Modeling Commercial Potential of Innovative Projects

KATARZYNA ROSTEK
Warsaw University of Technology Faculty of Management
Ul. Narbutta 85, 02-524 Warsaw, Poland
Email: k.rostek@wz.pw.edu.pl
Tel no: +48 22 234 85 10

Abstract

Purpose of the work is to develop effective assessment method of commercial potential of innovative projects, which will increase the clarity, intelligibility and provide repeatability of the evaluation process during applying for a funding. First was performed the analysis of the principles of financing innovative projects and identified the criteria for assessment of their commercial potential. Next was conducted quantitative research on a group of 150 SME sector dental clinics. Based on the collected data and information was developed the reference model for the evaluation of commercial potential of the research group. This model was subjected to regression analysis in order to identify the key factors of success / failure in the evaluation of commercial potential. All activities were included in the authors method of assessing the commercial potential of innovative projects. Based on conducted research should be stated that multi-criteria analysis and data mining methods can effectively assist selection of criteria and automating assessment of commercial potential of innovative projects. The result of their uses is optimizing the range of requirements addressed to applicants and also transparency, comprehensibility and reproducibility of assessments of applications. The possibility of automating the evaluation process leads to significant savings in time and contribution own labor of experts. The value of this work is the assessment model and assessment method of commercial potential of innovative projects, which can be used in competitions, grants, and other situations, where are many organizations applying for limited funds.

Key Words: *Commercial potential, innovative projects, method of assessing the commercial potential, multi-criteria analysis, data mining methods.*

Introduction

The key to innovative project evaluation is the incubation stage (Jolly, 1997; Midgley, 2009) when commercial viability of the proposed solutions is reviewed. The objective is to ensure that the idea behind the project and its conceptualization are attractive enough to obtain funding. There is a wide range of public and private funding programs both in Poland and internationally dedicated to innovative fundamental or applied research projects in a range of disciplines which are available to different types of beneficiaries such as research and business organizations, young researchers, doctoral and post doctoral students. Such funding sources can support the creation of R&D centers and / or R&D teams, the implementation of new business concepts or technologies into the market etc. One of the major challenges is the fact that each funding source and even each call for proposals apply their own criteria and methods for assessing the commercial potential of projects.
The structure and the grading scale used in the selection of innovation projects are subjective. This means that experts reviewing projects are limited by predetermined criteria and weights imposed by the sponsor. It should be noted that this inaccurate system and/or the limited number of review criteria result in strongly biased project reviews. Likewise, there are issues with the transparency of review criteria. Applicants do not usually know the project review criteria or their relative weights. The lack of knowledge and experience

in this field usually causes projects of high commercial potential to lose against substantially weaker competitors, provided the applicant is sufficiently experienced or has received assistance in preparing the grant application. The research conducted in this area focuses on key success/failure factors of innovative products in the process of commercialization and on good practice in innovative project management. This however does not increase the ability to raise funds by the applicants in the competition proceedings.

This implies the need to analyze the available methods allowing a more objective assessment of innovative projects, in particular formal procedures for the selection of significant assessment criteria and methods of multistage and multifaceted multi-criteria decision making. The goal of this analysis is to assess whether the innovative project assessment can be made more objective, less time consuming and more cost efficient. In this context, the following research hypothesis can be formulated:

RH: The implementation of a model of innovative project selection criteria and assessment algorithms will result in increased objectivity of assessment and significantly reduced cost.

The application of such a model at the project selection stage allows an automatic assessment by a suitably trained officer provided that project evaluation criteria are established in advance by experts in the field. This significantly reduces the time and cost of the process due to the limited involvement of experts in the identification of the most significant review criteria and their weights. In this context, an objective assessment method is sought, one which would apply a range of known and understood comparative criteria to allow a fair assessment of innovative projects and help select those with the greatest commercial potential.

In light of the above the paper presents an analysis of financing principles with regards to innovative projects and the key criteria for the evaluation of these projects (Section 2). The results of this analysis are included in proposed method of projects assessment and in the theoretical assessment model (Section 3). The utility of the method and model has been tested and the results are presented in by Case Study in Section 4. The paper ends with the conclusions of the research project and offers recommendations for follow-up (Section 5).

Research Questions within the Theoretical Framework

The definition of commercialization is converting or moving a technology into a profit-making competitive position (Siegel, et. al., 1995). It is a complex process that requires many complex management skills (Loftus, et al., 1994; Peter, 1999; Kathleen, 2003) and the ability to overcome a wide range of barriers (Bandarian, 2007). That is why commercialization of innovation is inherent to the costs that must be incurred in order to achieve and implement the project of innovation. This is often required in order to obtain additional sources of financing.

The award of financial support is dependent on the fulfillment of three conditions (C1-C3):

- C1 - Finding a potential funding source;
- C2 - Meeting all formal and substantive requirements set by the sponsor;
- C3 - Scoring higher than other competing project proposals.

With regards to condition C1, an important role is played by information sources such as the Internet, foundations, and national contact points for Research and Innovation Programs of the European Commission, relevant Government ministries etc. For example, this private website offers a wealth of information on the subject: http://www.fundsnetservices.com. This website was founded in 1996 in order to ensure access to free and comprehensive information on research funding and the development of science for non-profit organizations, universities and the research community.

Funding can be obtained from public and private sources. Making a choice entails the need to comply with condition C2, i.e. meeting all formal and substantive requirements of the donor. Among other things, it requires a detailed knowledge of the procedures and methods of assessment. However, applicants may find this type of information very difficult or even impossible to obtain. Typically, projects are turned down where there is little or no applicant's experience in writing grant applications and competing against those who have such skills in house or have used consultants is a major challenge. Hence, meeting condition C3 is unlikely where in fact the true competition be between innovative and less innovative projects with or without commercial viability. This raises the following research question (RQ1):

RQ1: What methods could be used to increase the transparency and objectivity of assessing innovative projects?

Question RQ1 is all the more important given that contemporary literature on innovative project management is often criticized for highlighting the importance of project selection without giving specific patterns of practical and effective implementation (Åstebro, 2003). Cooper (Cooper, & Edgget, 2010) characterizes this problem quite bluntly: *How to do project right without knowing which are the right projects?* Answers to this question should be considered within the decision making theory.

