A generic computerized method for the estimate of familial risks
Isabelle Colombet, MD, MPH, Yigang Xu, MS, Marie-Christine Jaulent, PhD, Daniel
Desages, MS, Patrice Degoulet, MD, PhD, and Gilles Chatellier, MD
Public Health and Medical Informatics Department, Broussais Hotel Dieu
University, Paris, France & ERM 202 INSERM
isabelle.colombet@spim.jussieu.fr

ABSTRACT

Most guidelines developed for cancers screening and for cardiovascular risk management use rules to estimate familial risk. These rules are complex, difficult to memorize, and need to collect a complete pedigree. This paper describes a generic computerized method to estimate familial risks and its implementation in an internet-based application.

The program is based on 3 generic models: a model of the family; a model of familial risk; a display model for the pedigree. The model of family allows to represent each member of the family and to construct and display a family tree. The model of familial risk is generic and allows easy update of the program with new diseases or new rules.

It was possible to implement guidelines dealing with breast and colorectal cancer and cardiovascular diseases prevention. First evaluation with general practitioners showed that the program was usable. Impact on quality of familial risk estimate should be more documented.

INTRODUCTION

The familial risk of prevalent chronic diseases, namely breast cancer, colorectal cancer and cardiovascular diseases, is multifactorial. As knowledge on genes involved in genesis of cancer increases (i.e. BRCA1 in breast cancer), genetic screening for multifactorial diseases is becoming more and more common [1]. However, an adequate triage of patients to be referred to genetic screening is crucial to avoid the harm of useless diagnostic test and false positives [2].

Most clinical practice guidelines, for prevention and screening of breast cancer, colorectal cancer and cardiovascular diseases, recommend an evaluation of familial risk [3].

During the typical 15-minutes consultation of a general practitioner (GP), it is almost impossible to record the family history, to draw a family tree with sufficient details and to apply rules allowing the diagnostic of a positive family risk for this disease. For example, in an American guideline for BRCA1

A generic computerized method for the estimate of familial risks
Isabelle Colombet, MD, MPH, Yigang Xu, MS, Marie-Christine Jaulent, PhD, Daniel
Desages, MS, Patrice Degoulet, MD, PhD, and Gilles Chatellier, MD
Public Health and Medical Informatics Department, Broussais Hotel Dieu
University, Paris, France & ERM 202 INSERM
isabelle.colombet@spim.jussieu.fr

ABSTRACT

Most guidelines developed for cancers screening and for cardiovascular risk management use rules to estimate familial risk. These rules are complex, difficult to memorize, and need to collect a complete pedigree. This paper describes a generic computerized method to estimate familial risks and its implementation in an internet-based application.

The program is based on 3 generic models: a model of the family; a model of familial risk; a display model for the pedigree. The model of family allows to represent each member of the family and to construct and display a family tree. The model of familial risk is generic and allows easy update of the program with new diseases or new rules.

It was possible to implement guidelines dealing with breast and colorectal cancer and cardiovascular diseases prevention. First evaluation with general practitioners showed that the program was usable. Impact on quality of familial risk estimate should be more documented.

INTRODUCTION

The familial risk of prevalent chronic diseases, namely breast cancer, colorectal cancer and cardiovascular diseases, is multifactorial. As knowledge on genes involved in genesis of cancer increases (i.e. BRCA1 in breast cancer), genetic screening for multifactorial diseases is becoming more and more common [1]. However, an adequate triage of patients to be referred to genetic screening is crucial to avoid the harm of useless diagnostic test and false positives [2].

Most clinical practice guidelines, for prevention and screening of breast cancer, colorectal cancer and cardiovascular diseases, recommend an evaluation of familial risk [3].

During the typical 15-minutes consultation of a general practitioner (GP), it is almost impossible to record the family history, to draw a family tree with sufficient details and to apply rules allowing the diagnostic of a positive family risk for this disease. For example, in an American guideline for BRCA1

A generic computerized method for the estimate of familial risks
Isabelle Colombet, MD, MPH, Yigang Xu, MS, Marie-Christine Jaulent, PhD, Daniel
Desages, MS, Patrice Degoulet, MD, PhD, and Gilles Chatellier, MD
Public Health and Medical Informatics Department, Broussais Hotel Dieu
University, Paris, France & ERM 202 INSERM
isabelle.colombet@spim.jussieu.fr

ABSTRACT

Most guidelines developed for cancers screening and for cardiovascular risk management use rules to estimate familial risk. These rules are complex, difficult to memorize, and need to collect a complete pedigree. This paper describes a generic computerized method to estimate familial risks and its implementation in an internet-based application.

