

Systems biology as a conceptual framework for research in family medicine; use in predicting response to influenza vaccination

Ljiljana Majnarić-Trtica¹ and Branko Vitale²

¹Department of Family Medicine, Medical School Osijek, University of Osijek, Osijek, Croatia

²Institute Ruder Bošković, Zagreb, Croatia

Aim: To introduce systems biology as a conceptual framework for research in family medicine, based on empirical data from a case study on the prediction of influenza vaccination outcomes. This concept is primarily oriented towards planning preventive interventions and includes systematic data recording, a multi-step research protocol and predictive modelling. **Background:** Factors known to affect responses to influenza vaccination include older age, past exposure to influenza viruses, and chronic diseases; however, constructing useful prediction models remains a challenge, because of the need to identify health parameters that are appropriate for general use in modelling patients' responses. **Methods:** The sample consisted of 93 patients aged 50–89 years (median 69), with multiple medical conditions, who were vaccinated against influenza. Literature searches identified potentially predictive health-related parameters, including age, gender, diagnoses of the main chronic ageing diseases, anthropometric measures, and haematological and biochemical tests. By applying data mining algorithms, patterns were identified in the data set. Candidate health parameters, selected in this way, were then combined with information on past influenza virus exposure to build the prediction model using logistic regression. **Findings:** A highly significant prediction model was obtained, indicating that by using a systems biology approach it is possible to answer unresolved complex medical uncertainties. Adopting this systems biology approach can be expected to be useful in identifying the most appropriate target groups for other preventive programmes.

Key words: family medicine; influenza vaccination; outcome; prediction; systems research

Received 25 June 2010; accepted 22 February 2011; first published online 19 May 2011

Introduction

Prevention, based on a person-centred approach, is a core task in family medicine and often involves a complex problem-solving approach (Sweeney *et al.*, 1998; Van Weel and Knottnerus, 1999). As evidence-based medicine draws primarily on

randomised controlled trials investigating specific questions about efficacy in narrowly defined population groups, it does not easily guide this area of decision-making. The challenge is therefore to develop research protocols that can address this task adequately (Rosser, 1999; Van Weel and Rosser, 2004). Mathematical models, by providing prognostic factors, specifically for the target outcome, or a patient group, might be helpful in this sense (Sweeney *et al.*, 1998; Campbell, 2006). Decisions on the choice of parameters for modelling involves researchers taking responsibility for

Correspondence to: Ljiljana Majnarić-Trtica, Department of Family Medicine, Medical School Osijek, University of Osijek, Strossmayerova 105, Osijek, Croatia. Email: ljiljana.majnarić@hi.t-com.hr

knowledge collection and integration. However, for many problems in every day primary care there are not clearly elaborated theories to guide this (Rosser, 1999; Van Weel and Rosser, 2004). This difficulty is even more apparent if one takes into account variation in local working environments, as well as the specific characteristics of local population groups. On the basis of an example of influenza vaccination outcome, we suggest the use of a systems biology methodology, considered as involving both a step-wise research protocol, allowing the research to start with a poorly proved theory, and the use of a systematic record of relevant health parameters.

During recent decades, efforts have been focused on defining markers that can identify individuals who are likely to respond poorly to influenza vaccine. This seemed important because of the conflicting results of published reports on the immune response to influenza vaccine in elderly persons (Webster, 2000). Actually, there has been uncertainty on whether differences in health status, or just in the age of those vaccinated were responsible for differences in immune responses between the elderly and younger people, observed in some, but not all studies. In addition, the effect of different vaccine status and/or differences in past infections with influenza viruses between these two generation groups, have been proposed as potential explanations for the differential response. Now that rapid progress in biotechnology is likely to ensure alternative vaccination approaches it is even more important to answer this question (Tosh and Poland, 2008). However, the single-disease oriented reductionist research methods currently in use are unable to cope with the multifactorial nature of this task (Brydak and Machala, 2000). Attempts to investigate the association of potentially relevant factors with post-vaccination antibody responses, using multivariate analyses are scarce. (Remarque *et al.*, 1996). A critical point in modelling is the choice of parameters used as the input, but in modelling influenza vaccination outcomes, a major difficulty is the wide range of factors related to chronic ageing diseases (Ligthart *et al.*, 1984; Brydak and Machala, 2000). In relation to this, the theoretical background is limited, as immunoregulatory disorders that might account for the deficient immune response to influenza vaccine observed in chronically ill elderly patients have not yet been found

(Gross *et al.*, 1989; Castle, 2000). It has been realised, for example, that differences in stages of a disease, comorbidity, lifestyle factors, or particular biochemical disorders can all contribute to the variation of immune response to influenza vaccine (Wick and Grubeck-Loebenstein, 1997; Brydak and Machala, 2000). To deal with the complexity of this task, we reached out for the concept of a systems biology, originally applied to analyse high-dimensional, non-linear data provided by new sophisticated diagnostic methods, such as genomics and proteomics (Larranaga *et al.*, 2005). From the theoretical background, this concept emerges from the science of complexity.

The science of complexity and a systems biology

According to the science of complexity, to truly understand the functioning of biological organisms, including humans and their diseases, they must be studied as complex systems (Goldberger, 1996). In a complex system, components respond to the environment by using internalised sets of rules that drive the action of the system. In other words, the behaviour of a complex system emerges as an effect of physiological networks.

