

## METHODICAL APPROACH TO DATA PROCESSING FROM A QUESTIONNAIRE SURVEY

Hana Tomášková\*, Hana Šlachťová, Anna Šplíchalová

Department of Health Impact Assessment, Institute of Public Health, Partyzánské nám. 7, 702 00 Ostrava, Czech Republic, e-mail: hana.tomaskova@zuova.cz

Received: June 20, 2003; Accepted in revised form July 10, 2003

Key words: Questionnaire survey/ Pre-research/ Repeatability study/ Quality of data

For conducting questionnaire studies there exist generally recommended methods that were also respected during the project funded by the Grant Agency MoH CR – “Subjective approach of inhabitants of Ostrava to their health in association with their life-style, socio-economic status and education”. The preparatory phase included the collection of literature and information on the investigated theme. The questionnaire had got five parts – A. General questions, B. Employment, C. Way of life, D. Health state and E. Personality. The validity of questionnaire was tested in the pre-research. In the main questionnaire study 3,000 questionnaires were sent. The total response rate was 21.1% (634 completed questionnaires). In the sample there were no differences in percentage rate by sex and age, but there are differences in the educational structure. After realizing the main questionnaire study, the repeatability study was carried out to find out the reliability of the answers. The response rate was 60.3% (181 questionnaires). The Kappa index and the total percentage of agreement were used for the evaluation of the repeatability study. The agreement was almost perfect and good in the total of 62.3 % of the questions. The quality of data was ensured by double data entry and by choosing the appropriate software. The selection of data for the complex analysis was based on the results of the repeatability study. On the basis of individual information from the questionnaire, new groups of individuals were generated. These groups of individuals were analysed further in relation to health and life-style by socio-economic factors in the models.

### INTRODUCTION

During questionnaire surveys it is necessary to keep generally recognized methods<sup>3, 7, 10, 11, 26</sup>. These were also used while putting into practice a project funded by the Grant Agency of the Ministry of Health of the Czech Republic No. 6139-3 „Subjective approach of inhabitants of Ostrava to their health in association with their life-style, socio-economic status and education”. The article draws one’s attention to the methods, which can improve quality and reliability of data (pre-research, repeatability study and double data entry), sources of bias and their solution, and methods of gathering information that are entered into models. Further, the article states statistical methods and software used while solving the aforementioned study.

#### Material, methods and results

The individual steps, which were put into practice in terms of the questionnaire survey, are pictured on the following scheme (Fig. 1.).

The first phase, which included forming the questionnaires, was time-consuming. It represented collection of information and study of literature with similar objectives<sup>20</sup>. The questionnaire, which could have been used in terms of this project was not found in available sources, which is why an appropriate questionnaire has

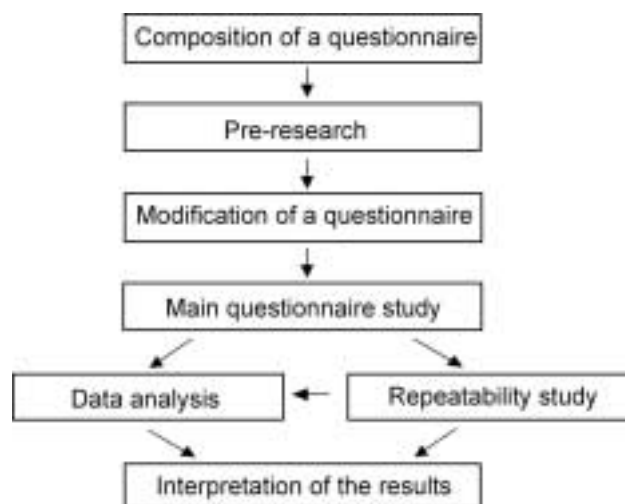


Fig. 1. Scheme of the questionnaire survey.

been formed. Further, the size of sample from a studied population was specified taking into account the expected response rate according to a form of distribution and collection of questionnaires. The calculation of size of a representative sample is mentioned in a lot of literary sources; for example<sup>3, 4, 11</sup>. The size of the sample of population in this study was defined according to

Katriak's<sup>11</sup> methodology. It recommends this range of addressing inhabitants: from 0.25 % of city inhabitants for cities with 1 million inhabitants to 1.5 % for a city of up to 100 thousand inhabitants. Ostrava has 319,000 inhabitants. Because of an assumed low response rate, circa 30 %, in the case of distribution and collection by mail<sup>23</sup>, 3,000 inhabitants were addressed by a random sample.