Decision making is a cognitive process of choosing alternative solutions to a specific problem. For each decision problem there exists at least one optimal decision, where it can be objectively determined whether there is no other better decision while maintaining neutrality in the decision process (Petersen, 2009). However, the problem consists in the choice of the best alternative which best meets the set of objectives identified with multiple criteria. This is a multi-criteria decision problem in which the limiting criteria can be mutually contradictory (Keeney, et. al., 1993; Belton, et. al., 2002). Thus improving the assessment of one of them can cause the deterioration of one or many others (Figueira, et. al., 2005). Then selection decision is a search for such an alternative which would be acceptable according to the sum total of all assessment criteria. For this purpose, it is necessary that the multi-criteria optimization method should be used.

The decision problem interpreted as a multi-criteria optimization task can be presented in the form of an equation (f1) allowing for searching to be performed in the set of acceptable decisions of one whose value reaches the proposed extreme (maximum or minimum) (Brinkhuis, et. al., 2005; Ehrgott, 2005):

$$f(x^*) = \text{extr}\{f(x): x \in X_0\} \quad \text{(f1)}$$

where:

$f(x)$ – the function of assessment criterion assigning value of assessment to the decision variable,
x – decision variable,
X_0 – the set of allowable decisions,
extr – the proposed extreme (maximum or minimum) of the function of evaluation criterion.

In this context, the assessment of the commercial potential of an innovative project, defined as a multi-criteria problem in a decision-making process, is dependent on the optimal choice of three basic elements:

- A set of evaluation criteria – X_0 - understood as a set of characteristics against which the project is assessed;
- Accorded evaluation criteria - $f(x)$ - understood as a collection of point and weight assessments, awarded by the examiner when evaluating the application;
- Function of aggregate evaluation - $f(x^*)$ - synthesizing the assessment criteria in a common aggregate assessment as a basis for ranking and selection between competing projects.

The set of assessment criteria in a given project is determined arbitrarily by the grant committee or experts depending on the specific innovation project at hand and the specific call for proposals. These features define the rules and procedures of selecting projects. They are contained in grant regulations. If the process of commercialization makes economic sense its implementation should be assessed in terms of the value of synthetic financial indicators such as the return on investment (ROI) or the internal rate of return (IRR). However, the assessment of economic impacts does not have to be decisive in the case of research and development projects. There are also other key assessment criteria for innovation projects such as:

- Project objective: its originality, significance and expected contribution to science, completeness of the overview of literature, adequacy of applied theoretical approach and the use of research methods, interdisciplinary approach, requirements vs. actual competence of the research team, expected effects and impact on the scientific and academic environment;
- Feasibility and scientific importance of the project: likelihood of effective and timely implementation of research objectives, adequacy of requested budget and justification for proposed expenditures, funding and in-kind contribution in kind from other sources, degree of exchange and dissemination of knowledge and experience within or outside scientific and academic environment, project timeframe;
- Probability of success: characteristics of scientific profiles of research team members: their research experience, team skills and team management, their publishing achievements, and practical skills and experience.

Nowadays, the selection of specific criteria and rules of assessment depend solely on the authors of competition / tender. In this approach, there is no mechanism to trace cause and effect relationships, linking values of arbitrarily fixed sets of assessment criteria with values of attributes ensuring real competitiveness of this application. This may result in bias both in terms of rating/scoring and the final selection of submitted project proposals. Therefore, it appears justifiable to search for methods of multi-criteria analysis to ensure an objective assessment of a single application and adequate selection of assessment criteria. This leads to another research question (RQ2):

RQ2: What model should be implemented to ensure high relevance of the selection of innovative projects while reducing process costs?

The work presented hitherto in this area addresses the assessment of the commercial potential in terms of technology, market and legal aspects, i.e. the areas where the project may be seen as innovative (Åstebro, 2003; Bandarian, 2007; Galbraith, et. al., 2012). This is a normal procedure with innovative projects or prototyping but it is very challenging organizationally and financially when it has to be applied a grant contest. That is why the main objective of this present project to propose a model of assessing commercial potential of innovation. The objective of such a model is to ensure transparency of the assessment of research projects both for examiners and projects promoters. Further, the model seeks to emphasize the importance of innovation in submitted projects and simplify, shorten and reduce the cost of the assessment process.

Theoretical Methods and Model

The analysis of the assessment of innovative R&D projects has been the subject of research for quite a long time. Researchers have studied the correlation between project characteristics and their success in the framework of empirical models based on statistical methods (e.g. (Cooper, 1981)), evaluation models based on operations research principles (e.g. (Souder, 1973)) or financial parameters based on the discounted cash flow (DCF) analysis (e.g. (Balachandra, et. al., 1997)). This analysis is made in three sections of factors: leading to success (e. g. (Yoon, et. al. 1985; Lester, 1998), (Goldenberg et al., 2001),), leading to failure (e. g. (Hopkins, 1981; Rackham 1998)) and separating success from failure (e. g. (Maidique, et. al., 1984; Cooper, 1985; Abratt, et. al., 1993; Van der Panne, et. al., 2003)). These authors focus on extracting key success/failure factors of new technologies or products in the market. Inspired by the earlier research, this paper proposes a method and a model for assessing the commercial potential of projects and will as vehicle for objective rating and selection.

With regard to question RQ1 methods are sought which will support the impact assessment of the values of various criteria in terms of the overall assessment of innovation in the project and help select among them only those whose meaning in the analysis is dominant. The goal is to make the assessment method used against innovative projects much more objective by reducing the number of assessment criteria and focusing on the most important characteristics. Such methods are used in data mining to analyze case classification (identifying characteristics of defined classification groups), clustering (identifying characteristics of identified groups of similar cases) or discriminant analysis (identifying distinguishing characteristics of individual groups of similar cases). This process allows a qualitative selection of results according to their utility (degree of significance with regard to the main objective of analysis) and value (level of impact on the change in the final value of the analyzed subject). Thus, the criteria used in the assessment of projects and their values can be tested. In the case at hand a method is sought which will strongly determine the cause and effect relationships between the analyzed criteria. Regression analysis may serve this purpose.

Regression analysis allows studying the impact of independent (explanatory) variables on the dependent (response) variable, which is the aim of this research project. It estimates such parameters of a theoretical equation so as to reflect the value, force and vector of such impact as accurately as possible. Regression analysis helps answer the following questions:

- What is the expected value of a dependent variable for a given configuration of the explanatory variables?
- How strong is the influence of explanatory variables on response variable?
- How much will the value of dependent variable change when the values of individual independent variables l change assuming that other explanatory variables do not change?
- how will the value of dependent variable change in next study period if the value of explanatory variable is X?