Although the science of complexity states that biological systems are complex, a systems biology can be defined as a quantitative analysis of how components in the network interact with each other to produce a function, or a phenotype (Kitano, 2002; Iris, 2008). We cannot predict the behaviour of a complex system with certainty; however, we can draw inferences by mathematical modelling. Although, in mathematical terms, a complex system can be determined by a range of numerical parameters, a working mathematical model may not need to include all possible parameters, as only a few of them are likely to control the outcomes of the system.

As based on a multitude of poorly proved parameters, a systems biology approach, unlike the classical, reductionist one, where only a few, recognisable parameters can be evaluated, is not strongly driven by the hypothesis, but is rather based on the use of research protocol (Figure 1). That means that background information for input are drawn out from the literature. The model is created on the basis of these pieces of information and tested by experiment, or a computer-based

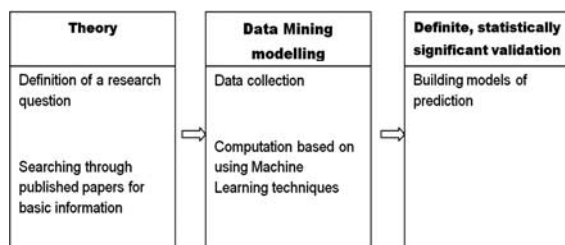


Figure 1 A step-wise research protocol used in the study

simulation (a model-building approach). The results are, in turn, used to make corrections to the model.

An example – the case study on prediction of influenza vaccination outcomes

We applied a systems biology approach to identify health parameters to use in models to predict responses to influenza vaccine. This aimed to support decision-making about vaccination strategies and provide insight into the value of adopting a systems biology approach in exploring complex health problems.

Methods

Population

The examined sample consisted of 93 volunteers, 35 male and 58 female, aged 50–89 years (median 69), out of 150 patients vaccinated against influenza in the season 2003–04, in a family practice located in the town of Osijek, Croatia, in a region with a high prevalence of chronic diseases. The sample was drawn from the high-risk population for influenza complications who often require vaccination, consisting of older patients with multiple medical conditions, primary health care attenders (Center for Disease Control and Prevention, 2007). The study protocol was approved by the local ethics committee.

Influenza vaccination

Trivalent inactivated split vaccine containing A/H1N1/New Caledonia/20/99-like, A/H3H2/Moscow/10/99-like and B/Hong Kong-330/2001-like influenza virus strains, recommended that season, was used. In addition, the influenza vaccine component type B was tested for heterologous

reaction against the virus antigen B/Sicuan 379/99, included in the recent past (Pyhala *et al.*, 1994). In order to measure vaccination outcomes, specific antibody production was determined by the haemagglutination inhibition (HI) test, with the measure for expression of the specific antibody production being taken as at least a fourfold increase (greater than or equal to four times) in antibody titres. Serological measurements were taken in the Department of Virology of The Croatian Public Health Institute, Zagreb.

The study design – a step-wise research protocol

We used the three-step research protocol (Figure 1). In the first step, we searched the MEDLINE/PubMed Journal database and also screened references to find out how chronic ageing diseases and age-related pathogenetic perturbations alter immune system functions. Our assumption was based on the novel theories of ageing stating that the mechanisms of ageing and age-related diseases could be better understood if viewed as networked reactions integrating all levels of the bodily organisation, from the molecular and cellular, to the personal (systems) one. In particular, this suggests that the diverse expression of ageing phenotypes can be explained through the dynamic interplay between three main control systems, (insulin-dependent metabolism, the neuroendocrine, and the immune systems), with chronic inflammation acting as an intermediate mechanism linking changes in these systems together (Franceschi *et al.*, 2000).

On the basis of this background information, in the second step of the protocol, we collected data on various health-related parameters indicating inflammation, nutritional, metabolic and neuroendocrine status, chronic renal impairment, latent infections, and humoral immunity (Pozzetto *et al.*, 1993; Wick and Grubeck-Loebenstein, 1997; Schroecksnadel *et al.*, 2002; Sipponen *et al.*, 2003; Trzonkowski *et al.*, 2003; Clarke *et al.*, 2004). Owing to the large number of input parameters, it was necessary to reduce this number before starting to build a prediction model. Non-linear data mining algorithms were used for this purpose, resulting in the limited pool of selected parameters, potential predictors of antibody responses to influenza vaccine (Figure 1).

To reach the final aim of our research, to develop a statistical model that may accurately predict responses to influenza vaccine, in the third step of the protocol, we combined previously selected candidate health parameters with information on past influenza exposure, using logistic regression to build a prediction model (Figure 1).

A data set – systematic health data record

We determined the health status of examined patients systematically, using 52 clinical parameters indicating age and sex, diagnoses of the main groups of chronic diseases (Table 1), anthropometric measures (indicating the nutritional status; Table 2) and haematological, and biochemical tests (Table 3). In order to be chosen, laboratory tests had to meet two criteria: to reflect the age-related health disorders found to have a negative impact on the immune system functions, and to be available in routine primary health care.

Blood sampling

From each patient, blood samples were collected twice, before and four weeks after vaccination.

Serum samples for HI antibodies were separated on each occasion and kept at -40°C until analysed. At first blood sampling, specimens were also obtained for laboratory tests. Haematological analyses were carried out from fresh blood samples while sera for biochemical analyses were separated and stored until assayed. Laboratory analyses were performed in the Central Biochemical Laboratory of the Osijek Clinical Hospital, using standard techniques.