The program is based on 3 generic models: a model of the family; a model of familial risk; a display model for the pedigree. The model of family allows to represent each member of the family and to construct and display a family tree. The model of familial risk is generic and allows easy update of the program with new diseases or new rules.

It was possible to implement guidelines dealing with breast and colorectal cancer and cardiovascular diseases prevention. First evaluation with general practitioners showed that the program was usable. Impact on quality of familial risk estimate should be more documented.

INTRODUCTION

The familial risk of prevalent chronic diseases, namely breast cancer, colorectal cancer and cardiovascular diseases, is multifactorial. As knowledge on genes involved in genesis of cancer increases (i.e. BRCA1 in breast cancer), genetic screening for multifactorial diseases is becoming more and more common [1]. However, an adequate triage of patients to be referred to genetic screening is crucial to avoid the harm of useless diagnostic test and false positives [2].

Most clinical practice guidelines, for prevention and screening of breast cancer, colorectal cancer and cardiovascular diseases, recommend an evaluation of familial risk [3].

During the typical 15-minutes consultation of a general practitioner (GP), it is almost impossible to record the family history, to draw a family tree with sufficient details and to apply rules allowing the diagnostic of a positive family risk for this disease. For example, in an American guideline for BRCA1
Modeling constraints

Since time is an important constraint for GPs and pedigree taken is not systematic in their daily practice, our main concerns in developing the program were the simplicity of interface and the rapidity to construct a pedigree.

Another concern was the easy update of the software, to allow either the addition of new diseases or updating of the knowledge concerning already described diseases. For this reason, we separated the modeling of the family and the description of the characteristics of the disease from the modeling of familial risk.

Such a software should also be able to improve the quality of the family data collection. One of potential biases is the poor memorization by the patient of his/her family history. To prevent it, asking someone...
about cancer for each member of his/her family, one after the other, is probably more appropriate than asking whether there was any breast cancer in his family. However, the first solution is more time consuming than the second. The solution proposed is a trade-off between these two solutions.

Description of models

The program is based on 3 generic models:
- A model of the family: it models all the relations among the members of a biological family.
- A model of familial risk: it is generic for representing all the rules to estimate the family risk of any disease. The description of the family is compared with the conditions of the rules.
- A display model for the pedigree: it allows to display the structure and the relations of the generated family under a tree format.

Model of the family

The family and its disease history is represented with a package of three classes (PersonModel) (see figure 1):
- class Person: to represent the characteristics of a member of the family, his/her diseases and his/her relatives. The family tree is not specifically represented in our models. The search for a person is completed thanks to the links towards his/her relatives.
- class GenealogyManager: to search for members in the family
- class PersonalATCD: to represent the description of one disease (where “ATCD”, i.e. “antecedent” is for “disease”)

Model of the familial risk

The familial risk for a disease is represented in a package of six classes (ATCDModel) (see figure 1):
- class ATCDManager and class FamilyATCD: to compare the description of diseases in the family with the conditions of rules
- class Rule, class RuleSet, class RuleItem class and class Qualifier: to represent the rules and qualifiers of the disease which are used in the rules (for example, “invasive” for breast cancer). A rule can be either elementary or complex. A complex rule is composed by several rules with logical operators (“AND”, “OR”, “FROM WHICH”).

An elementary rule is evaluated by searching the diseases of family members which fulfill the conditions of the rule. A typical condition such as “At least N first-degree relatives who have the disease X with certain qualifiers”, requires 3 searches:
- What is the target disease and its satisfied qualifiers?
- Who are the relevant people whose disease information can be used to compare with the condition?
- How many of these people satisfy the rule?

The evaluation of a complex rule consist first in evaluation of all elementary rules and second in computation of the results according to the logical operators.

The model of rules is entirely configurable. We use XML as the description format for the configuration of the rules.

Display of the pedigree

Since the tree is not specifically represented in the model of the family, each person is given a description of its position in the entire tree. This method of display allows to use different programming languages (Java, JavaScript, ...). The diseases of a person are represented on the tree thanks to color codes (see figure 2)

RESULTS

Implementation

Our models were applicable to represent all rules specified in cardiovascular guidelines and breast cancer guideline. It was more difficult to represent the rules to define the familial risk of colorectal cancer.

Figure 2: Display of the pedigree

Scenario for use

Building a pedigree

Information to build pedigree is entered for 1st degree and 2nd degree relatives (parents, siblings, uncles and aunts) thanks to 8 questions in one form:
- How many brothers do you have from same parents?
- How many sisters do you have from same parents?
- How many sons do you have?
- How many daughters do you have?
- How many brothers does your father have?
- How many sisters does your father have?
- How many brothers does your mother have?
- How many sisters does your mother have?