The mentioned above properties in many literature and practical examples (Nibler, 1997; Jiang-Liang, et. al., 2006; Asli 2007; Shih-Ming, et. al., 2009) confirm the relevance of regression analysis for the selection of arbitrarily adopted criteria and their evaluation values for such which actually determine the success or failure of the innovation project. Depending on the number of explanatory variables and the type of the dependent variable, there are three basic models of regression (Draper, et. al., 1966; Myers 2000): linear, multiple (or multivariate) and logistic. The value of the linear regression function is determined by the formula:

$$Y = f(x) = \beta_1 X + \beta_0 + \varepsilon \tag{3a}$$

where:

Y – dependent variable (response variable),
X – independent variable (explanatory variable),
f(x) – value of regression function,
β_1, β_0 – regression coefficients,
ε - random component of regression function.

Regression coefficients β_1 and β_0 are unknown. Their values are estimated based on the analyzed research sample of observations (x_i, y_i) for i = 1, 2,..., n. Estimating parameters of regression equation is conducted using the method of least squares, which seeks to ensure that the sum of squares of differences between observed values of dependent variable and its theoretical values designated by the model will be as low as possible. The values of these coefficients can be determined from the following formulas (Kim, et. al., 1981):

$$\beta_0 = \bar{Y} - \beta_1 \bar{X} \quad (3b)$$

$$\beta_1 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^{n}(X_i - \bar{X})^2} \quad (3c)$$

where:

Y_i – value of the dependent variable of the i-th observation from the research sample,
X_i – value of the independent variable of the i-th observation from the research sample,
$\bar{Y}$ – mean value of dependent variable for the research sample,
$\bar{X}$ – mean value of independent variable for the research sample.

Random component ε represents random interference in a functional relationship between values of the dependent and independent variable. This component expresses the impact of all existing factors, which besides X variable may affect Y variable. It is linked to a lack of full matching analytical form of regression function to actual connections between analyzed variables.

However, in the case of selecting assessment criteria for innovative projects not one but many explanatory variables should be considered. Then the function of multiple regression applies because it takes into account the effect of at least two explanatory variables on the response variable, which can be expressed by the following formula:

$$Y = f(x) = \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_n X_n + \beta_0 + \varepsilon \quad (3d)$$

Additionally, if the dependent variable is dichotomous a logistic regressions function can be used. The value of regression function in this case is a measure of the likelihood of achieving one of two possible states by the response variable. The value of the logistic regression function can be expressed by the following formula:

$$Y = f(x) \sim B(1, p) \quad (3e)$$

where:

B(1,p) - binomial distribution with probability of success p, where:

$$p = \frac{e^z}{1+e^z} \quad (3f)$$

$$z = \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_n X_n + \beta_0 \quad (3g)$$

In the logistic regression, the probability measure of success / failure is called odds ratio and is expressed by the formula:

$$odds(p) = \frac{p}{1-p} \quad (3h)$$

The odds ratio is the basis of probability distribution for success p in the logistic regression function, possible to estimate with following features:

- logit – expressed as a cumulative distribution function of logarithmic distribution:

$$logit(p) = \ln\frac{p}{1-p} = \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_n X_n + \beta_0 \quad (3i)$$

- probit – expressed as a standardized normal cumulative distribution function:

$$probit(p) = \frac{1}{\varphi(p)} = \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_n X_n + \beta_0 \quad (3j)$$

where φ - standardized normal cumulative distribution,

- loglog – expressed as s cumulative distribution function of double logarithmic distribution:

$$loglog(p) = \ln(-\ln(1-p)) = = \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_n X_n + \beta_0 \quad (3k)$$

The diversity of regression functions allows choosing the option that best fits to the conditions of making the assessment. This method allows selecting these criteria that actually determine the competitiveness of the assessed innovation project.

This means that based on the factors identified in the adopted structure and the relationships linking them it is possible to determine the vector assessment of the project. Thus, it is possible to perform a sensitivity analysis to assess the impact of changes in the value of individual factors on the efficiency of the innovation project. In the light of the above and with reference to RQ2 a commercial potential assessment model can be proposed along with a method of creating such a model.

The adopted method (Fig. 1) aims to develop such a model to assess the commercial potential in an unbiased and partially automated assessment process that warrants transparency and legibility to applicants.. This model should take into account any changes to evaluation criteria that may be adopted over time. Given the growing base of assessed projects the body of experience should be proactively used by grant committee examiners. The proposed approach should rely on data mining to provide for the capacity to easily adapt to the changing conditions of project implementation.

The characteristics of each stage of the proposed method are presented in Fig. 1. The pattern of project assessment evaluation is developed in step F1. It is based on regulations, rules of recruitment and assessment enforced by the contest sponsor. Step F2 is to identify the method of aggregation of step F1 results, expressed in a synthetic summarizing assessment. At this stage, the use of properly chosen multi-criteria analysis method is recommended.

