Knowledge-Oriented System in the Development of Functional Nutrition

Marina Nikitina, Irina Chernukha
V.M. Gorbatov Federal Research Center for Food System of RAS
Moscow, Russia
nikitinama@yandex.ru, imcher@inbox.ru

***Abstract*—A knowledge-oriented system of functional nutrition is described. Information technologies implemented using multi-criteria optimization and mathematical programming methods will allow adjusting and optimizing food rations according to various criteria, as well as structuring the resulting set of alternatives and determining the optimal adequate version of a food ration.**

I. Introduction

According to the World Health Organization (WHO, https://www.who.int/ru), a number of diseases are associated with an insufficiency or excess of certain components in the daily human diet. In Western Europe, 77% of all diseases are non-communicable diseases, while in 86% of cases they are the cause of death. The same trend is observed in Russia.

Maintaining and strengthening human health is impossible without adequate nutrition. Constant violation of dietary regime inevitably leads to pathological changes in vital functions. This is due to the deep impact of nutrition on all biochemical and physiological processes of the body. It is this fundamental influence that underlies the use of diet therapy - therapeutic, functional nutrition - for smoothing, treating and reducing the risk of various diseases. When recommending a particular diet, a dietitian should use not only biochemistry data (protein, carbohydrate, and lipid statuses, immune parameters, and biochemical blood analysis), physiology data (weight deficit, activity factors, and injuries), and food hygiene data (volume, weight, consistency, and temperature of food), but also take into account individual parameters (age, anthropometric, and caliper measure data). The mathematical apparatus is widely used in the analysis of nutrition problems [1-7], as well as principles of food combinatorics when designing combined food products [8-12]. To address the issue of adequate nutrition that meets the needs and capabilities of the human body and is balanced by all indicators of nutritional and biological value, it is necessary to process large amounts of data. In this regard, the issue of therapeutic, functional food can be solved by using a knowledge-oriented system of adequate nutrition, which will help the doctor to quickly and correctly design (select) an adequate diet, taking into account the state of nutrition, individual characteristics, and external factors.

II. Main part

The development of a knowledge-oriented system involves building a knowledge base that contains basic concepts and relations between them. In the case under consideration:

- region of residence;
- activity factor;
- injury factor (operations, bone fractures, traumatic brain injuries, etc.);
- current state of nutrition;
- anthropometric data;
- caliper measure indicators;
- indicators of trophological status;
- indicators of micro- and macronutrient status;
- indicators of vitamin status;
- database on the chemical composition of food products;
- database of biomedical nutritional requirements depending on the disease.

For example, the intake of vitamin B_1 influences carbohydrate metabolism, heart and lung functions; vitamin B_6 content - metabolism of amino acids and fatty acids, nervous system functions; folic acid (folate) – maturation of red blood cells, DNA and RNA synthesis; calcium - blood clotting, nervous and muscular system functions, heart functions; phosphorus - muscle and nervous system functions; magnesium - energy production, acid-base balance, etc. [13].

These data in the knowledge base can be presented as structural and parametric matrices of relations that include three main blocks: patient's physiological state, diet indicators, environmental factors, and many relations within blocks and between blocks and groups of factors.

Each individual block has its own parameters and attributes, for example, the block "Patient's physiological state" includes age, gender, anthropometric indicators – body weight, height, body mass index, shoulder circumference; caliper measure indicators – fat content, thickness of the fold 2 centimeters above the navel; thickness of the fold over the biceps, etc. Nationality and religious traditions are important. Expert estimates, correlation, simple and multiple regression coefficients, influence estimates found as a result of an active experiment, as well as possible functions, relations, and conversion algorithms can be used as formalized characteristics of relations between parameters of structural and parametric characteristics.

General diagnostics of the person's condition and decision-making in the knowledge-oriented system of adequate nutrition is carried out using the following algorithm. At the first stage, 69 physiological indicators of the patient (trophological status indicators, macro- and micronutrient composition) and 5 indicators of external factors (region of residence, religion, nationality, activity, injuries) are entered. At the second stage, the prognostic risk index, main energy exchange and actual energy expenditure are calculated in order to further determine needs of the person (patient) in proteins, fats, carbohydrates, etc. The main diagnostic procedure shall be further conducted. It forms a structural and parametric situational model of an abnormal human condition and searches for deviations from normal values (normative indicators of the institute of nutrition, FAO/WHO, etc.).

