

SECTION 3. COMPUTER SCIENCE

3.1 Informational and mathematical modeling in epidemiological studies of acute respiratory diseases

Combining computer modeling statistics and methods with dynamic methods can make modeling and the resulting analysis more realistic and reliable, and more accessible to understand the rules for the spread of infectious diseases.

The science of the patterns of origin, spread and course of diseases detected in population studies is called epidemiology. Epidemics are massive, short-term outbreaks. The disease is endemic if it depends on the population. Thus, epidemic models are used to describe a rapid outbreak that occurs within one year, while endemic models are used to study disease for a long period after which recovery occurs (including with the acquisition of temporary immunity). The spread of infectious diseases includes not only disease-related factors such as infectious agent, mode of transmission, latency, infectious period, susceptibility and persistence, but also a number of social, cultural, demographic, economic and geographical factors [125].

A number of studies continue to be conducted by domestic scientists related to the "pressure" on the epidemiology of "pure" mathematicians, who managed to create many abstract models, but with very limited epidemiological content. These are more complex mathematical models in which random factors of the epidemic process play a significant role, so most of these models are stochastic (probabilistic) in nature, and the working apparatus is the theory of probability and random processes. Research initiated by academicians continues to develop a new methodology for mathematical modeling of epidemics - epidemic dynamics. This methodology is based on the method of scientific analogy in the reflection of the epidemic process (the process of "transfer" of the pathogen from sick to healthy) with the process of "transfer" of matter (energy, momentum) in the equations of mathematical physics. Indeed, in the course of the epidemic, a complex process of "transfer" of the population of the self-sustaining

pathogen to the population of susceptible people is being formed among the population of the territory affected by the infectious disease.

The epidemiological content of this process is associated with its adequate reflection, as in calendar time « t », and in internal time « τ », which captures the development of an infectious disease in many people affected by the infection. The system of equations, which describes the development of the epidemic process, is a system of nonlinear equations in partial derivatives with corresponding initial and boundary conditions very "similar" to the equations of hydrodynamics [126].

Mathematical epidemiology has now become a separate area of population dynamics, parallel to mathematical ecology. Epidemiological models are used to combine data from different sources to analyze complex results. The most important value of epidemiological modeling is the basic number of reproductions (threshold value) R_0 , defined as the average number of people infected by one sick person. Value R_0 important for control and threshold values predict the outcome of the disease, whether the disease can invade a susceptible population if $R_0 > 1$. Using *SIR*-models for endemic diseases without age structure, estimates R_0 for different diseases show that some diseases spread faster than others, that is, they are less controlled.

Many models of epidemiology are currently used to study infectious diseases involving age structures, but mortality and contact depend on age. Thus, the base number of reproductions R_0 also depends on these epidemiological and demographic factors. When R_0 found by study, in models with positive (endemic) equilibrium with acceptable limits, and then tested.

Detailed temporal patterns of recurring epidemics can be described and analyzed using time series analysis tools. In terms of their ability to predict the patterns of recurrence in the future. Mechanistic mathematical models are extremely powerful for epidemiological research [128].

The mathematical model is the basis for explaining reasons, attitudes, and ideas. Mathematical models are very useful for obtaining conclusions that have a clear interpretation. Epidemiologists are helped by epidemiologists to master the conceptual

and intuitive understanding of pathological processes. Epidemiological modeling is an important part of the function of epidemiology to build and test theories. Mathematical and computer modeling of epidemics are the main tools in experimental epidemiology. The only data that are usually available from natural epidemics and from the natural incidence of endemic diseases, unfortunately, are not complete, as due to various factors are not recorded accurately. Because repetitive experiments and accurate data are generally not available in epidemiology, mathematical and computer models and simulations are used to perform the necessary theoretical experiments with different parameter values and data sets. With the help of a computer simulation, it is easy to find out what happens when one or more parameters have changed. Another advantage of epidemiological modeling is the availability of models for estimating quantitative hypotheses. For example, the model can be used to test the claim that two doses of measles vaccination program will lead to collective immunity while one dose of vaccine will not be useful. Epidemiological modeling refers to dynamic, deterministic modeling in which the population is divided into compartments according to their epidemiological status (eg susceptible, infected, and recovered), with the movement between compartments being defined as differential or differential equations.

Modeling can improve the quality of the epidemiological survey by providing a testing ground based on the analysis of the survey results. The design phase can help identify important issues in the emergence, spread and control of epidemics that need to be addressed to ensure that the results are adequate. The statistical aspects of the survey are currently being carried out before the surveys themselves are conducted. In addition, these aspects related to the estimation of parameters in the models are analyzed before conducting surveys to ensure that all relevant data are collected.