While forming the questionnaire, questions were phrased in the most understandable way possible, answers were put into scales and categorized as much as possible as recommended in the literature<sup>10</sup>. The visual part of a questionnaire is also very important<sup>5</sup>. For whom a questionnaire is meant must be taken into consideration. There are different opinions of the recommended number of questions, possibly of pages and the length of completion time for the questionnaire. Whereas the results of meta-analysis of 115 studies<sup>25</sup> have confirmed that a questionnaire shorter than four pages should increase the response rate, other authors have proved that the length of a questionnaire has no influence on the response rate<sup>1</sup>. Dillman's work<sup>5</sup> states that the response rate of a questionnaire survey is not influenced in a negative way by any length of a questionnaire up to 12 pages and 125 questions.

The questionnaire in this study was extensive because it included a broad spectrum of problems and consisted of 102 questions in total. The questions were divided into five parts according to the investigated problems – A. General questions, B. Employment, C. Way of life, D. Health state and E. Personality. The average length of completion time for the questionnaire was about 35 minutes<sup>18</sup>.

After forming the questionnaire, the verification of its validity in a pre-research was put into practice. In technical literature authors differ from each other in terminology concerning this part of a questionnaire survey. Apart from the term pre-research, other terms such as pilot study or pre-test have also been used. For example, Kapr<sup>10</sup> has been using the term pre-research from which he expects the verification of understandability and unambiguousness in forming questions, the verification of ways of manipulation with a questionnaire and of people's reactions at interviews and the verification of some particular hypotheses and technical processibility of data etc. On the other hand Žáček<sup>26</sup> has been using the term pilot survey (pilot study) which represents a probe done for a trial period and on a smaller scale, the purpose of which is to obtain preliminary experience from all the sections of the planned study, to clarify some circumstances of the research, to test the preconditions of using the chosen method and to have the correct idea about the costs and possible difficulties etc. Schneider and Koudelka<sup>16</sup> have been dividing the pre-research into a pilot study, which represents the introductory research aimed at gaining a general orientation in a given area of a problem, and into a pre-test whose aim is especially to test reliability and applicability of the chosen techniques. The number of cases for a pre-

test is not to be smaller than 25. Disman<sup>7</sup> is inclined to believe the definitions mentioned in most American technical publications, in which the aim of a pilot study is to find out whether a particular research in a given population is possible at all. On the other hand the purpose of pre-research is to test instruments (a questionnaire), which were constructed for a given research.

On the basis of this information the authors of the presented study were inclined to believe in using the term pre-research. Thirty people were addressed in the pre-research. It was a quota selection by sex, age and education from the total number of about 400 people. The response rate was 56.7 % (17 completed questionnaires). The correct understanding of questions and the sufficient range of categories were tested. The total reaction of respondents to a questionnaire was found out as well and the program for data entry was tested at the same time.

After the evaluation of the pre-research results the modification of the questionnaire followed so that all the inconsistencies were removed, for example modifications of the questions, the possibility of more answers, the enlargement of categories of answers etc. The manual for data entry and data cleaning was created too.

In the main questionnaire study 3,000 questionnaires were distributed. A cover letter with the information about the aims of the study and investigators, and envelopes with the return address and prepaid postage were sent together with the questionnaire. During the distribution and collection of questionnaires all the reactions of respondents were recorded and the course of response rate was recorded as well<sup>18</sup>. The total response rate was 21.1 %, which represents 634 completed questionnaires. Because of the relatively low response rate the comparison of the sample data (sample = the ones who handed in the completed questionnaire) and the data about the studied population (Ostrava inhabitants) was done by addition<sup>18</sup>. The literature states that among non-respondents in most cases<sup>6</sup> there are people with lower education, women and younger people. In our sample there are no differences in percentage rate by sex and age, but there are differences in the educational structure. Apprentices represented 33% of the respondents in the investigated sample, which approximates the percentage of people with apprenticeship education in Ostrava. There is lower interest in participating in studies in respondents with basic education, whereas people with secondary and university education are over-represented. It is about a selection bias<sup>9</sup>. It is necessary to take this fact into consideration while evaluating results and all the data should be adjusted for education<sup>22</sup>. A low response rate of the questionnaire distributed by mail is also mentioned in the technical literature – less educated people prefer to participate in interviews or telephone research which does not require demands of reading and writing<sup>6</sup>. A lot of questionnaire studies with a low response rate include also parts, which find out information about non-respondents. Some studies have confirmed that a low response