A. Identification algorithm

The main mechanism of the diagnostic procedure is a cyclic process of moving the matrix of structural and parametric characteristics along the diagonal (non-diagonal elements reflect relations between the patient's state indicators) with a comparison of values of diagonal elements with the norm level. If an indicator does not match the reference value, values of indicators that it is associated with are tracked. Upon the occurrence of a loop (i.e. repeated repetition in the cycle of the same indicator) all links of this diagonal element are artificially broken, while the above-mentioned procedure is repeated until the next indicator. All abnormal values of indicators are stored in an array of deviations from the norm. Based on its interaction with the knowledge base, the system user is provided with the conclusion of the diagnostic procedure.

The identification algorithm contains (Fig. 1) a block of formation of the situation matrix and the procedure for finding the causes of the anomalous state of the system.

The procedure is a cycle of iteration of independent deviations, within which the maximum element in a line is searched for, its ordinal number *p* is remembered and transition to the *p*-th line takes place followed by a new search for the maximum element of this line [14].

To detect possible looping of cause-effect relationships, an array of t_l indices of diagonal elements included in the interaction trajectory is formed, and when two elements of this array match, it is followed by a "cycle" signal. In this case, the cause may be inside or outside the cycle circuit. To exit the cause-effect cycle and continue to search for the original cause, the last link of feedback is broken, i.e. element $Sgp = 0$, with its value $f_g = S_{gp}$ and addresses memorized in index arrays Ind_{g1}; Ind_{g2}. Then when iterating over the elements of the *g*-th line, the procedure will either stop at the last link of the cycle (if the reason lies within the cycle circuit) or go further through the steps of a new cycle (Fig. 1). When moving to the detection of the causal chain of the next $k+1$-th consequence, the interrupted link of the *j*-th cycle of the previous trajectory of the links is restored, i.e. $S_{Indj1,\ Indj2} = f_j$.

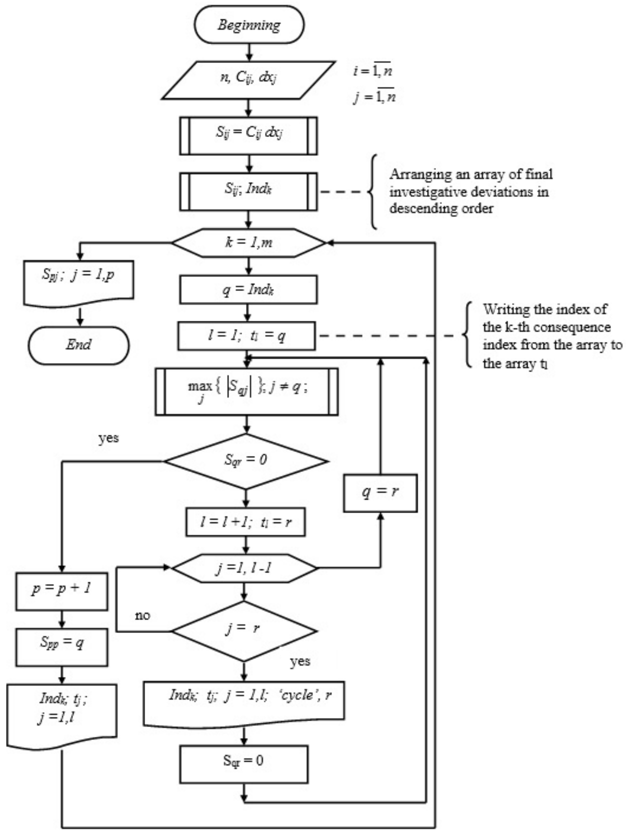

Fig. 1. Block diagram of the algorithm for diagnosing the abnormal state of nutrition system

To find the influence of other factors on the next *k*-th consequence, the first maximum contribution to its deviation is set equal to zero and the next largest element of the *k*-th line is selected, i.e. the next largest contribution to the *k*-th consequence.

All abnormal values of indicators are kept in the array of deviations from the norm and based on the knowledge base a person is provided with the initial selection of products in the recommended diet, which compensates for existing deviations taking into account individual characteristics of the patient and social conditions (personal perception of certain product, presence of allergies, as well as availability of specific products due to material or geographical factors).