Another goal of epidemiological modeling is to predict future cases. As a rule, it is not necessary to idealize epidemiological models and the accuracy of forecasts obtained with their help. However, with their help you can see that the epidemic can take different scenarios. One of the goals of epidemiological modeling is to reduce uncertainty in future forecasts [129].

The development of compartment models is based on the epidemiological

assumption that susceptible compartments (S) pass to the compartment of persons in the latent (or latent) phase of the disease (E and L) (individuals are already infected, but do not yet infect others), and then move on to the compartment of infected (I) (already infecting others) and, finally, to the compartment (R) persons who have recovered and have permanent immunity.

In many cases, it is sufficient to use the SIR model, because, for example, in this case the behavior remains the same as for the SEIR model. Parameters such as birth rate, mortality and rate of infection are stable in the models. The epidemiological model should answer the question of whether an epidemic of the disease will spread in a short period of time. In this case, the dynamic change in birth and death rates can be neglected. Otherwise, if the disease lasts for a long period (many years), you should take into account the dynamic changes in these indicators.

As the model becomes more complex, the number of compartments increases to match the real ways of developing the epidemic. The studied population is divided into disparate classes - compartments, the size of which changes over time t . In the main epidemiological models, it is assumed that the population is stable N , which is large enough so that the dimensions of each class can be considered as continuous variables. If the model includes changes in births and deaths over time (also called "life dynamics"), it is assumed that births and natural deaths occur at equal rates, and that all newborns are susceptible. The main concept in epidemiology is the existence of threshold levels. Endemic is a term in epidemiology that means the constant existence of a certain infectious disease in any area. Infection is endemic when it is constantly maintained in this population without the need for external sources. For an infection to be endemic, on average, each infected person must transmit the infection to another person. Assuming a perfectly susceptible population, this means that the baseline number of reproductions (R_0) infection should be equal to 1. If in the population some individuals have immunity, the baseline number of reproductions R_0 multiplied by the share of susceptible individuals in the population (s) should be equal to 1. $R_0 \times s = 1$

For a disease that is in a stable endemic state, the number of infected people does not decrease or increase [130].

The main types of models are proposed to describe ARI outbreaks. The three main deterministic models for infectious diseases that are spread through direct human contact with an infected person in a population are *MSEIR*, *SIS*, *SIR* – models. These models were used to obtain the main results, and then supplemented with additional compartments for refinement. Although the models presented are simple, they provide a basis for considering more complex models.

Table 1.1.

Designation to the model MSEIR,

M	Passive-immune compartment
S	Compartment susceptible
E	Compartment laten
I	Compartment of the infected
R	Compartment of people who have recovered with acquired immunity
β	Contact speed
μ	Average mortality rate
B	Average birth rate
$1/\varepsilon$	Average latency period
$1/\gamma$	The average infectious period
R_0	Baseline reproduction number (threshold value)
N	Total population
f	The rate of loss of immunity of recovering persons
δ	Average time of passive immunity

M – compartment containing persons with temporary passive immunity, ie acquired from a once infected mother through antibodies transmitted through the

placenta. Eventually, the antibodies disappear from the body and the person becomes receptive S . When a susceptible person comes in contact with an infected person, an infection occurs and the person becomes a compartment E . After the latent period the person passes to a compartment of the infected I . If the infectious period ends, the person goes to the compartment R , consisting of those who became ill and acquired immunity. The choice of these compartments for inclusion in the model depends on the characteristics of the specific disease. Passive-immune class M and latency E they are often omitted because they are not important. Abbreviated names of the epidemiological model are based on the movement between compartments, for example $MSEIR, MSEIRS, SEIRS, SIR, SIS$. Example $MSEIRS$ - the model will be similar to $MSEIR$, but immunity in the compartment R will be temporary, ie individuals return to the compartment of the susceptible when their temporarily acquired immunity disappears [131].

Model *SIR epidemic*. Because the epidemic occurs in a short period of time, temporary loss of immunity and natural births and deaths are ignored. Therefore, there is no flow from the compartment of healthy back to the compartment of susceptible, also passes the passive-immune compartment M . By the way, the model *SEIR* epidemic has a similar behavior as the basic model of the *SIR* epidemic. For this model, the main initial task is:

$$\begin{aligned} dS / dt &= \beta IS / N, & S(0) &= S_0 \geq 0, \\ dI / dt &= \beta IS / N - \gamma I, & I(0) &= I_0 \geq 0, \\ dR / dt &= \gamma I & R(0) &= R_0 \geq 0. \end{aligned} \quad (1.1)$$

There $S(t), I(t), R(t)$ satisfy the condition $S(t) + I(t) + R(t) = N$.

Model *SIR* is a special case of the model *MSEIR*.