rate leads to the reduction of the representativeness of a sample<sup>1,7,23</sup> however it does not have to concern all the investigated characteristics. For example Reijneveld and Stronks<sup>14</sup>, who were interested in the relationship between health state and the use of medical care in the Dutch population, studied the effect of non-respondents on validity of data. The data about the selected sample were compared with the information from health insurance companies. It concerned 2,934 respondents and 1,744 non-respondents. They have found out that from a statistical point of view both files were not significantly different in any basic characteristics such as sex, marital status, length of time spent at permanent address, nationality etc., but there was a significant difference in age structure. However, some papers have proved that respondents participating in a research after the first contact are not different from the respondents who joined the survey after a repeated contact. Siemiatycki and Campbell<sup>15</sup> were comparing the samples of respondents in two surveys while using various methods. In the first study the participants of the survey were contacted by mail; after reminding them by telephone or in person the response rate increased by another 12.5 %. The second study was a telephone survey; the questionnaire was sent by mail or delivered in person to the ones who had not been caught at home by a phone. The response rate increased by another 15.5 %. No significant differences were found by comparing the features of both samples in the first and the second studies. The authors came to the conclusion that increasing the response rate further was useless<sup>15</sup>.

While preparing the project funded by the Grant Agency the investigators did not plan to put the non-respondents study into practice because of high costs of postage and because of the fact that to contact the respondents was only possible through the address. A telephone contact has mostly been used for research of non-respondents but this method was not appropriate for the reason that only 69.9 % of households in the Czech Republic are equipped with telephone (the data of Czech Telecom from the year of 2001). The response rate was influenced in a negative way by the timing of the study which unexpectedly coincided with a population census. The population census was accompanied by a media campaign about the insufficient protection of personal data and about questioning the legitimacy of any collections of personal data. However, the planned repeatability study, which tested the reliability of data, was realized.

The repeatability study was done six weeks after finishing the main study. Three hundred respondents were chosen out of the total number of 600 people who returned the completed questionnaire. The selection was done in the following way. The respondents were sorted according to the date of delivery of the completed questionnaire in the main study; every other person was selected. This method of selection was used on purpose so that it evenly included the respondents who completed the questionnaire and sent it immedi-

tely after its delivery as well as the ones who hesitated with completing and sending it. A part of the delivery was again an explanatory letter so that the highest response rate was achieved. The response rate was higher than in the main study – 60.3 % (181 questionnaires). The literature<sup>7, 10</sup> states that not all the answers to the questions have to be verified – only the selected ones; however, in this study the questionnaire was used in its entirety. The method of calculation of the Kappa index and the calculation of total percentage of agreement were used for the evaluation of the repeatability study. The calculation of the Kappa index was first suggested by Cohen<sup>2</sup>. Its further modification was the purpose of the study by Landis and Koch<sup>13</sup>, and by Fleiss<sup>8</sup>. The calculation of the Kappa value is based on a total ratio of agreement and an expected ratio of agreement. The interpretation of agreement according to the Kappa value is divided into four groups (< 0.4 – poor; 0.41–0.6 – average; 0.61–0.80 – good; 0.81–1 almost perfect). The agreement was almost perfect and good in the total of 62.3 % of the questions. The questions were divided into two groups: factual questions (sex, age, marital status etc.) and questions containing a feature of motivation or evaluation; possibly the respondent's opinion could be influenced by a present situation (having bad mood because something had happened, current mood, health state etc.). The values of the percentage of agreement (86.8 %; 72.1 %) and the Kappa index (0.73; 0.48) were significantly higher in the factual questions than in the questions of the second group. For further processing, the questions with lower values of the Kappa index and the percentage of agreement were a) substituted for similar questions, b) the categories of answers were aggregated, c) excluded. The detailed results of the repeatability study are mentioned in a separate paper<sup>22</sup>.