Validation of this form gives a standard representation of family tree (figure 4).

Enter familial diseases

One click on any person belonging to the tree opens a dialog box allowing to describe the diseases of the person. Each disease is described by the age of onset, and other pertinent characteristics such as, for breast cancer, presence of an unilateral or bilateral cancer, of an invasive or a non-invasive cancer etc...

Estimating familial risk

Once all family data are stored, the function “estimate familial risk” trigger the comparison of diseases described in the pedigree to conditions described in the rules. A binary result is presented (presence or absence of familial risk) along with a link on some explanation of results, generated from the conditions and rules used and verified during the evaluation process.

The software is on free access on the Internet [5], as a part of a larger project called EsPeR (Personalized estimation of risks).

Qualitative evaluation

A qualitative evaluation of this program was performed with a group of 20 GPs, who used the program to work on a set of 12 structured clinical cases. They pointed us several issues with the usability of the program:

- For a trained user, it takes less than one minute to build the pedigree, enter the data for one disease and estimate the familial risk for this disease. However, the number of clicks and of windows to be opened is too large, and may imply re-learning for somebody who does not use regularly the program.
- When the program determines there is a familial risk, the explanation of rules could be more user-friendly. Conversely, the absence of familial risk should be explained in more details.
- A link to the source of knowledge is mandatory to increase acceptability of the results (i.e.: it is often difficult for a GP to accept that one case of cancer in a family does not constitute a familial risk).

DISCUSSION

We designed a program to help clinicians identify subjects with high familial risks of breast cancer, colorectal cancer and cardiovascular diseases, based on French national guidelines. For cancer, a high familial risk entails a specific genetic advice or screening. For cardiovascular diseases, familial risk contributes to the personal cardiovascular risk and is necessary to guide prevention strategies.

The objectives and functionalities of our program are partly similar to those of the RAGs program developed by Emery et al. in connection to the PROforma risk assessment system [6-8], although its strengths and limitations are different.

Our program is designed so that updates are easy to perform. The generic model allows to add diseases or rules without any important change in the original source code. We integrated it for use in a decision support system which implements a model of guideline adapted from GLIF [9]. However, it has been developed independently from this model and can be integrated with any other guideline-based decision support system. It can also be used as an autonomous program over the Internet.

An interesting asset of the RAG program is its risk assessment and communication capability. Our program outputs a YES/NO answer about familial risk of disease, whereas the concept of risk can be more complex. In fact, we follow the content of our sources of knowledge, the ANAES guidelines, which were developed according to an evidence-based methodology. By contrast, the RAGs software is based on guidelines that were not professionally validated, as stated by their author himself [10].

Several limitations of tools designed to evaluate familial risk actually pertain to the limitations of underlying knowledge. For example, we do not know how to weigh the risk according to the number of siblings in the family. Moreover, we had some difficulties to represent the rules which define the familial risk of colorectal cancer. Several reasons can explain these difficulties: 1) our source of knowledge was not a formal guideline but a consensus conference, 2) there are three types of familial risk of colorectal cancer: autosomal dominant in case of familial polyadenomatosis, probably monogenic in hereditary non polyposis colorectal cancer, and multifactorial in other familial cases. Our program was not designed to represent the mendelian transmission of a given disease, such as many genetic commercial programs [11, 12]. Our goal was to help physicians to identify patients with familial risk, in the most frequent cases, namely multifactorial familial risk.
In daily practice, such programs could have an impact to increase the appropriate identification of patients at highest risk of cardiovascular disease and cancer. Indeed, as stated by Sweet et al. this identification is presently poor because of poor data collection and inappropriate risk assessment by medical staff [13]. This impact is even more probable if the program is used by the clinician at the point of care [14]. Results of our qualitative evaluation confirms that simplicity of use is crucial for such programs. However, simplicity should not prevail over easy access to the underlying knowledge: in our case, this access is a pre-requisite for the acceptability of the program. Further evaluation is needed to measure the extend to which our system improves data collection and risk assessment and thereby influences clinical decisions. Finally, computer systems which implement familial risk assessment rules could also have a methodological impact on the development of the sources of knowledge. Experience drawn from development of such systems could help generate a formal framework to explicit familial history in guidelines [15].

Acknowledgements

We thank Prof. J. Menard (Paris, France) and the ESPER group for their support and criticisms.

The development of the EsPer project was possible thanks to a grant from the “Caisse Nationale des Professions Indépendantes”, Paris (CANAM).

REFERENCES

5. ESPER (Personalized Estimation of Risks) project,
   http://www.hegp.bhdc.jussieu.fr/esper/