If there is insufficient compensation for deviations by selecting the desired products and dishes included in the diet, a search should be made for their optimal quantitative ratios (structural optimization) with the possible introduction of additional products and dishes, depending on the current deviations of parameters from the norms, or the development of an individual combined product that minimizes residual deviations.

B. Issue of multi-criteria selection of adequate nutrition

The development of optimal diets and adequate individual meals with a large degree of uncertainty of parametric descriptions is related to the solution of multi-objective optimization issues [15] by the criteria: food (F1), biological (F2) values, vitamin compliance (F3), mineral compliance (F4), cost (F5), etc. taking into account medical and biological requirements considering individual parameters of the determined group (age, gender, physiological state, digestibility of food products, labor intensity, regional, ethnic and environmental characteristics).

General statement of the issue of multi-criteria selection of an optimal diet has the following form: $a = (a_1, a_2, ..., a_n) \in A$ – a set of alternatives (diets); $F_i(a_j)\, i = \overline{1,m},\ j = \overline{1,n}$ – a set of criteria for evaluating diets.

If the best alternative *(aj)* corresponds to the minimum of the $F_i(a_j)$, criterion, then the search for the optimal product formulation is reduced to minimizing the vector criterion *F(A)* on the set of alternatives *A*:

$$F(A) = \{F_1(a_j), F_2(a_j), ..., F_i(a_j), ..., F_m(a_j)\} \rightarrow min \quad a \in A \tag{1}$$

When comparing two diets $a_i \in A$ and $a_j \in A$, provided $F_i(a_i) \leq F_j(a_j)$ for all $i, j = \overline{1,m}$, it is obvious that the a_i diet is not worse than the a_j diet and, upon the availability of such $i = \overline{1,m}$, that $F_i(a_i) < F_i(a_j)$, then the a_i diet is better than the a_j diet.

If it is impossible to set a preference for a pair of diets, then these diets are called incomparable. A solution is Pareto-optimal if there is no other solution that improves the value of one of the criteria without degrading the other criteria. Since the Pareto-optimal solution may not be the only one, the concept of a Pareto-optimal set of solutions arises as a set of non-dominanted alternatives.

If *WP(A)* is a set of incomparable diets, then there will always be such criterion $F_i(a)\, i = \overline{1,m}$ for $a^* \in A$ and any other $a \in A$ with $F_i(a^*) < F_i(a)$, then it can be argued that diets represent a Pareto-optimal set

$$WP(A) = \{a^* \mid \forall a \in A (\exists i = \overline{1,m} [F_i(a^*) < F_i(a)])\} \tag{2}$$

In this regard, we get a set of alternative solutions for a given diet that correspond to the extreme values of individual criteria with the inability to further improve other criteria without worsening this one.

To evaluate and select the best alternative (diet), it is necessary to structure alternatives according to the adequacy criteria. The set of evaluated alternatives is entered in the "criteria table".

TABLE I. CRITERIAL MATRIX OF ALTERNATIVES

	F_1	F_2	…	F_i	…	F_m
A_1	x_{11}	x_{12}	…	x_{1i}	…	x_{1m}
A_2	x_{21}	x_{22}	…	x_{2i}	…	x_{2m}
…	…	…	…	…	…	…
A_j	x_{j1}	x_{j2}	…	x_{ji}	…	x_{jm}
…	…	…	…	…	…	…
A_n	x_{n1}	x_{n2}	…	x_{ni}	…	x_{nm}

where x_{ij} – evaluation of the *i*-th alternative (diet) a_i by the *j*-th criterion $(F_j)\, i = \overline{1,m};\ j = \overline{1,n}$.

To build a binary preference, an expert in this field shall set the P_i weight for each of the F_i criteria characterizing its importance.

To compare alternatives (diets) a_i and a_j the complete set of criteria is divided into three subsets $\{I_+, I_=, I_-\}$:

$I_+(a_i, a_j)$ – criteria for which a_i diets are superior to a_j diets;

$I_=(a_i, a_j)$ – criteria for which a_i diets and a_j diets are equivalent;

$I_-(a_i, a_j)$ – criteria for which a_i diets are superior to a_j diets.