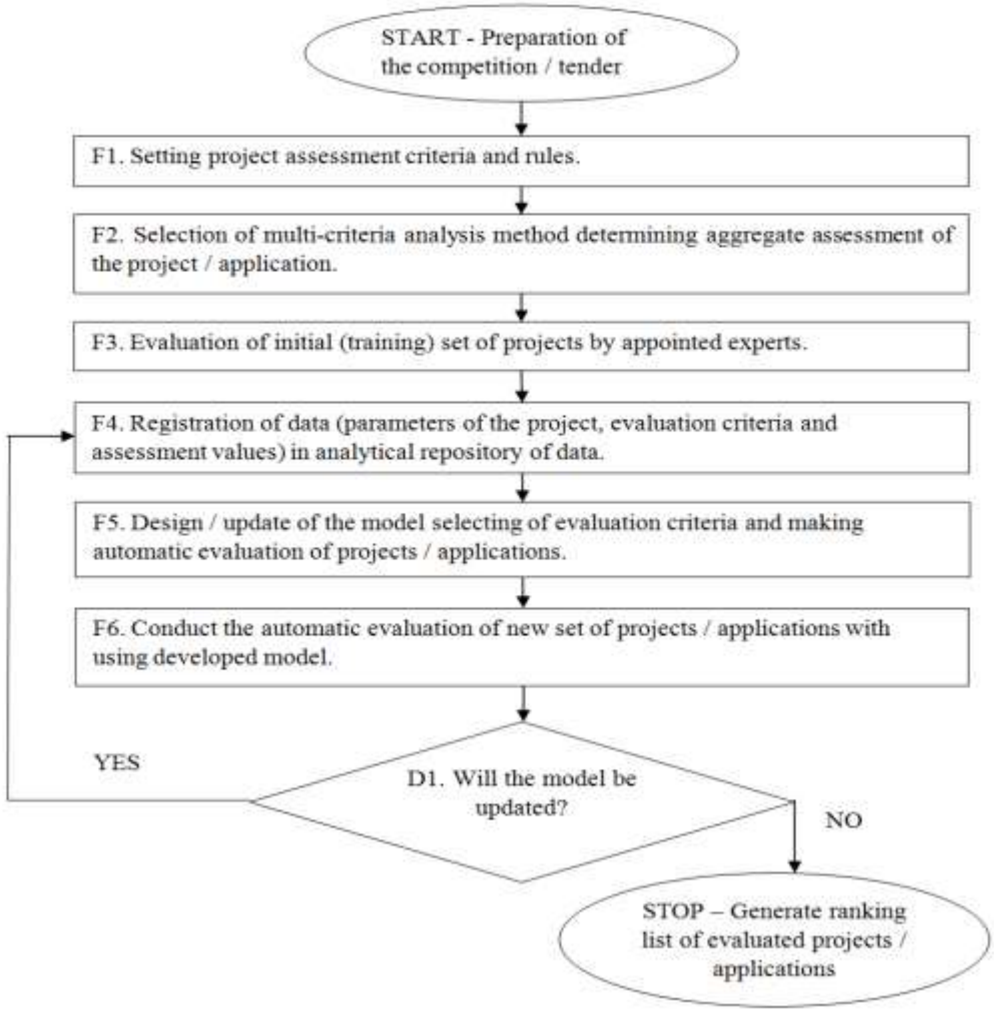

Figure 1 The method of modeling and assessment of the commercial potential of innovative projects (Source: own research)

The development of a data mining model is possible on the basis of an initial (called as training) set of projects which should be evaluated in step F3 by a designated group of experts, as agreed upon in stages F1 and F2. Details of projects and their evaluations determined in step F3, are recorded in an analytical data repository in step F4. This repository is a source of data that populate the automatic project assessment engine.

Step F5 is used for the construction of a data mining procedure for the selection of key assessment criteria (among those assessed by the expert) and the automatic assessment of a new set of projects. The final result in step F6 is a ranking list of assessed projects

The data mining model may be periodically updated along with the updates of the data set on projects and their assessments to ensure alignment with the changing socio-economic conditions of the assessment. This is of great importance for the continuity, consistency and transparency of the assessment and selection.

Case Study

The proposed method and a prototype version of the model were tested on a body of data collected in a research project of dental clinics of the SME sector [1]. The study was designed to create a reference model of the factors of competitiveness for the selected group of clinics to effectively support the development of their competitive positions (Rostek 2012). The assumption adopted for this project is that the data supplied by clinics will form the criteria of assessing their project proposals and the competitive position will be the measure of such assessment. A set of 70 grant applications to support the development of competitiveness filed by SME dental clinics were examined. The applicants:

- offer dental or mixed medical services;
- employ dentists;
- are located only in major Polish cities: Gdansk, Gdynia, Katowice, Krakow, Lublin, Lodz, Poznan, Sopot, Warsaw and Wroclaw;
- are SMEs, i.e.:
  - they employ 2 to 250 staff;
  - their annual revenue does not exceed PLN 210 million.

The goal of this case study is to examine an initial set of 50 grant applications to help maintain competitiveness (hereinafter referred to as the training set) in order to develop a data mining model which will allow an automatic assessment of the remaining 20 applications in the sample. In reality, all 70 applications have already been assessed by experts, so it will be possible to verify the effectiveness of the proposed method and the developed model.

Determining a Set of Assessment Criteria and Designating an Aggregate Assessment of the Application

The selection of assessment criteria is key to optimizing the process of project/application assessment. It should be based both on the experience of a dedicated group of experts and the experience gained in other contests. While assessing applications aiming at identifying the competitive position of SMEs the focus was on factors that determine the competitive position of the clinics. Desk research was conducted to identify key competitiveness factors (study reports of Pentor [2] and PKPP Lewiatan[3]) and a separate survey was completed in the SME sector (questionnaire filled by 150 dental clinics).

Taking into account the results of the above analysis it was found that competitiveness factors of medical clinics can be considered in three major areas of measurable effects: E_1 - *modernity and quality of medical*

[1] Scientific research financed by funds for science in 2009-2011 as a research project No. 0078/B/H03/2009/37.

[2] Report on the Condition of Small and Medium-sized Enterprises in Poland in : 2005–2006, 2006-2007; 2007-2008; 2008-2009. PARP, Warsaw.

[3] The Competitiveness of the SME Sector. Research Report. Lewiatan, Warsaw, 2005; 2006; 2007; 2008.

services, E_2 - *ability to meet the needs of patients*, E_3 - *sales performance*. Each element of the set of measurable effects: $E_i = \{E_1\ E_{2,}\ E_3\}$ was assigned a characteristic set of competitiveness factors: $C_{ij} = \{C_{11}, C_{12}, C_{21}, C_{22}, C_{31}, C_{32}, C_{33}, C_{34}\}$. Then for each competitiveness factor a set of competitiveness measures were assigned: $M_{ijk} = \{M_{111}, M_{112}, M_{121}, M_{122}, M_{123}, M_{211}, M_{212}, M_{221}, M_{222}, M_{223}, M_{224}, M_{311}, M_{312}, M_{313}, M_{321}, M_{322}, M_{323}, M_{324}, M_{325}, M_{331}, M_{332}, M_{341}, M_{342}, M_{343}\}$. Thus, hierarchical structure of competitiveness factors was developed for the dental clinics at hand (tab. 1) and it was used for the assessment of the grant applications

Table1 The structure of competitiveness criteria in a selected group of dental clinics (Source: own research)