Data analysis

Selection of health parameters

On the prepared data set, we applied data mining algorithms based on Machine Learning methods to identify age-related health disorders with the potentially largest negative impact on serological responses to influenza vaccine. This was the key step in the research protocol, resulting in the selection of relevant health parameters (Figure 1).

Data mining is a group of robust computational techniques able to extract and interpret information contents (patterns) from massive biomedical databases (Witten and Frank, 2005). In this study, we used algorithms of the ILLM (Inductive

Table 1 Parameters collected: prevalence of the main groups of chronic diseases

Parameters	Positive cases <i>n</i> (%)	Diagnosis making criteria
Hypertension	77 (82.8)	Medical records Blood pressure
Diabetes mellitus	23 (24.7)	Medical records, fasting glucose, OGTT if borderline fasting glucose
Cardiovascular diseases (myocardial infarction, angina, history of revascularisation, stroke, transient ischaemic cerebral event, and peripheral vascular disease)	24 (25.8)	Medical records and clinical examination
Gastroduodenal disorders (gastritis, ulcer)	32 (34.4)	Medical records
Chronic urinary tract disorders (recurrent cystitis in women, symptoms of prostatism in men)	51 (55.4)	Medical records, clinical criteria
Chronic obstructive pulmonary disease	13 (13.9)	Positive spirometry
Allergy (rhinitis, asthma)	9 (9.6)	Symptoms, drug treatment, positive spirometry
Severe osteoarthritis	23 (25.0)	Severe symptoms or frequent use of non-steroidal anti-inflammatory drugs Currently not under active treatment
Malignancy	14 (15.0)	Medical records, DEXA
Osteoporosis	30 (40.0)	Medical records, psychotropic medication, MMSE
Neuropsychiatric disorders (anxiety/depression, Parkinson's disease, and cognitive impairment)	53 (57.0)	Medical records, physical examination
Chronic skin disorders (chronic dermatitis, dermatomycosis)	23 (25.0)	

OGTT = Oral Glucose Tolerance Test; DEXA = dual-energy-X-ray-absorptiometry, the standard to diagnose osteoporosis (in 10 cases the data were missing); MMSE = Mini Mental State Examination Score, standard screening test on cognitive impairment (maximum score 30, <24 indicates positive on dementia).

Table 2 Collected parameters: anthropometric measures

Parameters	Minimum	Maximum
Body mass index (kg/m ²)	20.2	43.1
Underweight		<20
Normal weight		20–25
Overweight		26–30
Obesity		31–40
Severe obesity		>40
Waist-to-hip ratio	0.76	1.1
Male		1.0
Female		0.8
Triceps skinfold thickness (with caliper; mm)	20.2	50.0
50–65 years		
Male		8–23
Female		10–35
Mid-upper arm circumference (cm)	24.0	39.0
Male		54 ± 11
Female		30 ± 7

Table 3 Laboratory tests performed^a

Inflammation: WBC count, WBC differential (% neutrophils, lymphocytes, eosinophils, and monocytes), CRP, and serum proteins electrophoresis (α_1 , α_2 , β , γ -globulins)
Nutritional status: RBC count, haemoglobin, MCV, iron, serum albumin, folic acid, vitamin B ₁₂ , and homocysteine
Metabolic status: Fasting glucose, HbA _{1c} , total cholesterol, HDL-cholesterol and triglycerides
Chronic renal impairment: Creatinine clearance
Latent infections: <i>Helicobacter pylori</i> specific IgA and IgG, and cytomegalovirus specific IgG
Humoral immunity: IgE and ANA
Neuroendocrine status: Blood cortisol in the morning, TSH, fT ₃ , fT ₄ , and prolactin

WBC = white blood cell; CRP = C-reactive protein; RBC = red blood cell; MCV = mean cell volume; HbA_{1c} = glycosylated haemoglobin; HDL = high-density lipoprotein; ANA = antinuclear antibodies; TSH = thyroid-stimulating hormone; fT₃ = free triiodothyronine; fT₄ = free thyroxine.

^aLaboratory tests were performed by using standard techniques. Descriptive statistics were not shown, as being of the minor relevance for the topic.

Learning by Logic Minimization) system, developed in the Laboratory for Information Systems, Institute Rudjer Bošković, Zagreb, because of the availability and good classification and pattern recognition properties of this method (Gamberger and Šmuc, 2001; Gamberger *et al.*, 2003). The result of applying the ILLM algorithms to the prepared data set is a cluster of six parameters, most strongly associated with the target outcome value, with the first parameter on the list ranking most important. Statistical measures ‘sensitivity’ (the accuracy of the true positive results of the classification procedure) and ‘specificity’ (the accuracy of the true negative results of the classification procedure) are used for expression of the statistically significant properties of the parameters selected in the cluster (Gamberger

Primary Health Care Research & Development 2011; **12**: 310–321

and Šmuc, 2001; Gamberger *et al.*, 2003). *As no unique definition of the target outcome value (low antibody response to influenza vaccine) is possible, because influenza vaccines are trivalent and factors related to past influenza viruses exposure strongly affect vaccination outcomes, we set up a maximum number (four) of reasonable definitions (not presented), allowing selection, from the initial dataset, of the four sets of health parameters. In making definitions, we tried to maximally exclude the influence of factors related to past influenza virus exposure, to allow health-related parameters to gain their full effect.*

Generating prediction models

For binary outcomes, such as in this study, cases with good or poor antibody induction after

influenza vaccination, the classical approach is to develop prognostic logistic regression models (Campbell, 2006). We performed full model regression, with all parameters included, and two reduced forms, with the forward and backward parameter selection, using StatSoft, Inc. (2008), STATISTICA (data analysis software system), version 8.0, www.statsoft.com.