Further, the relative importance of P_{+ij}, $P_{=ij}$, P_{-ij} of each of these subsets is determined by expert analysis. In addition, a certain threshold *C* is set and it is assumed that the a_i diet is superior to the a_j diet only in the case when the f_s function (concordance index) complies with the condition

$$f_s(P_{+ij}, P_{=ij}, P_{-ij}) \geq C \tag{3}$$

The importance of groups of criteria $I_+(a_i, a_j)$, $I_=(a_i, a_j)$, $I_-(a_i, a_j)$ is defined as the sum of weights of the criteria included in them

$$P^* = \sum p_l, \quad * \in \{+, -, =\}; l \in I^*(i) \tag{4}$$

Additionally, conditions that take into account not only the order of a_i and a_j diets by adequacy criteria $F_i(A)$, but also values of their differences shall be formed

$$d_{ij} = \{F_k(a_i) - F_k(a_j)\}, i \neq j; i = \overline{1,m}; k = \overline{1,n} \tag{5}$$

These conditions, called the discordance index, can be evaluated in the form of $d_{ij} \leq d$, where *d* is the threshold value of the discordance index.

The ratio of preference (R) to the optimality of diets is determined by the condition

$$a_i\ R\ a_j \Leftrightarrow \lfloor f_s(P_{+ij}, P_{=ij}, P_{-ij}) \geq C \rfloor \wedge \lfloor d_{ij} \leq d \rfloor \tag{6}$$

It is proposed not to allocate non-dominant variants of diets, but to expand the subset by selecting a kernel in the original set, all alternatives of which are incomparable, while any variant that is not included in the kernel is dominated by at least one alternative of the kernel. Further decrease in variants can be achieved by setting other more stringent restrictions, such as increasing the threshold value of the concordance index *C* and decreasing the threshold of the discordance index.

C. System's information background

In order to support a healthy lifestyle and maintain health, it is necessary to develop a decision support system for the development and correction of adequate dietary regimes taking into account human metabolism. The authors propose the creation of a computer system of optimization of the current diet and preparation of a new diet for a person (user) based on the principles of adequate nutrition, mathematical methods and information technologies.

The information basis of the system is a database (DB) of products and dishes, being most common and sold in large cities and metropolises. The construction of multidimensional parametric models begins with the creation and filling of a database (DB) of reference information necessary both for constructing parametric and mathematical models of healthy food products and for assessing adequacy.

The database structurally displays physical and chemical parameters of raw materials, products of animal and plant origin, optimization criteria and adequacy assessment, recommendations and norms of food nutrients and energy consumption, ensuring the selection of raw materials that meet the specified requirements.

The structure of the information subsystem has been developed - a database of the main physicochemical characteristics of food products, such as calorie content, vitamin quantitative composition, chemical composition (proteins, fats, carbohydrates, and minerals), fatty acid compositions, and essential amino acids per unit mass of products, which is presented in Fig. 2.

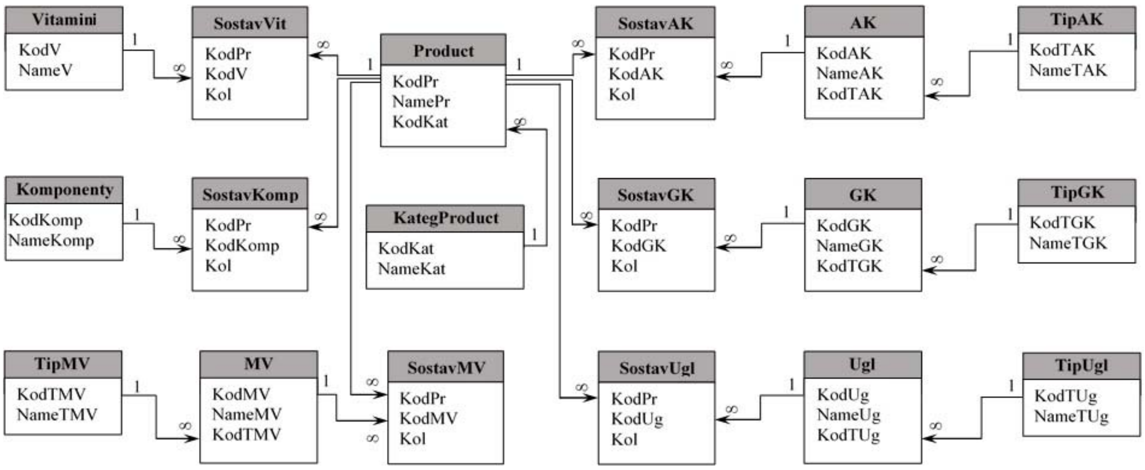

Fig. 2. Logical structure of the database

III. Conclusion

A knowledge-oriented system of adequate nutrition cannot exist without a mechanism for acquiring knowledge. Results of numerous studies and experience of researchers in the development of functional, specialized food products, and nutritionists will be used as the basis for creating a computer knowledge base and information technologies for identifying and predicting body state.