After gaining cleaned data, the data analysis followed. Starting points were basic descriptive statistics such as frequency tables, arithmetic mean, median, standard deviations and intervals of reliability etc<sup>11, 17</sup>. In the case of small frequencies the aggregation of categories of answers followed, through the analysis of relations between two variables using contingency tables, odds ratios, t-tests and regression analysis, leading to constructions of models. Further, new variables were formed and used in the models. Logistic regression<sup>17</sup> was used for the evaluation of the models.

All activities during these steps were recorded in detail.

While interpreting the results it is necessary to take into account selection bias, which was formed as a result of the low response rate. Some socio-economic factors as well as the results of the repeatability study have already been taken into consideration in the results of the models.

The important task is to guarantee the quality of entered data. The quality in terms of the presented study was guaranteed by double data entry followed by data cleaning. The people who entered the data had been qualified before and were following a methodical

manual. The examples of percentage of mistakes in various studies are mentioned in the Table 1. The per-

centage of mistakes in this study was about 3.3 %. The percentage of mistakes also depends on types of data.

**Table 1.** Percentage of mistakes.

Study	No. of records	No. of variables	Total	No. of mistakes	% of mistakes
A	50	263	13150	434	3.3
B	287	5	1435	17	1.2
C	46	160	7360	79	1.1
D	46	80	3680	38	1.0
E	55	14	770	3	0.4

The study:

- A. The questionnaire (general questions, employment, way of life, health state and personality).
- B. The sample with the results of measurements of heart frequency, lung capacity and energy expenditure.
- C. The questionnaire – life-styles of students.
- D. The results – basic internal examination of EKG, biochemical and hematological examinations, anthropometric examination, spirometric examination using the method of flow – volume and load test on bicycle ergometer.
- E. The furans – the results of measurements and coded values in numbers.

Some mistakes in quantitative items can be unimportant, for example in the last decimal number. However, a mistake of up to several degrees can appear which, if not found out before, can cause a considerably skewed result. Mistakes in qualitative items (for example sex) are serious as well. These can hardly ever be found out without checking double data entry or by another checking of entered data.

In the case of mistakes it is necessary to distinguish careless mistakes of the person who entered the data

and mistakes caused by an incorrect understanding or an incorrect completion by respondents. The correction is simple in the case of careless mistakes. However, in the second case the mistake must be solved by a defined method, which will be used in all similar cases. The solution must be suggested by an investigator not by the person who enters the data.

Table 2 shows the list of software, which was used during the solution of the project funded by the Grant Agency.

**Table 2.** The list of using software.

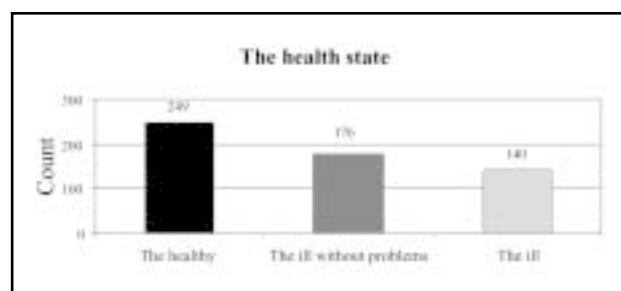
Process	Software
Data entry, validation	EpiInfo v.6
Data transfer	Stat/Transfer v.6
Data Analysis	Stata v.7
GIS	ArcView v.3.2
Classification and decision tree	AnswerTree v.3
Documentation	MS Word
Graphs	MS Excel
Presentations	MS Power Point

The software EpiInfo, version 6. cz<sup>4</sup>, was used for data entry and validation. The software Stat/Transfer was used for the transfer of data into the statistic system Stata, version 7<sup>17</sup>. The geographic information was processed by the software ArcView, version 3.2. The software SPSS AnswerTree, version 3 was used for the method of classification and decision tree. Further, the software from the MS Office package was used.

The following part of this paper contains the method of forming new items, which were necessary for further analysis and construction of models. It is about the items, which were created on the basis of answers to chosen questions.