A negative outcome for the model (poor response to influenza vaccine) was defined by a positive result of HI test for only one, or none, of three vaccine components (A/H1N1, A/H3N2, and B), whereas a positive outcome (good vaccination response) was defined by positive results of HI tests for two, or all three vaccine components. The model was based on 56 eligible patients.

In the model candidate health parameters, which had been selected from the data set using data mining methods, were combined with information on past influenza virus exposure, including the number of past vaccinations, pre-existing

antibody titres and heterologous reaction, indicated by the vaccine component B/Sicuan. The parameter 'the number of past vaccinations', was expressed as categories: vaccinated for the first time ($n = 37$), previously vaccinated once ($n = 19$), previously vaccinated two to three times ($n = 13$), and four or more previous vaccinations ($n = 24$).

Results

Data mining models

By applying ILLM algorithms to the prepared database four recognisable patterns (clusters) in the data associated with low serological response to influenza vaccination were identified (Table 4). Owing to the partial overlap among the patterns, the initial 24 selected parameters were further reduced to 16 parameters (Table 5).

Within this pool of 16 selected parameters, four of them, ranking best in a cluster, or according to

Table 4 Results of data mining modelling. Patterns of parameters selected

Attribute ranking ^a	Attributes	Cut-off value	Sensitivity %	Specificity %
Model 1				
1	Monocyte %	>8.0%	90.0	70.8
2	Vitamin B ₁₂	≤212.0 pmol/L	80.0	75.0
3	Homocysteine	>12.7 μmol/L	80.0	75.0
4	fT ₄	≤13.6 nmol/L	70.0	79.1
5	Creatinine clearance	≤1.5 mL/s/1.73 m ²	70.0	75.0
6	Skinfold thickness	≥32.5 mm	80.0	62.5
Model 2				
1	Monocyte %	>7.8%	71.4	73.6
2	γ-Globulins	>13.0 g/L	64.2	78.9
3	MCV	>90.5 fL	78.5	63.1
4	<i>Helicobacter pylori</i> IgA	>11.8 IU/mL	78.5	63.1
5	Prolactin	>90.2 mIU/L	85.7	57.8
6	β-Globulins	>8.5 g/L	64.2	73.6
Model 3				
1	Lymphocyte %	≤35.1%	65.6	63.6
2	fT ₄	≤13.6 nmol/L	59.3	68.1
3	Fasting glucose	≤5.4 mol/L	50.0	77.2
4	β-Globulins	≥8.0 g/L	53.1	72.7
5	Monocyte %	>7.9%	65.6	56.8
6	Serum albumin	<45.3 g/L	75.0	54.5
Model 4				
1	Lymphocyte %	≤35.4%	56.7	89.4
2	Monocyte %	>7.9%	59.7	84.2
3	Skinfold thickness	≤34.5 mm	65.6	73.6
4	fT ₄	≤14.5 nmol/L	71.6	63.1
5	Age	>65.5 years	71.6	63.1
6	TSH	>1.3 IU/mL	59.7	68.4

fT₄ = free thyroxine; MCV = mean cell volume; TSH = thyroid-stimulating hormone.

^a Attribute ranking: strength of association with a poor response to vaccination.

Table 5 A pool of 16 selected parameters

Data mining models	Parameters selected in a particular model	Parameters overlapping between two or more models
Model 1	Creatinine clearance Homocysteine <i>Helicobacter pylori</i> IgA	Monocyte %, vitamin B ₁₂ , fT ₄ , triceps skinfold thickness
Model 2	γ -Globulins Prolactin	Monocyte %, MCV (vitamin B ₁₂) β -Globulins
Model 3	Fasting glucose Serum albumin	Monocyte %, lymphocyte %, fT ₄ , β -globulins
Model 4	Age TSH	Monocyte %, lymphocyte %, fT ₄ , Triceps skinfold thickness

fT₄ = free thyroxine; MCV = mean cell volume; TSH = thyroid-stimulating hormone.

the statistical measures ‘sensitivity’ and ‘specificity’, including ‘monocyte %’, ‘lymphocyte %’, ‘vitamin B₁₂’, and ‘homocysteine’, can be especially important (Table 4). On the basis of the existing knowledge, these four parameters are likely to indicate two pairs of disorders, including increased percent of monocytes and decreased percent of lymphocytes in white blood cell (WBC) differential (indicating the switch from the specific to nonspecific immune reaction), and mutually related metabolic disorders, vitamin B₁₂ deficiency, and hyperhomocysteinaemia.

Logistic regression models

Three types of the logistic regression models predicted 76.9% (full model) and 75.8% (both, forward and backward types) of the total results of responses to influenza vaccine ($P = 0.00$, 0.04 and 0.00, respectively; Table 6).

Among parameters included in the model, the one indicating older age was not selected as an independent predictor (Table 6). Factors related to past influenza virus exposure showed the greatest influence, especially the number of past vaccinations (Table 6). In particular, past vaccinations of two to three times are likely to have the beneficial effect (OR 0.06, 95% CI 0.00–0.63; Table 6, full model). Immune reaction to influenza vaccine component from the recent past (heterologous reaction), as in this case study with the B/Sicuan influenza vaccine component, also showed significant, albeit negative effect (OR 1.05, 95% CI 1.00–1.10; Table 6, full model). In contrast to the factors related to past influenza exposure, parameters indicating chronic health

disorders showed only minor effect (negative; Table 6, full model). The results of the forward and backward model types showed that only a few health parameters are sufficient for prediction, including relative lymphopaenia (decreased percent of lymphocytes in WBC differential; OR 0.94, 95% CI 0.88–0.99), vitamin B₁₂ deficiency (OR 0.99, 95% CI 0.99–1.00), and hyperhomocysteinaemia (OR 1.15, 95% CI 0.99–1.32; Table 6).