In general *SIR* – The model is suitable for viral diseases such as measles, mumps, smallpox, also describes some bacterial diseases such as ARI, meningitis, plague and sexually transmitted diseases, the simplest diseases such as malaria.

Statement of control problems in epidemiological mathematical models. Consider the Makinde model:

$$\begin{aligned}
\frac{dS(t)}{dt} &= (1-P)\gamma N - \beta I(t)S(t)/N - \mu S(t), \\
\frac{dI(t)}{dt} &= \beta I(t)S(t)/N - (\eta + \mu)I(t), \\
\frac{dR(t)}{dt} &= P\gamma N + \eta I(t) - \mu R(t)
\end{aligned} \tag{1.2}$$

It is considered η, γ, β, μ all positive constant parameters with $0 \leq p < 1$

Consider the problem of optimal control to minimize the target functionality

$$J(u_1, u_2) = \int_0^{t_{end}} \left[S(t) + I(t) + \frac{(A_1 u_1^2(t) + A_2 u_2^2(t))}{2} \right] dt \tag{1.3}$$

for the system

$$\begin{aligned}
\frac{dS(t)}{dt} &= (1-P)\gamma N - \beta I(t)S(t)/N - \mu S(t) - u_1(t), \\
\frac{dI(t)}{dt} &= \beta I(t)S(t)/N - (\eta + \mu)I(t) - u_2(t), \\
\frac{dR(t)}{dt} &= P\gamma N + \eta I(t) - \mu R(t) + u_1(t)
\end{aligned} \tag{1.4}$$

with initial conditions $S(0) = S_0, I(0) = I_0, R(0) = R_0$.

In system (1.4) we have three state variables $S(t), I(t)$ and $R(t)$. Variable control $u(t) = (u_1(t), u_2(t)) \in U$. The problem of optimal control to limit control $0 \leq u_1(t) + u_2(t) < N$, where $u_1(t) \geq u_2(t) \geq 0$, for all $t \in [0, t_{end}]$.

The theoretical foundations of optimal control on which dynamic models are based are given by ordinary differential equations. This theory was developed by Pontryagin and his collaborators. The theory of applications and corresponding numerical algorithms has steadily progressed. By using the Pontryagin maximum principle, and extending its corresponding numerical methods, control is adjusted in the model to achieve the goal [132].

Analytical results for optimal management are applied in a simple epidemic *SIR* - models, including the calculation of optimal management taking into account epidemic control measures, vaccination and quarantine.

Today there are many software tools for use in epidemiology. The site [<http://www.ehdp.com/links/episw.htm>] lists software products used in epidemiological studies.

Cytel [<http://www.cytel.com>] develops statistical software for planning any type of clinical trial, including adaptive planning. This software is used by data monitoring commissions to recruit patients to monitor and implement adaptive clinical trials.

EpiData opposes open source software for entering and analyzing epidemiological data [<http://www.epidata.dk>]. EpiData Entry is used to easily enter or document data from another program and document it. Data entry is based on simple forms, optimization of documentation and has the ability to detect database errors. EpiData Analysis performs basic statistical analysis, graphing and data management.

EpiGram [<http://www.brixtonhealth.com/epigram.html>] is designed to build charts for inclusion in articles and reports. It was designed to track epidemiological contacts for inclusion in epidemiological reports of epidemic outbreaks. Although in fact the program is a universal tool for charting. The Centers for Disease Control and Prevention (USA) offers one of the most popular and versatile software tools in epidemiological research – Epi Info™, which is a simple tool for social workers with little experience in information technology - to quickly create data sets and their analysis, visualization and preparation of reports involving methods of epidemiology. Epi Info™ – it is actually a software package that provides basic epidemiological functionality without the hassle of large-scale commercial applications.

Epi Info™ is easy to use in places with limited network connections or limited resources for commercial software and professional IT support. Epi Info™ is flexible, scalable and freely available in data collection, thorough statistical analysis and compatibility with geographic information systems (GIS). Users of this system are registered in 181 countries (and Antarctica). The system has a version available for non-specialists (public domain version). Sentinel Research & Consulting Corp has developed a mobile application for iOS devices. iEpidemiology is an easy-to-use calculator for analyzing epidemiological research data (Figure.1.1). The program allows you to enter research data (epidemic / non-epidemic, epidemic curve).

