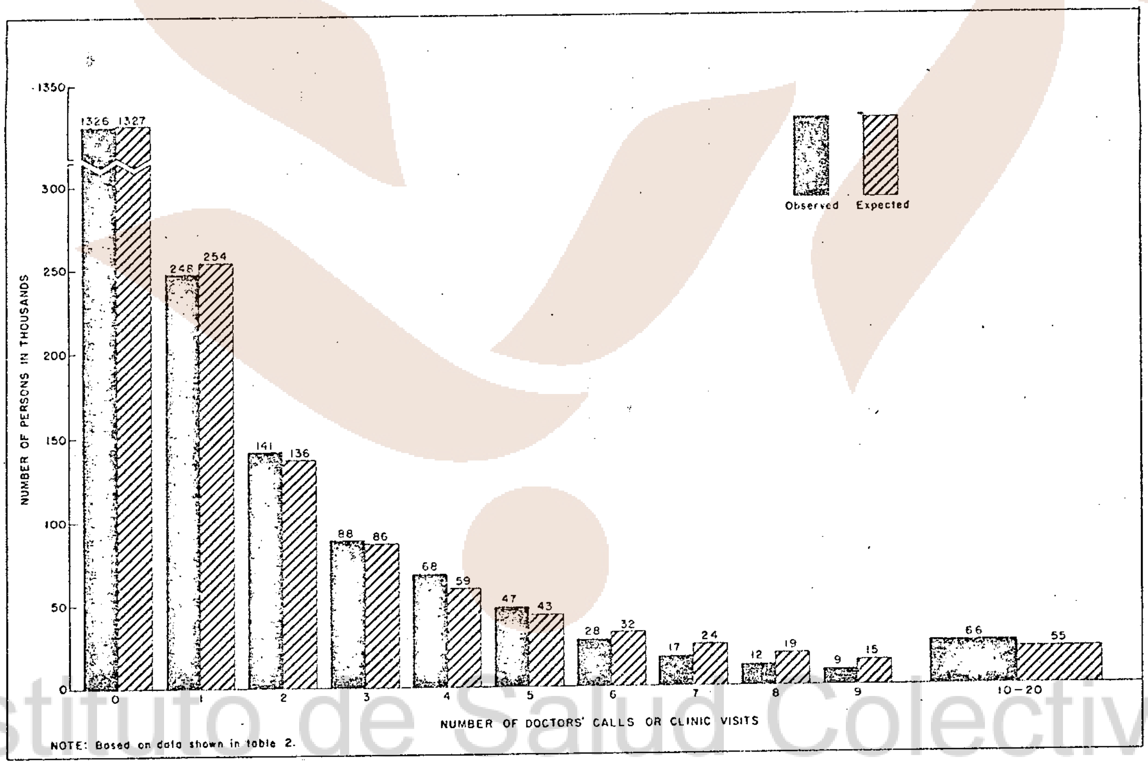

Figure 2. Observed and expected number of persons 15-24 years of age by number of doctors' calls or clinic visits in a year.