The main purpose of the system is to form structural and parametric models for a specific (individual) person in the field of healthy nutrition, based on the knowledge that it has.

In general, the system should serve as an assistant to a dietitian who is not a "narrow" specialist in this field.

Acknowledgment

The article was published as part of the scientific research theme No. 0585-2019-0008 under the state assignment of the Federal state budgetary scientific institution V.M. Gorbatov Federal Research Center for Food Systems of RAS

References

[1] P.V. Sukhatme, "The World's Hunger and Future Needs in Food Supplies", *Journal of Royal Statistical Society*, Series A, vol. 124, 1961, pp. 463-525.

[2] W. Edwardson, "The Design of Nutritional Food Products for a Developing Country", *A Thesis for the Degree of Ph. D. in Product Development, Massey University*, 1974.

[3] A.M. Anderson and M.D. Earle, "Diet Planning in the Third World by Linear and Goal Programming", *J. Opl. Res. Soc.*, vol.34, 1983, pp. 9-16.

[4] F. Alpaslan, "Türkiye'de 6 Büyük İlde Doğrusal Programlama ile Optimum Beslenme Maliyetinin Minimizasyonu (1994-1997)", *Ondokuz Mayıs University, Fen-Edebiyat Fakültesi Araştırma Fonu. Yayın,* No: F.150, 1996, pp. 6-8.

[5] E. Kaldırım and Z. Köse, "Application of a Multi-objective Genetic Algorithm to the Modified Diet Problem", *Genetic and Evolutionary Computation Congress (GECCO), Undergraduate Student Workshop*. Seattle, USA, 2006.

[6] Y. Lv, "Multi – Objective Nutritional Diet Optimization Based on Quantum Genetic Algorithm"; *in Proc. ICNC,* vol. 4, 2009, pp. 336-340.

[7] S.A. Sahingoz and N. Sanlier, "Compliance with Mediterranean Diet Quality Index (KIDMED) and nutrition knowledge levels in adolescents. A case study from Turkey", *Appetite*, vol. 57, is. 1, Aug. 2011, pp. 272-277.

[8] N.N. Lipatov, "Methods for quantifying and modelling the amino acid balance of meat products". *The 31st European congress of meat scientists*, Sofia, 1985, pp. 158-161.

[9] Yu.A. Ivashkin, S.B. Yudina, M.A. Nikitina and N.G. Azarova, "Information technologies for food design", *Meat Industry*, no. 5, 2000, pp. 40-41.

[10] M.A. Nikitina, I.M. Chernukha and D.E. Nurmukhanbetova, "Principal approaches to design and optimization of a diet for targeted consumer groups", *News of the National Academy of Sciences of the Republic of Kazakhstan. Series of Geology and Engineering*, vol. 1, no. 433, 2019, pp. 231-241

[11] Lv. Youbo, "Multi-objective nutritional diet optimization based on quantum genetic algoritm", *Fifth international conference on natural computation*, 2009, pp. 336-340.

[12] C.-H. Chen, M. Karvela, M. Sohbati, T. Shinawatra and C. Toumazou, "Person – personalized expert recommendation system for optimized nutrition", *IEEE Transactions on biomedical circuits and systems*, vol.12, no. 1, Feb. 2018, pp. 151-160.

[13] *The merck manual of medical information Whitehouse Station*. St.Petersburg: Norint, 2008.

[14] Yu.A. Ivashkin, *System analysis and operations research in applied biotechnology*. Moscow: MGUPB, 2005.

[15] M.A. Nikitina and I.M. Chernukha "Multi-criteria optimization of a product recipe composition", *Theory and practice of meat processing*, vol.3, no. 3, 2018, pp. 89-98.