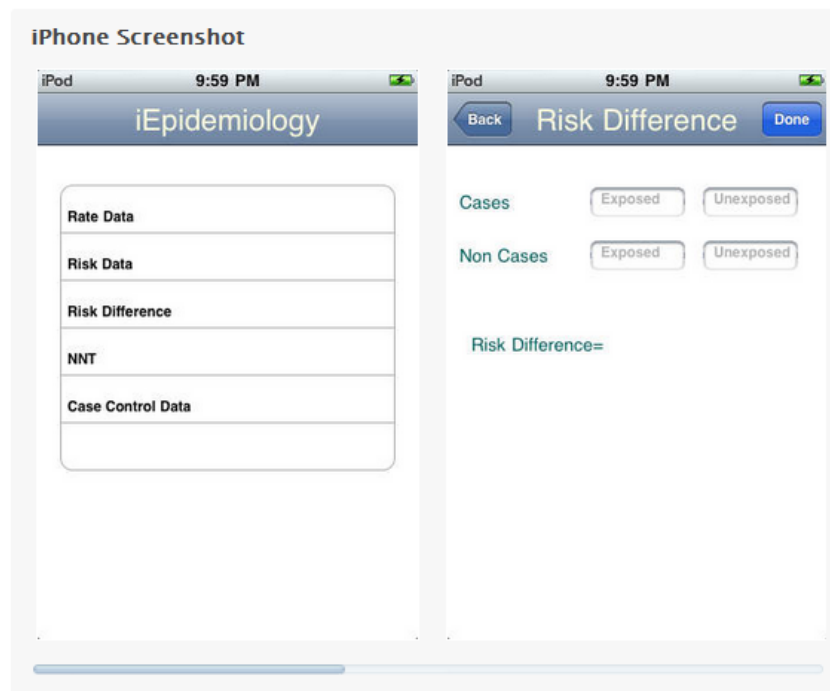


Figure.1.1. Mobile application interface for iOS devices from Sentinel Research & Consulting Corp.

The OpenEpi package [<http://www.openepi.com/>] is open source software for epidemiological research. It can run on a server, or it can be downloaded and run without a web connection. Programs are written in JavaScript and HTML, they are compatible with Linux, Mac in PC browsers. It also declares browser support for most iPhones and Android. The OpenEpi menu is shown in Figure. 1.2. The program provides statistics for descriptive and analytical studies, calculation of confidence intervals, sample sizes, sensitivity and specificity, R x C - tables, chi-square criteria for the problem "dose - effect".

The NetEpi package [<https://code.google.com/p/netepi/>] is a set of free, open source software that is also fully integrated with the Internet. To date, there are two main NetEpi applications.

NetEpi Collection is a tool for data collection and management for use in the study of outbreaks and other epidemiological studies. The program is written in Python and uses only open source software components and infrastructure, including the PostgreSQL database.

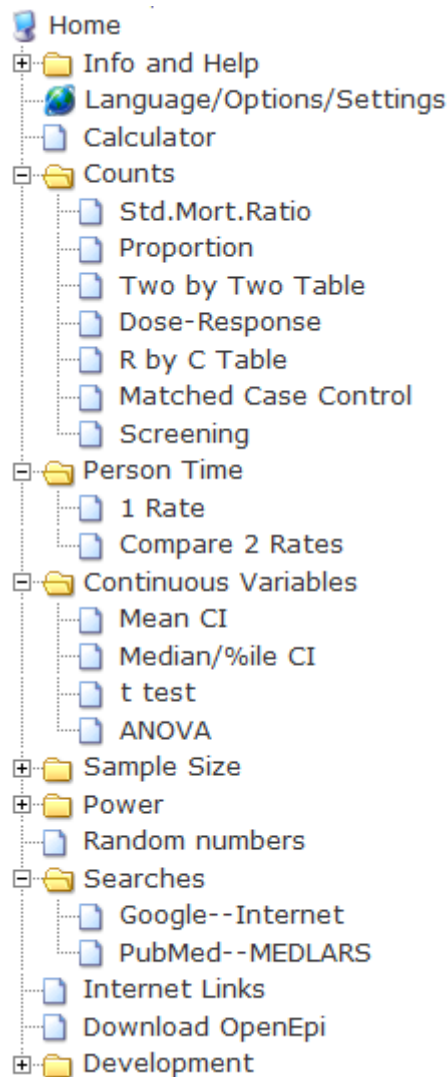


Figure.1.2. "Tree" menu of the OpenEpi program

NetEpi Analysis is a tool for interactive data analysis of large population datasets (10 to 100 million records). The product is written mainly in Python, it is possible to use Numeric Python (NumPy) and open source statistical environment R. NetEpi Analysis uses a simple but innovative approach to filtering, reducing and summarizing data, including object-oriented implementation of fast multiple operations on sorted inverted ordinal mappings. Both a web browser interface and a Python API are available. From the given review of software products in epidemiology it is visible that there is an orientation exclusively on statistical processing of epidemiological data with their visualization. At the same time, there are no specialized tools to support decision-making - through such tools as compartmental modeling, qualitative analysis of epidemiological systems, building knowledge structures.