Discussion

Modelling responses to influenza vaccine

In this study, we have shown that by using a systems biology methodology approach it is possible to identify health parameters that can be used to build useful models to predict responses to influenza vaccine. Good model performances, including the high likelihood level for prediction (indicated by significant P -value) and good predictive accuracy (of 76.9%) support its practical usefulness, although obtained on a small sample (Table 6; Campbell, 2006). In addition, narrow confidence intervals of health parameters imply their mean values as if the sample was large, instead of small (Table 6; Campbell, 2006). The latter characteristic may be due to the process of health parameters being pre-selected using Machine Learning methods. However, there are also some limitations in terms of model’s applicability. As the patients in the sample were recruited from the local area in Croatia where a high prevalence of chronic diseases have been recorded, the model should best be applied in this local population group. To be valid as a practical

Table 6 Results of the linear regression analysis – OR and 95% CI

Parameters included ^a	OR	95% CI	P-value
Full model			
Age			
Fasting glucose	1.07	0.96–1.20	0.20
Triceps skinfold	0.83	0.59–1.17	0.30
Thickness	1.07	0.95–1.20	0.26
Past vaccinations			
Zero			0.08
One time	0.30	0.03–2.49	0.26
Two to three times	0.06	0.01–0.63	0.01
More than or equal to four times	0.13	0.01–1.65	0.11
<i>Helicobacter pylori</i> IgA	1.00	0.98–1.01	0.97
Monocyte %	1.23	0.84–1.79	0.27
Lymphocyte %	0.96	0.88–1.04	0.36
MCV	1.02	0.90–1.16	0.71
Albumin	1.29	0.96–1.73	0.08
Creatinine clearance	2.53	0.29–21.7	0.39
Homocysteine	1.10	0.87–1.39	0.38
β -Globulins	0.65	0.30–1.39	0.26
γ -Globulins	1.33	0.93–1.88	0.11
Vitamin B ₁₂	0.99	0.99–1.00	0.38
Prolactin	1.00	0.99–1.01	0.39
TSH	2.01	0.93–4.36	0.07
fT ₄	0.85	0.62–1.18	0.34
Pre-existed titre			
H1N1	0.97	0.92–1.02	0.25
H3N2	1.00	0.99–1.01	0.49
B/Sicuan	1.05	1.00–1.10	0.04
B/Hong Kong	1.00	0.99–1.01	0.41
Type forward			
Past vaccinations			
Zero			0.01
One time	0.127	0.02–0.64	0.01
Two to three times	0.050	0.01–0.30	0.00
More than or equal to four times	0.105	0.01–0.77	0.03
Lymphocyte %	0.942	0.88–0.99	0.04
Homocysteine	1.150	0.99–1.32	0.05
Type backward			
Past vaccinations			
Zero			0.02
One time	0.18	0.03–0.99	0.05
Two to three times	0.06	0.01–0.36	0.00
More than or equal to four times	0.12	0.01–0.91	0.04
Lymphocyte %	0.94	0.88–1.00	0.05
Vitamin B ₁₂	0.99	0.99–1.00	0.27
B/Sicuan pre-existed titre	1.02	0.99–1.04	0.08

^a MCV = mean cell volume; TSH = thyroid-stimulating hormone; fT₄ = free thyroxine.

screening tool in other settings, the model should be retested with a larger sample, relevant to the settings in which it might be used.

Our results indicate that factors related to past influenza exposure are preferable for prediction, compared with factors related to chronic health disorders (Table 6). The reason the parameter

‘older age’ remained unselected could be its contribution in logistic regression as a confounder, relative to the two other factors, past influenza exposure and chronic health disorders, both known as being age-dependent (Webster, 2000). From a practical point of view, our results also indicate that only a few health parameters, such as those

Primary Health Care Research & Development 2011; **12**: 310–321

indicating B-vitamin deficiency, hyperhomocysteinaemia, and relative lymphopaenia, are needed for the model to achieve a reasonable level of prediction (Table 6, forward and backward model types). Parameters indicating metabolic disorders, B-vitamin deficiency, and hyperhomocysteinaemia, although not proved as causally related with poor antibody responses to influenza vaccine, are likely to provide common mechanisms to link the burden of chronic ageing diseases with lymphopaenia and other age-related immune system dysfunctions.

There are two lines of evidence supporting this assumption. The first, derived from the literature, is that these parameters can serve as markers of decreased turnover of immunocompetent cells and of the switch from the specific to non-specific and cellular immune response (in our results indicated by increased percent of monocytes and decreased percent of lymphocytes in the differential WBC; Fenech *et al.*, 1997). In a broader context, these changes can be considered as markers of an impaired methylation reaction, a biochemical process, which when impaired can manifest as DNA damage, genome instability, impaired cell proliferation and insufficient neurotransmitter synthesis, intermediate mechanisms during the development of chronic ageing diseases (Schroecksnadel *et al.*, 2002).