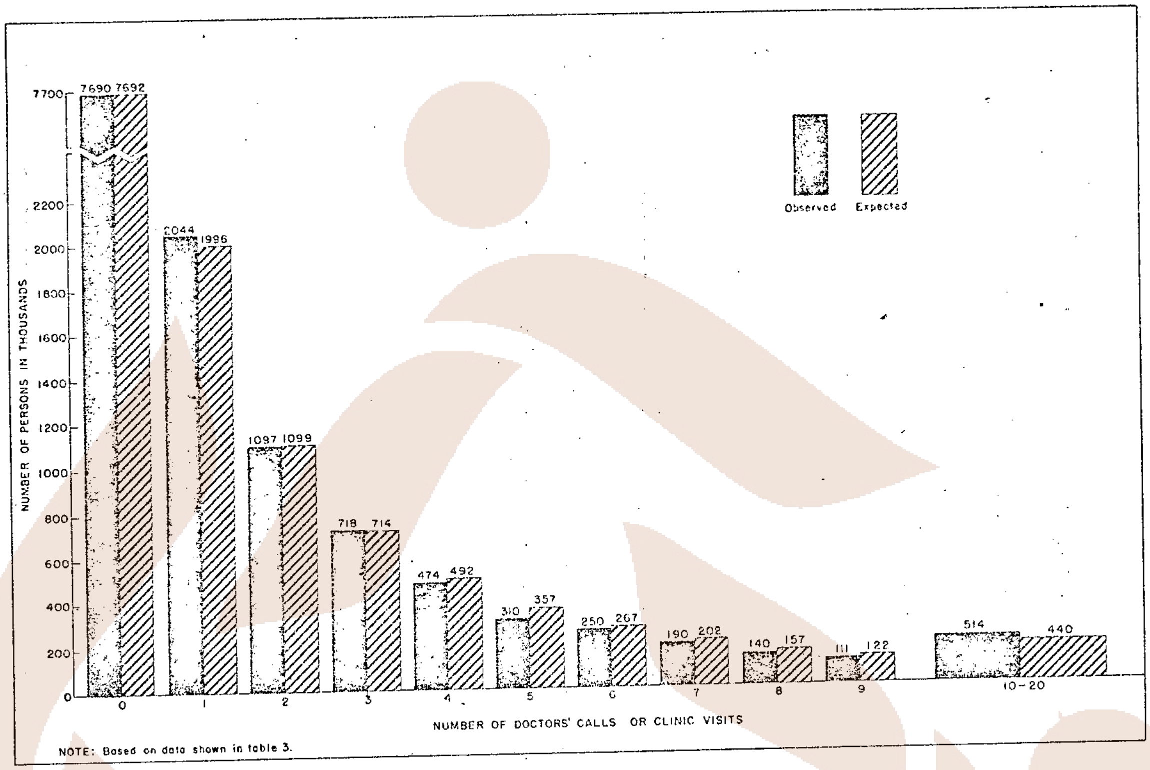

Figure 3. Observed and expected number of persons of all ages by number of doctors' calls or clinic visits in a year.