Symbol of measure	Description of measure	Symbol of factor	Description of factor	Symbol of area	Description of area
M_{11}	*sales volume of innovative medical services as % of services sales volume*	C_1	*technological level*	E_1	*modernity and quality of medical services*
M_{12}	*investment and development expenditure as % of services sales volume*				
M_{21}	*the number of complaints as % of the number of provided medical services*	C_2	*quality of services*		
M_{22}	*the value of complaints as % of services sales volume*				
M_{23}	*the number of registered patients per 1 medical professional employed*				
M_{31}	*the average waiting time for patients to visit*	C_3	*timely delivery of services*	E_2	*ability to meet the needs of patients*
M_{32}	*the average duration of the visit*				
M_{41}	*the number of returning patients as % of the total number of patients*	C_4	*lasting relationships with patients*		
M_{42}	*the number of regular patients as % of the total number of patients*				
M_{43}	*the number of visiting patients as % of the total number of patients*				
M_{44}	*the number of foreign patients as % of the total number of patients*				
M_{51}	*the number of services sold per 1 medical professional employed*	C_5	*sales*	E_3	*obtained sales results*
M_{52}	*the value of medical services sold per 1 medical professional employed*				
M_{53}	*return on sales*				
M_{61}	*the average wage rate of medical personnel*	C_6	*costs and expenses*		
M_{62}	*the average wage rate of administrative personnel*				
M_{63}	*the cost of labor of administrative personnel as % of labor costs of the medical personnel*				
M_{64}	*the cost of labor of medical personnel as % of the services sold*				
M_{65}	*promotion and marketing costs as % of services sold*				
M_{71}	*the total value of fixed assets as % of services sold*	C_7	*the use of fixed assets*		
M_{72}	*the value of medical equipment as % of services sold*				
M_{81}	*the value of medical equipment per one person of medical professional employed*	C_8	*personnel productivity*		
M_{82}	*the value of profits attributable to one medical professional*				
M_{83}	*% of employees subject to some form of training*				

The factors presented in tab. 1 were evaluated by experts in accordance with the following procedure (tab. 2):

1. Competitiveness measures values M_{ijk} were computed based on the basis provided data in considered applications;
2. The values of measures were aggregated to the value of individual competitiveness factors C_{ij};
3. Competitiveness factors were synthesized in the assessment of individual areas of measurable effects E_i;
4. Finally, on the basis of individual assessments of measurable effects areas has been designated aggregated evaluation as competitive position CP, occupied by each dental clinic in research group.

The procedure of determining the aggregate assessment CP starts calculating the value of individual competitiveness measures M_{ijk}, arising from the formulas fM_{ijk} associated with each measure. On the basis of competitiveness measures the values of weighted competitiveness factors are determined (according to the formula of competitiveness factor fC_{ij}):

$$fC_{ij}: C_{ij} = \sum_{k=1}^{n} M_{ijk} * w_{ijk} \quad (4a)$$

where:

n – number of competitiveness measures defined for particular competitiveness factor,
C_{ij} – value of the j-th competitiveness factor in i-th measurable effects area,
M_{ijk} – the k-th competitiveness measure of the j-th competitiveness factor in the i-th measurable effects area,
w_{ijk} – weight assigned to competitiveness measure M_{ijk}, where the sum of weights w_{ijk} designated for each competitiveness factor C_{ij} is 1:

$$\bigwedge_{c_{ij}} \sum_{k=1}^{n} w_{ijk} = 1 \quad (4b)$$

The ranking positions in research group (PC_{ij}) are determined based on the values of competitiveness factors. They form the basis of determining the weighted value of each measurable effects area (according to the formula of measurable effects area fE_i):

$$fE_i: E_i = \sum_{j=1}^{n} PC_{ij} * w_{ij} \quad (4c)$$

where:

n - number of competitiveness factors defined for a particular measurable effects area;
E_i – aggregated value of the i-th measurable effects area;
PC_{ij} – ranking position of the j-th competitiveness factor in the i-th measurable effects area;
w_{ij} – weight assigned to the j-th competitiveness factor in the i-th measurable effects area, where the sum of weights w_{ij} designated for each measurable effects area is 1:

$$\bigwedge_{E_i} \sum_{j=1}^{n} w_{ij} = 1 \quad (4d)$$

The final element of the procedure is calculating the value of the clinic's competitive position CP (according to the formula of competitive position fCP):

$$fCP: CP = \sum_{i=1}^{n} E_i * w_{i,} \quad (4e)$$

where:
n – number of measurable effects areas,
CP – value of competitive position of the clinic in research group,
E_i – aggregated value of the i-th measurable effects area,

w_i – weight assigned to the i-th measurable effects area, where the sum of weights w_i designated for competitiveness position is 1:

$$\sum_{i=1}^{n} w_i = 1 \quad (4f)$$

Table 2 Reference model for identifying aggregated assessment of competitive position of individual dental clinics in selected SMEs group (Source: own research)

Assessment criteria for applications[1]	The function of competitiveness measure	The value of competitiveness measure	The function of competitiveness factor	The value of competitiveness factor	The function of measurable effects area	The value of measurable effects area	The function of competitive position	The value of competitive position
SUN, SB	fM_{111}	M_{111}	fC_{11}	C_{11}	fE_1	E_1	fCP	CP
WIR, SB	fM_{112}	M_{112}						
LR, LWO	fM_{121}	M_{121}	fC_{12}	C_{12}				
WR, SB	fM_{122}	M_{122}						
LPO, PM	fM_{123}	M_{123}						
SCO	fM_{211}	M_{211}	fC_{21}	C_{21}	fE_2	E_2		
SCT	fM_{212}	M_{212}						
LSP, LWO	fM_{221}	M_{221}	fC_{22}	C_{22}				
L3L, LWO	fM_{222}	M_{222}						
LPZM, LWO	fM_{223}	M_{223}						
LPZG, LWO	fM_{224}	M_{224}						
LWO, PM	fM_{311}	M_{311}	fC_{31}	C_{31}	fE_3	E_3		
SB, PM	fM_{312}	M_{312}						
ZS, SN	fM_{313}	M_{313}						
SPM	fM_{321}	M_{321}	fC_{32}	C_{32}				
SPA	fM_{322}	M_{322}						
KRA, KRM	fM_{323}	M_{323}						
KRM, SB	fM_{324}	M_{324}						
KMR, SB	fM_{325}	M_{325}						
CWST, SB	fM_{331}	M_{331}	fC_{33}	C_{33}				
WS, SB	fM_{332}	M_{332}						
WS, PM	fM_{341}	M_{341}	fC_{34}	C_{34}				
ZS, PM	fM_{342}	M_{342}						
PS, PO	fM_{343}	M_{343}						

This method was used for determining aggregated assessment CP for each submitted application. This assessment is the basis of the ranking and selection of applications to be approved for funding.

The Procedure for Selecting Project Assessment Criteria and Automating the Assessment

The assessment examined source data contained in the applications (Appendix A, Tab. 7) based on the established formulas and weight values necessary to determine aggregate assessment CP for each application. As a result of combining these two elements a training data set was created for the procedure of selecting key competitiveness factors and determining an automatic assessment of each submitted application. The set includes all data from the submitted applications, i.e. 32 independent variables (Appendix A, Tab. 7), aggregated assessment of CP (Formula 4e) and an additional variable: a binary assessment of OC.