The second line of evidence, suggesting there may be a causal relationship between key health parameters and the vaccination response that the model predicts, arises from the theoretical background of the complex systems science that the study is based on. Accordingly, identified key components may be reflective of the functional integration of the elements within the common biological network (in this case, linking the burden of chronic ageing diseases with the immune system dysfunction; Kitano, 2002; Iris, 2008). The rest of the pool of selected 16 parameters can also be useful, although in some other situations of vaccinations. However, this statement has yet to be proved, by the application of these results in different situations of vaccinations and on different samples of the defined high-risk population for influenza complications.

Actually, those selected parameters overlapping between two or more data mining models are likely to indicate common intermediate mechanisms linking chronic diseases with the immune system dysfunctions (Table 5). Parameters specifically

selected in particular models are likely to indicate more specific, relatively well defined clinical conditions (clinical domains), very likely associated with poor responsiveness to influenza vaccine. By using past knowledge, these clinical conditions may link with impaired renal function, especially syndrome characterised with hyperhomocysteinaemia (model 1), chronic gastritis, caused by *Helicobacter pylori* infection and accompanied with chronic nonspecific immune reaction (model 2), a syndrome composed of glucose metabolism impairment and protein malnutrition (model 3), and ageing of the hypothalamus and the pituitary gland, accompanied with the neuroendocrine system dysfunction (model 4; Table 5). Although laboratory tests are preferred over diagnoses of chronic diseases, these results, implicating clusters of pathogenetic disorders, can provide physicians with information needed for initial screening of older patients who are potentially at higher risk for deficient responses to influenza vaccine.

Implications of implementing a systems biology methodology approach in research in family medicine

Risk charts and scores have been developed to assess the risk for cardiovascular events. The major risk factors were identified some time ago in large prospective cohort studies, but as evidence accumulates there is a tendency to add new risk factors into revised risk scores (Cooper *et al.*, 2005). Another emerging field, also using simple clinical parameters, is the early detection of type 2 diabetes (Rahman *et al.*, 2008). There is, however, a challenge to know whether a uniform, generally applicable risk assessment tool can be developed, as risk functions depend on the characteristics of the studied population (The European Society on Cardiology and The European Association for the Study of Diabetes Guidelines, 2007). This observation is due to the findings that each population may have a different distribution of risk factors and that the same risk factors may not have the same effect in determining diseases in different population groups. In addition, there are changes in trends over time, as well as the accumulation of new knowledge, which suggest there is a need for a more dynamic and adaptable framework to prepare effective risk scores (Majnarić-Trtica *et al.* (2007; 2010a)). Efforts to develop

simple, practical risk scores in some other clinical disciplines, such as prognosis of particular types of cancer, dementia, or influenza vaccination outcomes, are faced with even more difficulties, since for these medical problems simple clinical and demographic prognostic markers have not yet been defined (Jellema *et al.*, 2010; Landau, 2010).

On the basis of the results of this study, we propose that a systems biology methodology, considered as both, a systematic health parameters record and a step-wise research protocol, allowing research to begin from poorly proved theory – might be that feasible research framework capable of solving many uncertainties when planning preventive interventions in family medicine. This is made possible by recent advances in information technology such as Machine Learning. In addition, due to the possibility of the modelling process of being separately performed for each problem, or data set, research can be specifically tailored to fit the needs of the local environment.

Subsequently, over time, a common health data set, appropriate for modelling of many medical problems and reflecting the specific features of local population groups, can be selected. This is likely to be due to the characteristics of chronic ageing diseases, such as an overlap between risk factors and clinical expression of related diseases, shared pathogenetic mechanisms, as well as the tendency of these diseases to appear in a cluster (several diseases and morbid conditions occurring in the same person; Buchanan *et al.*, 2006). If information from other sources, such as those on family history, socio-economic status, local environment, occupation, or specific genetic traits, are added to the basic health database, various comprehensive conclusions, based on the modelling, can be drawn (Griffiths, 1998; Majnarić-Trtica *et al.*, 2009). This may substantially facilitate preventive programmes implementation.

The increasing complexity of health care implies a need for new research methods (Van Weel and Knottnerus, 1999; Plsek and Greenhalgh, 2001). Decision-making relies on multiple factors, some of which are as yet undefined, suggesting that a systems biology methodology may prove useful. This is particularly appropriate in family medicine, where electronic health records provide the opportunity for data collection and integration by using advanced computer-based

techniques (Majnarić-Trtica *et al.*, 2010b). We propose that a limited number of practices should focus on research and models construction, with computer programmers and mathematicians working as a part of research teams, whereas the remaining practices can be the places where models are tested for their utility in everyday practice.

An ongoing process of models construction and application

Systems biology can be considered as a cyclic, long-term and *ongoing research protocol* (Iris, 2008; Figure 2). Construction of models, especially for problems lacking in evidence, as in this case with influenza vaccination outcome, can only be the first, *exploratory phase* of research where parameters potentially relevant to the problem are mapped in the large, unknown input space. During the second, *training phase* of the protocol, the challenge is to repeat the same procedure of selecting parameters on other samples, to find whether the same factors are extracted when the sample is changed. The next phase of the protocol is *testing the feasibility of models* by means of