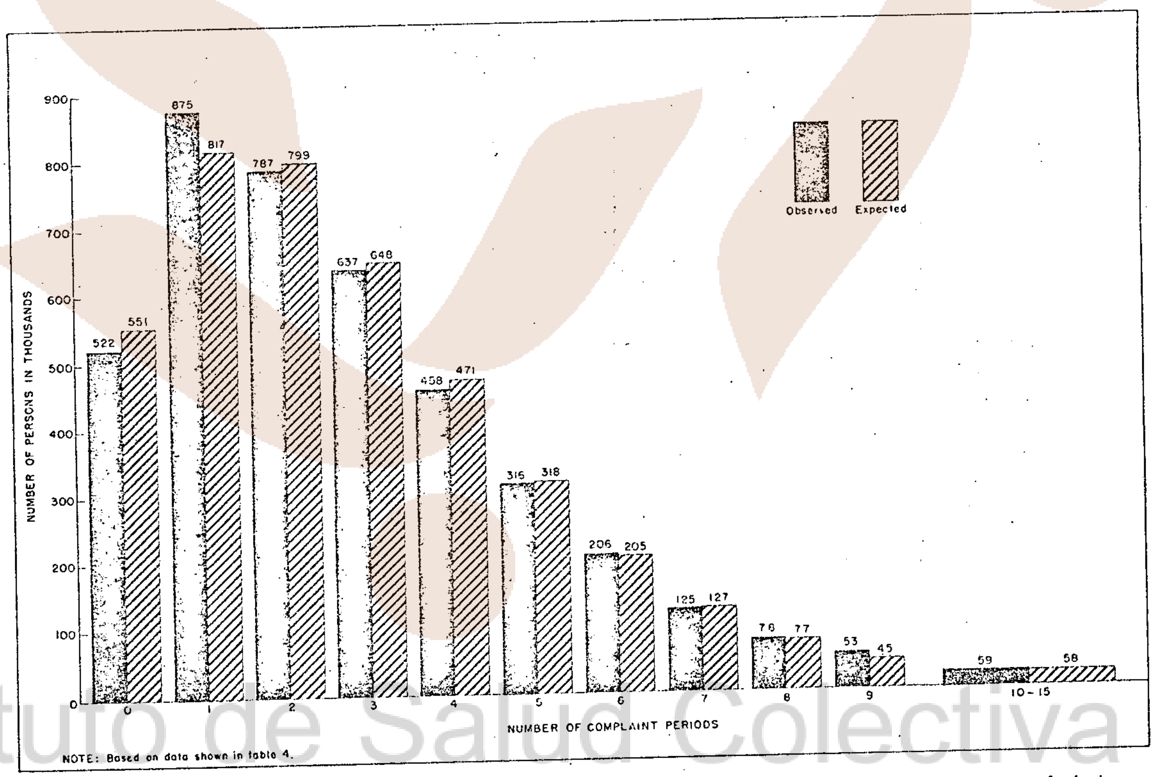

Figure 4. Observed and expected number of persons under 15 years of age by number of complaint periods in a year.

DETAILED TABLES

Table 1. Observed and expected number of persons under 15 years of age, by number of doctors' calls or clinic visits in a year

Number of doctors' calls or clinic visits, n	Observed f_n	Expected F_n	Difference $f_n - F_n$	$\frac{(f_n - F_n)^2}{F_n}$
	Number of persons in thousands[1]			
Total-------------------------	4,116	4,116	0	[1]3.899
0 call or visit---------------------	2,367	2,379	-12	0.060
1 call or visit---------------------	749	715	+34	1.617
2 calls or visits-------------------	350	372	-22	1.301
3 calls or visits-------------------	222	221	+1	0.004
4 calls or visits-------------------	136	140	-4	0.114
Subtotal, 5-9 calls or visits--	239	242	-3	0.037
5 calls or visits-------------------	95	91	+4	
6 calls or visits-------------------	64	61	+3	
7 calls or visits-------------------	41	41	0	
8 calls or visits-------------------	25	29	-4	
9 calls or visits-------------------	14	20	-6	
Subtotal, 10+ calls or visits--	53	47	+6	0.766
10 calls or visits------------------	12	14	-2	
11 calls or visits------------------	11	10	+1	
12 calls or visits------------------	9	7	+2	
13 calls or visits------------------	8	5	+3	
14 calls or visits------------------	5	3	+2	
15+ calls or visits-----------------	8	8	0	

$\bar{N} = 1.163$ $\hat{\alpha} = 0.405$ $\chi^2 = 3.899$

$S^2 = 4.500$ $\hat{\beta} = 0.348$ $d.f. = 4$

Source: Observed f_n were calculated from percent distributions shown in table 7-C, page 27 of Reference 4, and population totals shown in table 114, page 193 of Reference 3. To estimate the parameters involved in the model, subtotals shown were distributed by the number of calls or visits to obtain f_n for each n in the respective groups.

[1]For justification in using thousand as a single count in computing the χ^2, see text on page 4.

Table 2. Observed and expected number of persons 15-24 years of age, by number of doctors' calls or clinic visits in a year

Number of doctors' calls or clinic visits, n	Observed f_n	Expected F_n	Difference $f_n - F_n$	$\frac{(f_n - F_n)^2}{F_n}$
	Number of persons in thousands[1]			
Total	2,050	2,050	0	[1]6.955
0 call or visit	1,326	1,327	-1	0.001
1 call or visit	248	254	-6	0.142
2 calls or visits	141	136	+5	0.184
3 calls or visits	88	86	+2	0.047
4 calls or visits	68	59	+9	1.373
Subtotal, 5-9 calls or visits	113	133	-20	3.008
5 calls or visits	47	43	+4	
6 calls or visits	28	32	-4	
7 calls or visits	17	24	-7	
8 calls or visits	12	19	-7	
9 calls or visits	9	15	-6	
Subtotal, 10+ calls or visits	66	55	+11	2.200
10 calls or visits	8	11	-3	
11 calls or visits	7	9	-2	
12 calls or visits	7	7	0	
13 calls or visits	6	6	0	
14 calls or visits	5	5	0	
15 calls or visits	4	4	0	
16 calls or visits	4	3	+1	
17 calls or visits	3	3	0	
18 calls or visits	3	2	+1	
19 calls or visits	2	2	0	
20+ calls or visits	17	3	+14	

$\bar{N} = 1.347$ $\hat{\alpha} = 0.223$ $\chi^2 = 6.955$

$S^2 = 9.478$ $\hat{\beta} = 0.166$ $d.f. = 4$

Source: Observed f_n were calculated from percent distributions shown in table 7-C, page 27 of Reference 4, and population totals shown in table 114, page 193 of Reference 3. To estimate the parameters involved in the model, subtotals shown were distributed by the number of calls or visits to obtain f_n for each n in the respective groups.