The value of binary evaluation OC indicates that the application will (for OC=1) or will not be (for OC=0) approved. The assessment OC value is a consequence of the received aggregated assessment value CP in accordance with the formula:

$$OC_i = \begin{cases} 1 \text{ for } CP_i > \frac{\sum_{i=1}^{n} CP_i}{n} \\ 0 \text{ for } CP_i \leq \frac{\sum_{i=1}^{n} CP_i}{n} \end{cases} \quad (4g)$$

where n – number of assessed applications.

This helped create a procedure for selecting project assessment criteria that selects only those whose impact on the value of aggregated assessment CP and binary evaluation OC is significant. Data mining methods applied to build this procedure are in fact the various variants of regression. The SAS Enterprise Miner tool was used for their implementation.

It was assumed that response variable (dependent variable) would be variable OC for all variants of models and that it would take binary values: 1 - the application is approved for funding , 0 - the application is not approved for funding. It was also decided that at the level the analysis aimed at identifying the significance of the introduced variables the following types of the regression method will be used, as outlined in Tab. 3:

M1 - Logistic regression with probit function;
M2 - Regression model with parameterization imposed by SAS Enterprise Miner software;
M3 - Two-stage and sequential model of regression and decision trees;
M4 - Logistic regression with logit function;
M5 - Partial Least Squares (PLS) regression.

Table 3 Characteristics of particular variants of regression methods (Source: own research)

Model options	**M1**	**M2**	**M3**	**M4**	**M5**
Type of regression model	logistic regression	regression with parameterization imposed by software	two-stage and sequential model: regression + decision trees	logistic regression	PLS regression
Selection method of variables to the model	all available variables	all available variables	R-square and Chi-square method	R-square and Chi-square method	PCA (Principal Components Analisys)
Function combining average response with linear predictor	probit	-	-	logit	-
Selection method of model effects	back propagation	R-square method	a posteriori probability in classification	stepwise propagation	significance of variables

Model M1 obtained the best scores (Tab. 4). Based on it the key criteria of assessing applications suggested by this model are: SCT (average visit duration), ZS (profit generated in the period at hand) and the interaction of variables CWST (total value of fixed assets) and KM (the cost of medical supplies). They all have a positive impact on the dependent variable value, what means that an increase in their value increases the probability of OC = 1.

Table 4 The comparison of scores obtained by M1-M5 models (Source: own research)

Model	Validation: the degree of misclassification	Learning: the degree of misclassification	Validation: error function	Learning: error function	Validation: the number of misclassification	Learning: the number of misclassification	Validation: mean square error	Learning: mean square error
M1	0,125	0,08824	48,3296	9,04237843	2	3	0,142413	0,04581
M5	0,1875	0,00000	11,4129	7,29306881	3	0	0,467744	0,55639
M4	0,1875	0,05882	25,3240	16,99398079	3	2	0,180918	0,07147
M3	0,1875	0,08824	37,7574	15,31365666	3	3	0,179546	0,07084
M2	0,3125	0,00000	188,0652	0,00000005	5	0	0,295661	0,00000

M5 scored slightly lower in the qualitative assessment. However, it also scored lowest on error function value in the validation set. There is a probability that a combination of results of model M1 and model M5 could significantly have a major impact on the effectiveness of predictive assessment of applications in evaluation OC. A method of averaging complex results of M1 and M5 was used to verify this hypothesis. The assessments of the effectiveness of the results of validation model M6 (Tab. 5) revealed that it could be accepted as the final model of selecting criteria and automatic assessment of applications for funding for competitiveness development projects in dental clinics in the SME sector.

Table 5 Comparison of scores obtained by M6 vs. M1 and M5 models (Source: own research)

Model	Validation: the degree of misclassification	Learning: the degree of misclassification	Validation: error function	Learning: error function	Validation: the number of misclassification	Learning: the number of misclassification	Validation: mean square error	Learning: mean square error
M6	0,0625	0,00000	8,2018	7,10735595	1	0	0,077084	0,02220
M1	0,125	0,08824	48,3296	9,04237843	2	3	0,142413	0,04581
M5	0,1875	0,00000	11,4129	7,29306881	3	0	0,467744	0,55639

Based on model M6, an automatic evaluation of 20 new grant applications will be made (Fig. 2). It will take place without the expert, but only by using the prepared model.