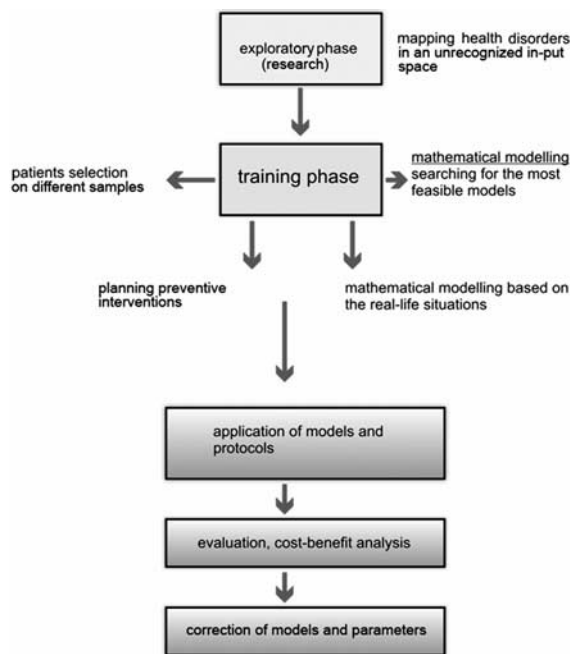


Figure 2 An ongoing computer-based research protocol

their predictive accuracy, cost-effectiveness and the availability of the parameters that make up the model.

The practical usefulness of constructed models, when applied to programmes such as patient selection for immunisation, should be assessed through *the ongoing process of real-life models application*, by means of observational studies (Hannoun *et al.*, 2004). During this process, models and parameters can further be corrected (Figure 2).

Adopting a systems biology approach in research in family medicine can generally be expected to be useful in identifying the most appropriate target groups for preventive interventions implementation.

Acknowledgement

We would like to express our grateful thanks to Dr. Dragan Gamberger, Head of the Laboratory for Information Systems, Institute Ruđer Bošković, Zagreb, for his efforts and time spent in performing Data Mining models, necessary for data processing.

References

- Brydak, L.B. and Machala, M. 2000: Humoral immune response to influenza vaccination in patients from high risk groups. *Drugs* 60, 35–53.
- Buchanan, A.V., Weiss, K.M. and Fullerton, S.M. 2006: Dissecting complex disease: the quest for the Philosopher's stone? *International Journal of Epidemiology* 35, 562–71.
- Campbell, M.J. 2006: *Statistics at square two*, second edition. Oxford: Blackwell Publishing.
- Castle, S.C. 2000: Clinical relevance of age-related immune dysfunction. *Clinical Infectious Diseases* 31, 578–85.
- Center for Disease Control and Prevention (CDC), Immunization Practices Advisory Committee on Immunization Practices. 2007: *MMWR Morb Mortal Wkly Rep* 56, 1–54.
- Clarke, R., Grimley, E.J., Schneede, J., Nexo, E., Bates, C. and Fletcher, A. 2004: Vitamin B12 and folate deficiency in later life. *Age Ageing* 33, 34–41.
- Cooper, J.A., Miller, G.J. and Humphries, S.E. 2005: A comparison of the PROCAM and Framingham point-scoring systems for estimation of individual risk of coronary heart disease in the Second Northwick Park Heart Study. *Atherosclerosis* 181, 93–100.
- Fenech, M.F., Dreosti, I.E. and Rinaldi, J.R. 1997: Folate, vitamin B12, homocysteine status and chromosome damage rate in lymphocytes of older men. *Carcinogenesis* 18, 1329–36.
- Franceschi, C., Valensin, S., Bonafe, M., Paolisso, G., Yashin, A.I., Monti, D. and De Benedictis, G. 2000: The network and the remodeling theories of aging: historical background and new perspectives. *Experimental Gerontology* 35, 879–96.
- Gamberger, D. and Šmuc, T. 2001: *Data mining server*. Zagreb: Institute Ruđer Bošković. Laboratory for Information System. Retrieved from <http://dms.irb.hr/>
- Gamberger, D., Šmuc, T. and Lavrač, N. 2003: Subgroup discovery: on-line data mining server and its application. In Arnez, Z.M., Brebbia, C.A., Solina, F. and Stankowski, V., editors, *Simulations in biomedicine*, fifth International conference. Ljubljana: V.WIT press, 433–42.
- Goldberger, A.L. 1996: Non-linear dynamics for clinicians: chaos theory, fractals and complexity at the bedside. *Lancet* 347, 1312–14.
- Griffiths, F. 1998: General practice and the new science emerging from the theories of 'chaos' and complexity. *British Journal of General Practice* 48, 1697–99.
- Gross, P.A., Gerald, V.Q., Weksler, M.E., Setia, U. and Douglas, R.G. 1989: Relation of chronic disease and immune response to influenza vaccine in the elderly. *Vaccine* 7, 303–308.
- Hannoun, C., Megas, F. and Piercy, J. 2004: Immunogenicity and protective efficacy of influenza vaccination. *Virus Research* 103, 133–38.
- Iris, F. 2008: Biological modeling in the discovery and validation of cognitive dysfunctions biomarkers. In Turck, C.W., editor, *Biomarkers for psychiatric disorders*. New York: Springer, 473–522.
- Jellema, P., van der Windt, D.A.W.M., Bruinvels, D.J., Mallen, C.D., van Wajenberg, S.J.B., Mulder, C.J. and de Vet, H.C.W. 2010: Value and symptoms and additional diagnostic tests for colorectal cancer in primary care: systematic review and meta-analysis. *British Medical Journal*. Retrieved 1 April 2010 from <http://www.bmj.com/content/340/bmj.c1269.full>
- Kitano, H. 2002: Computational systems biology. *Nature* 420, 206–10.
- Landau, S.M., Harvey, D., Madison, C.M., Reiman, E.M., Foster, N.L., Aisen, P.S., Petersen, R.C., Shaw, L.M., Trojanowski, J.Q., Jack, C.R., Weiner, M.W. and Jagust, W.J. 2010: Comparing predictors of conversion and decline in mild cognitive impairment. *Neurology* 75, 230–38.
- Larranaga, P., Calvo, B., Santana, R., Bielza, C., Galdiano, J., Inza, I., Lozano, R.A., Santafe, G., Perez, A. and Robles, V. 2005: Machine learning in bioinformatics. *Briefings in Bioinformatics* 7, 86–112.
- Ligthart, G.J., Corberand, J.X. and Fournier, C. 1984: Admission criteria for immunogerontological studies in man: the SENIEUR protocol. *Mechanisms of Aging & Development* 28, 47–55.
- Majnarić-Trtica, L.J., Vitale, B. and Martinis, M. 2007: Is it time for a new approach in cardiovascular risk assessment? *Periodicum Biologorum* 110, 45–50.