[1]For justification in using thousand as a single count in computing the χ^2, see text on page 4.

Table 3. Observed and expected number of persons of all ages, by number of doctors' calls or clinic visits in a year

Number of doctors' calls or clinic visits, n	Observed f_n	Expected F_n	Difference $f_n - F_n$	$\frac{(f_n - F_n)^2}{F_n}$
	Number of persons in thousands[1]			
Total	13,538	13,538	0	[1]24.072
0 call or visit	7,690	7,692	-2	0.001
1 call or visit	2,044	1,996	+48	1.154
2 calls or visits	1,097	1,099	-2	0.004
3 calls or visits	718	714	+4	0.022
4 calls or visits	474	492	-18	0.658
Subtotal, 5-9 calls or visits	1,001	1,105	-104	9.788
5 calls or visits	310	357	-47	
6 calls or visits	250	267	-17	
7 calls or visits	190	202	-12	
8 calls or visits	140	157	-17	
9 calls or visits	111	122	-11	
Subtotal, 10+ calls or visits	514	440	+74	12.445
10 calls or visits	97	96	+1	
11 calls or visits	79	77	+2	
12 calls or visits	57	61	-4	
13 calls or visits	45	49	-4	
14 calls or visits	33	39	-6	
15 calls or visits	29	31	-2	
16 calls or visits	24	26	-2	
17 calls or visits	18	20	-2	
18 calls or visits	13	16	-3	
19 calls or visits	11	14	-3	
20+ calls or visits	108	11	+97	

$\bar{N} = 1.631$ $\hat{\alpha} = 0.307$ $\chi^2 = 24.072$

$S^2 = 10.286$ $\hat{\beta} = 0.188$ $d.f. = 4$

Source: Observed f_n were calculated from percent distributions shown in table 7-C, page 27 of Reference 4, and population totals shown in table 114, page 193 of Reference 3. To estimate the parameters involved in the model, subtotals shown were distributed by the number of calls or visits to obtain f_n for each n in the respective groups.

[1]For justification in using thousand as a single count in computing the χ^2, see text on page 4.

Table 4. Observed and expected number of persons under 15 years of age, by number of complaint periods in a year

Number of complaint periods, n	Observed f_n	Expected F_n	Difference $f_n - F_n$	$\frac{(f_n - F_n)^2}{F_n}$
	Number of persons in thousands[1]			
Total	4,116	4,116	0	[1]7.871
0 complaint period	522	551	-29	1.526
1 complaint period	875	817	+58	4.118
2 complaint periods	787	799	-12	0.180
3 complaint periods	637	648	-11	0.187
4 complaint periods	458	471	-13	0.359
5 complaint periods	316	318	-2	0.013
6 complaint periods	206	205	+1	0.005
7 complaint periods	125	127	-2	0.031
8 complaint periods	78	77	+1	0.013
9 complaint periods	53	45	+8	1.422
Subtotal, 10+ complaint periods	59	58	+1	0.017
10 complaint periods	31	26	+5	
11 complaint periods	14	15	-1	
12 complaint periods	5	8	-3	
13 complaint periods	4	5	-1	
14 complaint periods	2	3	-1	
15+ complaint periods	3	1	+2	

$\bar{N} = 2.826$ $\hat{\alpha} = 3.113$ $\chi^2 = 7.871$

$S^2 = 5.392$ $\hat{\beta} = 1.102$ $d.f. = 8$

Source: Table 31, page 122 of Reference 3. To estimate the parameters involved in the model, the subtotal shown was distributed by the number of complaint periods to obtain f_n for each n in the group.

[1]For justification in using thousand as a single count in computing the χ^2, see text on page 4.