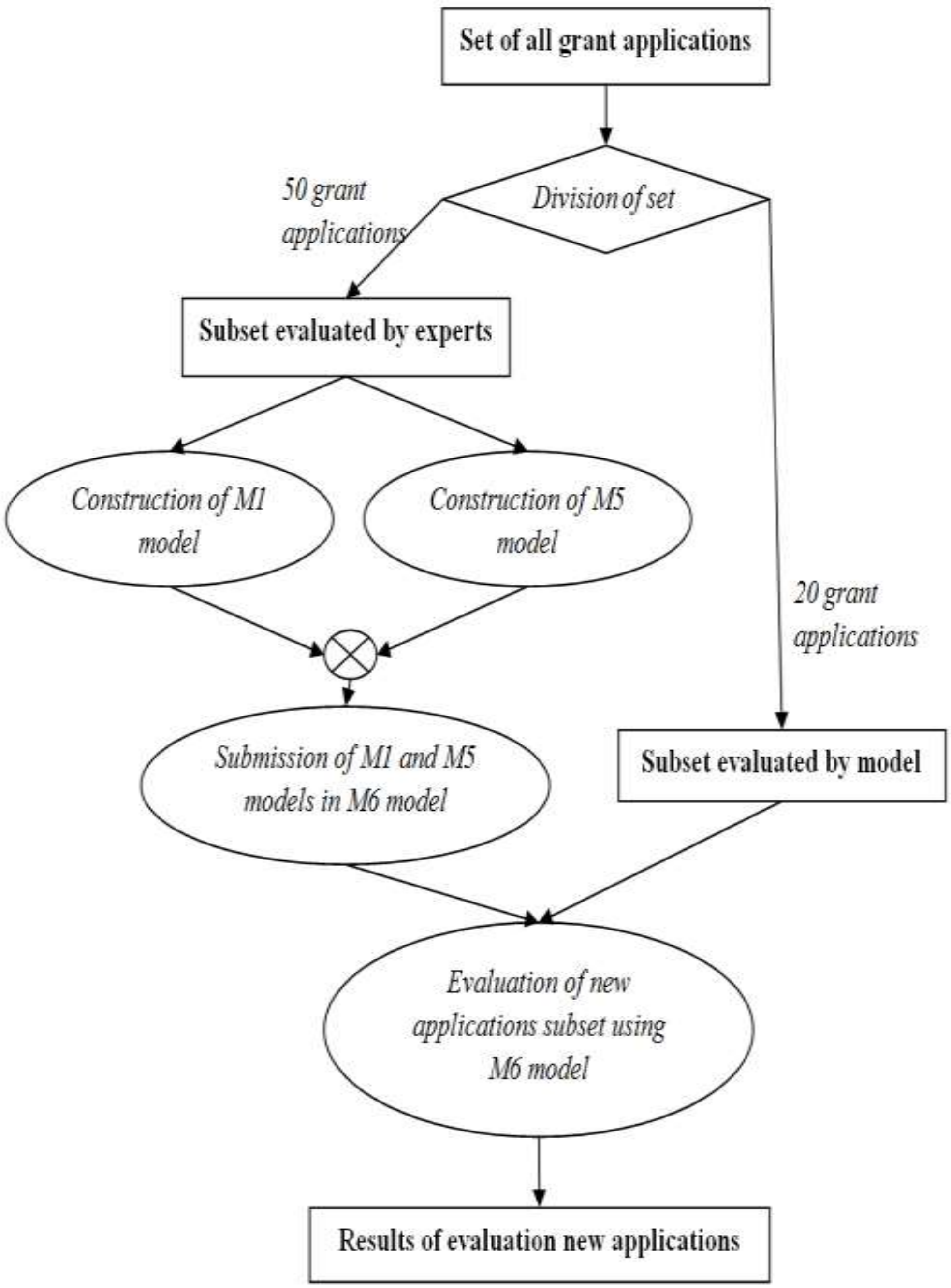

Figure 2 The structure and application of M6 model (Source: own research)

Tab. 6 shows the results of predictive values of OC on the subset of 20 new grant applications. The degree of misclassification reached 4 cases (cases no. 7, 15, 17 and 20), i.e. 20% of entire set of applications. This is a relatively high error rate and it is caused by an insufficient number of cases in the training and validation set, which makes it impossible to generalize the model.

Table 6 The results of predicting the OC values in the subset of new applications (Source: own research)

No. of case	CP	OC	Predictive OC
1	5.0949	0	0
2	4.9284	0	0
3	6.0273	0	0
4	4.8951	0	0
5	4.8951	0	0
6	4.9617	0	0
7	**4.2291**	**1**	**0**
8	3.7629	1	1
9	4.9284	0	0
10	4.2291	1	1
11	4.8951	0	0
12	5.5611	0	0
13	6.2271	0	0
14	5.0949	0	0
15	**4.5954**	**1**	**0**
16	5.2614	0	0
17	**4.5621**	**1**	**0**
18	4.0293	1	1
19	4.5954	1	1
20	**4.6287**	**0**	**1**

These results suggest that training and validation of this model on a wider set of cases (at least 100 cases in line with the applicability conditions of the regression function) will generate satisfactory results both in terms of key selection criteria for application/project assessment and their automatic assessment by a trained set.

Conclusions

This piece of research has demonstrated that innovative projects have not always been assessed against objective criteria but rather against scores that reflect a singular preference of the source of funding. This implies that choices are not free from bias. By analyzing the significance of the criteria using the available data mining methods from the group of discriminant, associative and grouping analysis, it is possible to identify only those which have a dominant influence on the final assessment.. This means that despite arbitrarily selected quantitative and qualitative assessment criteria, it is always possible to objectively assess projects, which is evidenced by the method and its use case presented in this paper.
It should also be noted that the design of the model of criteria selection and automation and the method of assessing grant applications allow the use of backward propagation and search for the answer to the following question: *What values and combinations of assessment criteria ensure the success of the assessment process?* Such an approach allows the applicant to make informed decisions with regard to the project proposal both in terms of the applicant's individual capabilities and the requirements of project sponsor.

With regard to RQ1, this current research project has revealed that the multi-criteria analysis and data mining can effectively assist the selection of criteria and automating the assessment of the commercial potential of innovative projects. They contribute to improved optimization of the range of requirements imposed on applicants and builds transparency, comprehensibility and reproducibility of application assessments. The possibility of automating the evaluation process leads to savings in time and effort contributed by experts, which responds to research question RQ2. Finally, the positive answers to research questions RQ1 and RQ2 support the research hypothesis RH.

References

Abratt, R., & Lombard, A. A. (1993). Determinants of Product Innovation in Specialty Chemical Companies. *Industrial Marketing Management,* 22 (1993), 169-175. DOI: 10.1016/0019-8501(93)90002-O.

Asli, D. A. T. (2007). Assessment of Factors Influencing Destination Image Using a Multiple Regression Model. *Tourism Review*. 62 (2), 23 – 30. DOI: 10.1108/16605370780000311.

Åstebro, T. (2004). Key Success Factors for R&D Project Commercialization. *Engineering Management, IEEE Transactions*, 51(3), 314-321. DOI: 10.1109/TEM.2004.830863.

Balachandra, R., & Friar, J. (1997). Factors for Success in R&D Projects and New Product Innovation: A Contextual Framework. *IEEE Transactions on Engineering Management*, 44(3), 276-87. DOI: 10.1109/17.618169.

Bandarian, R. (2007). Evaluation of Commercial Potential of a New Technology at the Early Stage of Development with Fuzzy Logic. *Journal of Technology Management & Innovation*, 2(4), 73-85. url: http://www.jotmi.org/index.php/GT/article/viewArticle/418.

Belton, V., & Stewart, T. J. (2002). *Multiple Criteria Decision Analysis: An integrated approach*. Dordrecht: Kluwer Academic Publishers.

Brinkhuis, J., & Tikhomirov, V. (2005). *Optimization: Insights and Applications*. New York: Princeton University Press.

Cooper, R. G. (1981). An Empirically Derived New Product Project Selection Model. *IEEE Transactions on Engineering Management*, 28(3), 54-61. DOI: 10.1109/TEM.1981.6448587.

Cooper, R. G. (1985). Selecting Winning New Product Projects; Using the NewProd System. *Journal of Product Innovation Management,* 2(1), 34–44. DOI: 10.1016/0737-6782(85)90014-1.

Cooper, R. G., & Edgett S. J. (2010). Developing a Product Innovation and Technology Strategy for Your Business. *Research Technology Management,* 53(3), 33-40. url: http://www.ingentaconnect.com/content/iri/ rtm/2010/00000053/00000003/art00006.

Draper, N., & Smith, H. (1966). *Applied Regression Analysis*. New York: John Wiley & Sons Inc. Publ.