Three groups of individuals were created according to the objective and subjective evaluations of health state (Fig. 2): a) the healthy – without chronic diseases and subjectively evaluating their health states as good,

b) the ill without problems – in their diagnoses there is a chronic disease but they subjectively evaluate their health states as good, c) the ill who suffer from chronic diseases and evaluate their health states as bad.



**Fig. 2.** Distribution of the respondents according to the health state.

Further, the respondents were divided into active and passive ones. Physical and social activities during free time, weekends and holidays were the basis of these labels. Nine questions which refer to the individual's activities were chosen:

1. does sports and hiking in free time,
2. seeks cultural and social life in free time,
3. devotes free time to hobbies,
4. spends weekends in an active way – by doing sports, hiking or going on trips,
5. devotes weekends to social and cultural activities,

6. uses holidays in an active way – by doing sports, hiking, sightseeing and gardening,
7. usually spends weekends out of Ostrava,
8. spends holidays at cottages or weekend houses, travels in the Czech Republic or abroad,
9. keeps frequent contacts with friends.

According to the number of positive answers to these questions the individual was placed into one of three categories – active in a minimal, average and maximal way (Table 3).

**Table 3.** Distribution of the respondents according physical and social activities.

Categories	Active in a minimal way	Active in a average way	Active in a maximal way	Total
Count of positive answers	0–1	2–3	4–7	max. 9
Count of respondents	98	278	259	635
Percentage of respondents	15.4	43.8	40.8	100

The same method was also chosen in the case of passive behaviour. Similarly, a group of questions was chosen documenting the individual's passivity:

1. often watches TV,
2. spends free time by reading,
3. spends weekends by reading,
4. does work in connection with one's job at weekends,
5. spends only one weekend in three months out of Ostrava,

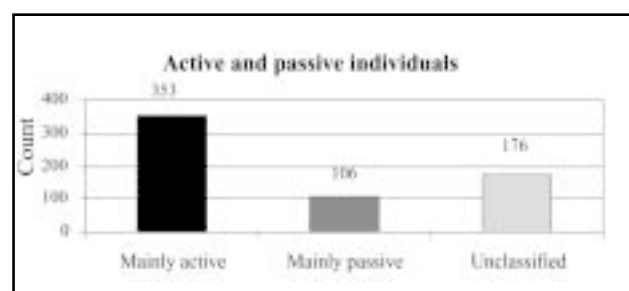
6. does not leave the city at all,
7. spends holidays at home,
8. does not do household tasks at all,
9. is not in contact with friends.

According to the number of positive answers to these questions the individual was placed into one of three categories – passive in a minimal, average and maximal way (Table 4).

**Table 4.** Distribution of the respondents according physical and social passivity.

Categories	Passive in a minimal way	Passive in a average way	Active in a maximal way	Total
Count of positive answers	0–1	2–3	4–7	max. 9
Count of respondents	273	305	58	635
Percentage of respondents	42.8	48.0	9.2	100

Based on the created features of behaviour the individuals were divided into two groups – mainly passive and mainly active individuals. The individuals who did not belong to neither the first nor the other group, were dropped from the further analysis (Fig. 3).



**Fig. 3.** Distribution of the respondents according activity.

Furthermore the individuals according well-being were characterized. The categories were based on the number of positive answers to these questions in the individual level (Table 6):

1. had to solve serious problems during the last year,
2. has conflict relationships to people,
3. is susceptible to stress,
4. cannot cope with stressful situations,
5. is not satisfied with his (her) economical situation,
6. often feels tired and irritable,
7. feels dissatisfied in general.

An individual with risk behaviour is characterized by a lack of interest in physical activity, unhealthy diet and by neglecting health problems:

1. does not devote time to regular exercising, sports and hiking (after excluding the people whose health problems unable them doing these activities),

**Table 6.** Distribution of respondents according well-being.

Categories	Well-being in a minimal way	Well-being in a average way	Well-being in a maximal way	Total
Count of positive answers	0–1	2–3	4–7	max. 7
Count of respondents	273	305	58	635
Percentage of respondents	42.8	48.0	9.2	100

2. drinks three and more cups of black coffee with caffeine per day,
  3. smokes daily,
  4. does not eat regularly in the course of the day,
  5. does not even have one warm meal a day,
  6. evaluates his (her) food as unhealthy or does not pay attention to it,
  7. does not pay attention to health problems when they appear, or follows the advice of family and friends and does not seek medical treatment,
  8. goes to work when he (she) has a cold with fever,
  9. if a doctor orders him (her) to be on sick-leave, he (she) does not take it,
  10. does not go to preventive check-ups,
  11. he (she) does not limit the consumption of food which can cause a health risk for him (her) and he (she) is aware of that.
- According to the number of positive answers the degree of risk behaviour was defined (Table 7).