☆ U.S. GOVERNMENT PRINTING OFFICE : 1965 O—770-162

REPORTS FROM THE NATIONAL CENTER FOR HEALTH STATISTICS
Public Health Service Publication No. 1000

Series 1. Programs and collection procedures
- No. 1. Origin, Program, and Operation of the U.S. National Health Survey. 35 cents.
- No. 2. Health Survey Procedure: Concepts, Questionnaire Development, and Definitions in the Health Interview Survey. 45 cents.
- No. 3. Development and Maintenance of a National Inventory of Hospitals and Institutions. 25 cents.

Series 2. Data evaluation and methods research
- No. 1. Comparison of Two Vision-Testing Devices. 30 cents.
- No. 2. Measurement of Personal Health Expenditures. 45 cents.
- No. 3. The One-Hour Glucose Tolerance Test. 30 cents.
- No. 4. Comparison of Two Methods of Constructing Abridged Life Tables. 15 cents.
- No. 5. An Index of Health: Mathematical Models.

Series 3. Analytical studies
- No. 1. The Change in Mortality Trend in the United States. 35 cents.
- No. 2. Recent Mortality Trends in Chile. 30 cents.

Series 4. Documents and committee reports
No reports to date.

Series 10. Data From the Health Interview Survey
- No. 1. Acute Conditions, Incidence and Associated Disability, United States, July 1961-June 1962. 40 cents.
- No. 2. Family Income in Relation to Selected Health Characteristics, United States. 40 cents.
- No. 3. Length of Convalescence After Surgery, United States, July 1960-June 1961. 35 cents.
- No. 4. Disability Days, United States, July 1961-June 1962. 45 cents.
- No. 5. Current Estimates From the Health Interview Survey, United States, July 1962-June 1963. 35 cents.
- No. 6. Impairments Due to Injury, by Class and Type of Accident, United States, July 1959-June 1961. 25 cents.
- No. 7. Disability Among Persons in the Labor Force, by Employment Status, United States, July 1961-June 1962. 40 cents.
- No. 8. Types of Injuries, Incidence and Associated Disability, United States, July 1957-June 1961. 35 cents.
- No. 9. Medical Care, Health Status, and Family Income, United States. 55 cents.
- No. 10. Acute Conditions, Incidence and Associated Disability, United States, July 1962-June 1963. 45 cents.
- No. 11. Health Insurance Coverage, United States, July 1962-June 1963. 35 cents.
- No. 12. Bed Disability Among the Chronically Limited, United States, July 1957-June 1961. 45 cents.
- No. 13. Current Estimates From the Health Interview Survey, United States, July 1963-June 1964. 40 cents.
- No. 14. Illness, Disability, and Hospitalization Among Veterans, United States, July 1957-June 1961. 35 cents.
- No. 15. Acute Conditions, Incidence and Associated Disability, United States, July 1963-June 1964.
- No. 16. Health Insurance, Type of Insuring Organization and Multiple Coverage, United States, July 1962-June 1963.
- No. 17. Chronic Conditions and Activity Limitation, United States, July 1961-June 1962.

Series 11. Data From the Health Examination Survey
- No. 1. Cycle I of the Health Examination Survey: Sample and Response, United States, 1960-1962. 30 cents.
- No. 2. Glucose Tolerance of Adults, United States, 1960-1962. 25 cents.
- No. 3. Binocular Visual Acuity of Adults, United States, 1960-1962. 25 cents.
- No. 4. Blood Pressure of Adults, by Age and Sex, United States, 1960-1962. 35 cents.
- No. 5. Blood Pressure of Adults, by Race and Region, United States, 1960-1962. 25 cents.
- No. 6. Heart Disease in Adults, United States, 1960-1962. 35 cents.
- No. 7. Selected Dental Findings in Adults, United States, 1960-1962. 30 cents.

Series 12. Data From the Health Records Survey
No reports to date.

Series 20. Data on mortality
No reports to date.

Series 21. Data on natality, marriage, and divorce
- No. 1. Natality Statistics Analysis, United States, 1962. 45 cents.
- No. 2. Demographic Characteristics of Persons Married Between January 1955 and June 1958, United States.

Series 22. Data from the program of sample surveys related to vital records
No reports to date.