Ehrgott, M. (2005). *Multicriteria Optimization*. Berlin, Heidelberg: Springer.

Figueira, J., Greco, S., & Ehrgott, M. (2005). *Multiple Criteria Decision Analysis: State-of- the- Art Surveys*. Berlin, Heidelberg: Springer.

Galbraith, C. S., DeNoble, A. F., & Ehrlich, S. B. (2012). Predicting the Commercialization Progress of Early-Stage Technologies: An Ex-Ante Analysis. *IEEE Transactions on Engineering Management*, 59(2), 213-225. DOI: 10.1109/TEM.2010.2068050.

Goldenberg, J., Lehmann, D. R., & Mazursky, D. (2001). The Idea Itself and the Circumstances of Its Emergence as Predictors of New Product Success. *Management Science,* 47(1), 69-84. DOI: 10.1287/mnsc.47.1.69.10670.

Hopkins, D.S. (1981). New-product winners and losers. *Research Management,* 24(3), 12–17.

Jiang-Liang, H., Hsiu-Yan, L. (2006). A Multiple Regression Model for Patent Appraisal. *Industrial Management & Data Systems,* 106(9), 1304 – 1332. DOI: 10.1108/02635570610712591.

Jolly, V. K. (1997). *Commercializing New Technologies: Getting From Mind To Market*. Boston: Harvard Business School Press.

Kathleen, A. R. (2003). *Bringing New Technology to Market.* New Jersey: Prentice Hall.

Keeney, R.L., & Raiffa, H. (1993). *Decisions with Multiple Objectives: Preferences and Value Trade-offs. 2nd ed*. Cambridge: Cambridge University Press.

Kim, J., & Feree, G. (1981). Standardization in Causal Analysis. *Sociological Methods and Research,* 10(2), 187-210. DOI: 10.1177/004912418101000203.

Lester, D.H. (1998). Critical Success Factors for New Product Development. *Research Technology Management,* 41(1), 36–43. DOI: 27AF403.

Loftus, B. S., & Meyers, P. W. (1994). Launching Emerging Technologies to Create New Markets: Identifying Industrial Buyers. *Logistics Information Management*, 7(4), 27-34. DOI: 10.1108/09576059410066408.

Madique, M.A., & Zirger, B.J. (1984). A Study of Success and Failure in Product Innovation: The Case of the U.S. Electronics Industry. *IEEE Transactions on Engineering Management*, 31(4), 192–203. DOI: 10.1109/TEM.1984.6447537.

Midgley, D. (2009). *The Innovation Manual: Integrated Strategies and Practical Tools for Bringing Value Innovation to the Market.* New York: John Wiley & Sons Inc. Publ.

Myers, R.H. (2000). *Classical and Modern Regression with Applications. 2nd ed.*, Pacific Grove: Duxbury Press.

Nibler, R. (1997). A Regression Model for Matching Parallel Systems. *International Journal of Quality & Reliability Management*, 14(2), 176 – 185. DOI: 10.1108/02656719710165437.

Peter, B. P. (1999). *The Valuation of Technology: Business and Financial Issues in R&D*. New York: John Wiley & Sons Inc. Publ.

Peterson, M. (2009). *An Introduction to Decision Theory*. Cambridge: Cambridge University Press.

Rackham, N. (1988). From Experience: Why Bad Things Happen to Good New Products. *Journal of Product Innovation Management*, 15(3), 201–207. DOI: 10.1111/1540-5885.1530201.

Rostek, K. (2012). The Reference Model of Competitiveness Factors for SME Medical Sector. *Economic Modelling*, 29(5), 2039-2048. DOI: 10.1016/j.econmod.2012.03.002.

Shih-Ming, Y., & Chin-oh C. (2009). Weight Regression Model from the Sales Comparison Approach. *Property Management*, 27(5), 302 – 318. DOI: 10.1108/02637470910998465.

Siegel, R. A., Hansén, S. O., & L. H., Pellas (1995). Accelerating the Commercialization of Technology: Commercialization through Co-operation. *Industrial Management & Data Systems*, 95(1), 18 – 26. DOI: 10.1108/02635579510079425.

Souder W. E. (1973). Utility and Perceived Acceptability of R&D Project Selection Models. *Management Science*, 19(12), 1384-94. DOI: 10.1287/mnsc.19.12.1384.

Van der Panne, G., Van Beers, C., & Kleinknecht, A. (2003). Success and Failure of Innovation: A Literature Review. *International Journal of Innovation Management*, 7(3), 309-337. DOI: 10.1142/S1363919603000830.

Yoon, E., & Lilien, G.J. (1985). New Industrial Product Performance. The Effects of Market Characteristics and Strategy. *Journal of Product Innovation Management*, 2(3), 134–144. DOI: 10.1016/0737-6782(85)90033-5.

Appendix A

Table 7 Characteristics of criteria for assessing grant applications for competitiveness development in dental clinics in the SME sector (Source: own research)

Source data	Desription of source data
Data about patients:	
LPO	Total number of patients
LSP	Number of patients who used the services of the company at least 3 times a year
L3L	Number of patients who have been patients for at least 3 years
LPZM	Number of patients living more than 50 km from Warsaw
LPZG	Number of foreign patients
Data about sales of services:	
SB	Value of gross sales of services in PLN thousands
SN	Value of net sales of services in PLN thousands
SUN	Gross value of sales of innovative services in PLN thousands
ZS	Profit
LWO	Number of visits in the period
SCO	Average patient's wait time for a visit in days
SCT	Average duration of a visit in minutes
LR	Number of complaints reported by patients
WR	Gross value of complaints reported by patients in PLN thousands
Data about employees:	
PO	Total number of employees
PM	Number of medical employees
PA	Number of administrative personnel
PS	Number of workers benefiting from any form of training
LRM	Number of man-hours worked by medical personnel
LRA	Number of man-hours worked by administrative personnel
SPM	Average wage rate of medical personnel
SPA	Average wage rate of administrative personnel
KRM	Labor cost of medical personnel in PLN thousand
KRA	Labor cost of administrative personnel in PLN thousands
Data about costs incurred:	
KM	Cost of medical supplies in PLN thousands
KN	Costs of inspections and repairs of medical equipment in PLN thousands
WS	Value of medical equipment in PLN thousands
WZM	Value of stocks of medical supplies in PLN thousands
KMR	Marketing and promotion costs in PLN thousands
WIR	Investment and development expenditure in PLN thousands
CWST	Total value of fixed assets in PLN thousands
WB	Value of buildings in PLN thousands