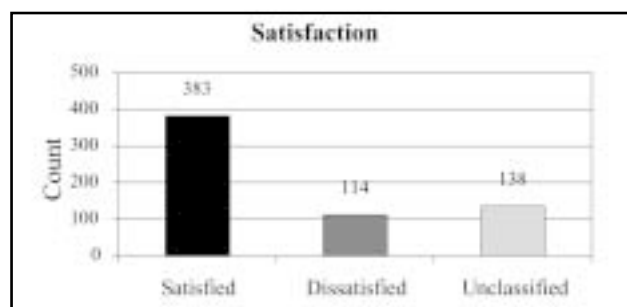
**Table 7.** Distribution of the respondents according risk behaviour.

Categories	Risk behaviour in a minimal way	Risk behaviour in a average way	Risk behaviour in a maximal way	Total
Count of positive answers	0–2	3–4	5–7	max. 11
Count of respondents	286	211	138	635
Percentage of respondents	45.0	33.2	21.7	100

The evaluation of satisfaction of respondents was done on the basis of the answers to the following questions:

1. He (she) is satisfied with the economical situation of his (her) family.
2. He (she) feels rested after his (her) holiday.
3. He (she) considers the time devoted to sleep as satisfactory.
4. He (she) considers his (her) food situation as satisfactory.
5. He (she) seldom feels tired or irritable.
6. He (she) evaluates his (her) physical condition as very good.
7. He (she) feels satisfied in general.

In the case of four or more positive answers the respondent was labelled as satisfied. In the case of four or more negative answers the respondent was labelled as dissatisfied. The other respondents were not included (Fig. 4).

**Fig. 4.** Distribution of the respondents according a satisfaction.

## CONCLUSION

The questionnaire survey, which was put into practice in terms of this project funded by the Grant Agency, consisted of these important steps. The preparatory phase included the collection of literature and information on the investigated issue. Further, the questionnaire was formed, the validity of which was tested in the pre-research. After realizing the main questionnaire study, the repeatability study was carried out to find out the reliability of the answers. The quality of data was guaranteed by double data entry and by choosing the appropriate software. The emphasis was put on recording all the steps and changes which happened during data processing. The selection of data for the complex analysis was based on the results of the repeatability study. On the basis of individual information from the questionnaire, new groups of individuals were formed – the healthy, the ill without problems and the ill; active and passive individuals; in well-being; with risk behaviour; satisfied and dissatisfied. These groups of individuals were analysed further in relation to health and life-style according to socio-economic factors in the models. The results of the models are the topic of further papers<sup>23, 26</sup>.

## REFERENCES

1. Asch DA, Jedrziwski MK, Christakis NA. (1997) Response Rates to Mail Surveys Published in Medical Journals. *J Clin Epidemiol* 10, 1129–36.