- Majnarić-Trtica, L., Vitale, B., Kovačić, L. and Martinis, M.** 2009: Trends and challenges in preventive medicine in European Union countries. Comment on the state in Croatia. *Periodicum Biologorum* 111, 5–12.
- Majnarić-Trtica, L., Vitale, B., Martinis, M. and Reiner, Ž.** 2010a: A view at the future – a dynamical, protocol-based and computationally intensive approach in cardiovascular risk assessment. *Collegium Antropologicum* 34, 437–45.
- Majnarić-Trtica, L., Zekić-Sušac, M., Šarlija, N. and Vitale, B.** 2010b: Prediction of influenza vaccination outcome by neural networks and logistic regression. *Journal of Biomedical Informatics* 43, 774–81.
- Plek, P.E. and Greenhalgh, T.** 2001: The challenge of complexity in health care. *British Medical Journal* 323, 625–28.
- Pozzetto, B., Odelin, M.F., Bienvenu, J., Defayolle, M. and Aymard, M.** 1993: Is there a relationship between malnutrition, inflammation, and post-vaccinal antibody response to influenza viruses in elderly? *Journal of Medical Virology* 41, 39–43.
- Pyhala, R., Kumpulainen, V., Alanko, S. and Forsten, T.** 1994: HI antibody kinetics in adult volunteers immunized repeatedly with inactivated trivalent influenza vaccine in 1990–1992. *Vaccine* 12, 947–52.
- Rahman, M., Simmons, R.K., Harding, A.-H., Wareham, N.J. and Griffin, S.J.** 2008: A simple risk score identifies individuals at high risk of developing type 2 diabetes: a prospective cohort study. *Family Practice* 25, 191–96.
- Remarque, E.J., Cools, H.J.M., Boere, T.J., van der Klis, R.J., Masurel, N. and Ligthart, G.J.** 1996: Functional disability and antibody response to influenza vaccine in elderly patients in a Dutch nursing home. *British Medical Journal* 312, 1015.
- Rosser, W.W.** 1999: Application of evidence from randomized controlled trials to general practice. *The Lancet* 353, 661–64.
- Schroeksnadel, K., Frick, B., Wirleitner, B., Winkler, C., Schennach, H. and Fuchs, D.** 2002: Moderate hyperhomocysteinemia and immune activation. *Current Pharmaceutical Biotechnology* 5, 107–18.
- Sipponen, P., Laxen, F., Huotari, K. and Harkonen, M.** 2003: Prevalence of low vitamin B12 and high homocysteine in serum in an elderly male population: association with atrophic gastritis and *Helicobacter pylori* infection. *Scandinavian Journal of Gastroenterology* 38, 1209–16.
- Sweeney, K., Mac Auley, D. and Gray, D.P.** 1998: Personal significance: the third dimension. *The Lancet* 351, 134–36.
- The European Society on Cardiology (ESC) and The European Association for the Study of Diabetes (EASD).** 2007: Guidelines on diabetes, pre-diabetes, and cardiovascular diseases: full text. *European Heart Journal* 28, 88–136.
- Tosh, P.K. and Poland, G.A.** 2008: Emerging vaccines for influenza. *Expert Opinion on Emerging Drugs* 13, 21–40.
- Trzonkowski, P., Mysliwska, J., Szmit, E., Wieckiewicz, J., Lukaszuk, K. and Brydak, L.B.** 2003: Association between cytomegalovirus infection, enhanced proinflammatory response and low level of anti-haemagglutinins during the anti-influenza vaccination – an impact of immunosenescence. *Vaccine* 21, 3826–36.
- Van Weel, C. and Knottnerus, J.A.** 1999: Evidence-based interventions and comprehensive treatment. *The Lancet* 353, 916–18.
- Van Weel, C. and Rosser, W.W.** 2004: Improving health care globally: a critical review of the necessity of family medicine research and recommendations to build research capacity. *Annals of Family Medicine* 2, S5–16.
- Webster, R.G.** 2000: Immunity to influenza in the elderly. *Vaccine* 18, 1686–89.
- Wick, G. and Grubeck-Loebenstein, B.** 1997: Primary and secondary alterations of immune reactivity in the elderly: impact of dietary factors and disease. *Immunological Reviews* 160, 171–84.
- Witten, I.H. and Frank, E.** 2005: *Data mining. Practical machine learning tools and techniques*, second edition. San Francisco: Elsevier, Morgan Kaufmann Publishers.