2. Cohen J. (1960) A coefficient of agreement for nominal scales. *Educational and Psychological Measurement* 20, 37–46.
3. Červenka J. *Základy statistiky*. Martin: Osveta, 1975.
4. Dean AG, Dean JA, Coulombier D, Brendel KA, Smith DC, Burton AH, Dicker RC, Sullivan K, Fagan RF, Arner TG. Epi Info, Version 6 – a word processing, database, and statistics program for epidemiology on microcomputers, Centers for Disease Control and Prevention, U.S.A.: Atlanta, Georgia 1994.
5. Dillman DA. Procedures for Conducting Government-Sponsored Establishment Surveys: Comparisons of the Total Design Method (TDM), A Traditional Cost- Compensation Model, and Tailored Design. <http://survey.sesrc.wsu.edu/dillman/papers/ICESII%20TDM%20paperProceedings.pdf>
6. Dillman DA, Phelps G, Tortora R, Swift K, Kohrell J, Berck J. Response Rate and Measurement Differences in Mixed Mode Surveys Using Mail, Telephone, Interactive Voice Response and the Internet. [http://survey.sesrc.wsu.edu/dillman/papers/Mixed%20Mode%20ppr%20\\_with%20Gallup\\_%20POQ.pdf](http://survey.sesrc.wsu.edu/dillman/papers/Mixed%20Mode%20ppr%20_with%20Gallup_%20POQ.pdf)
7. Dismar M. Jak se vyrábí sociologická znalost. Praha: Univerzita Karlova, Karolinum, 1993.
8. Fleiss JL. *Statistical Methods for Rates and Proportions*, 2<sup>nd</sup> ed. New York: John Wiley and Sons, 1981.
9. Janout V. *Epidemiologické studie*. Brno: IDPVZ, 1996.
10. Kapr J, Šafář Z. Sociologie nebo zdravý rozum?: praxe sociologického průzkumu. Praha: Mladá fronta, 1969.
11. Katriak M. *Metodologické zásady sociologického výzkumu*. Bratislava, 1968.
12. Kubánková V. *Statistika pro zdravotníky*. Avicenum Praha, 1986.
13. Landis JR, Koch GG. (1977) The measurement of observer agreement for categorical data. *Biometrics* 33, 159–74.
14. Reijneveld SA, Stronks K. (1999) The impact of response bias on estimates of health care utilization on a metropolitan areas: the use of administrative data. *Int J Epidemiol* 6, 1134–40.
15. Siemiatycki J, Campbell L. (1984) Nonresponse bias and early versus all responders in mail and telephone surveys. *Am J Epidemiol* 2, 291–301.
16. Schneider M, Koudelka F. *Úvod do základů sociologických výzkumů*. Olomouc: Univerzita Palackého, filozofická fakulta, 1993.
17. STATA manual version 7.0, Texas: Stata Press, College Station, 2001.
18. Šlachťová H, Máchová T, Tomášková H, Michalík J. Hodnocení návratnosti dotazníkové studie použitím GIS a standardních metod. Sborník z Vědecko-odborné konference s mezinárodní účastí Životní podmínky a zdraví. Štrbské Pleso 15.–17. 10. 2001. Bratislava, 2002. 125–134.
19. Šlachťová H, Tomášková H, Šplíchalová A. Analýza životního stylu ve vztahu k socioekonomickému statusu. Sborník z Vědecko-odborné konference s mezinárodní účastí Životní podmínky a zdraví. Štrbské Pleso 11.–13. 9. 2002.
20. Šplíchalová A., Šlachťová H. Faktory životního stylu a socioekonomického statusu ve vztahu ke zdraví jako předmět zájmu epidemiologických a sociologických studií. Sborník z Vědecko-odborné konference s mezinárodní účastí Životní podmínky a zdraví. Štrbské Pleso 15.–17. 10. 2001. Bratislava, 2002. 119–24.
21. Šplíchalová A, Tomášková H, Šlachťová H. Analýza subjektivně udávaného zdraví ve vztahu k socioekonomickým faktorům. Sborník z Vědecko-odborné konference s mezinárodní účastí Životní podmínky a zdraví. Štrbské Pleso 11.–13. 9. 2002.
22. Textbook – Advanced course in epidemiology. Epidemiology in Public Health, Cracow Summer school, 1999. London School of Hygiene & Tropical Medicine, Collegium Medicum, Jagiellonian University, Cracow, World Health Organization European Centre for Environment and Health, Billhoven, 1999.
23. Tomášková H. (2003) Lifestyle and socio-economic status – verification of data reliability from questionnaire surveys. *Cent Eur J Publ Health* 2, 21–6.
24. Vytlačil J. *Výběrová šetření v praxi*. Praha: SEVT, 1969.
25. Yammarino FJ, Skinner SJ, Childers TL. (1991) Understanding Mail Survey Response Behaviour. A meta-analysis. *Public Opinion Quarterly* 55, 613–39.
26. Žáček A. *Metody studia zdraví a nemocí v populaci*. Avicenum Praha, 1984.