COMPARING SAMPLING FRAMES

Ildikó Zakariás*1*, Renáta Németh*2*,

Hungarian National Centre for Epidemiology, *Budapest*

An alternative of sampling directly from a list of individuals is to draw an address-based sample: take a sample of households first then select the respondent within household at a second stage. Researchers employing the later often report the under-sampling of males and over-representation of elderly people; in such cases practical realization problems are usually blamed for the phenomenon. Previous theoretical results that reveal the inherent connection between the sampling design and representativity problems will be given further empirical support by analysing data from four Hungarian health surveys. Systematic comparison of the frames, planned and realized samples allows for separating sampling and non-response error, thus comparing the performance of the two designs. As in several European countries both sampling designs are possible to implement, according to our findings, in such cases it may be worth considering the aspects highlighted in the paper.

Key words: Leslie Kish grid, address-based sample, population-based sample

1 INTRODUCTION

In cases when the target population is the adult population of a certain area, it is reasonable to use an accessible list of adults for the sampling design and select respondents directly from the population. Selection methods of this type resulting in a population register sample are usually multi-stage, stratified designs with equal sampling probabilities of individuals (Design A). However in certain cases this type of survey design is impossible to apply; the same holds in cases when there is no such list of individuals available. This problem is often sorted out by using a different type of two-stage design: households are selected first with almost equal sampling probabilities, and then one adult member of each selected household is chosen applying a quasi-random procedure e.g. Leslie Kish grid, last birthday method etc (Design B).

1 zakariasi@oek.antsz.hu
2 nemethr@oek.antsz.hu

Relying on theoretical considerations as much as on empirical data, our present article aims to give a comprehensive comparison on the representativity of samples obtained by the two methods.

In Section 1 the principle aspects usually influencing the choice between the two methods will be described, also recent examples of their application will be listed underlining the relevance of the problem. In section 2 a major representativity bias inherently connected to Design B will be revealed in theory, while in Section 3 four Hungarian health surveys will be compared focusing on representativity and other related aspects such as weighting. Special attention will be given to the National Health Interview Survey carried out in 2000 and 2003 by the Hungarian National Centre for Epidemiology. The paper will conclude in Section 4.

2 GENERAL PRINCIPLES OF APPLICATION

A frequently referred cause of preferring an address-based design instead of population register based design is that – due to non-registration and mobility – in several countries electoral registers are bad-quality databases of individuals but are good-quality databases of households. Therefore the register is often applied to construct a sample of flats or households, and the sample of adults is obtained later in some other way. A list of individuals may also be missing at the implementation of area sampling: in cases when the target population is located in a special geographical area, such as a city, the frame at the first stage often consists of a list of districts, followed by a list of streets, then by a list of blocks, finally by a list of households.*3* The list of individuals may neither be available in telephone surveys: as usually random digit dialling is applied in order to overcome the under-registration problems, the respondent has to be selected within household at a second stage.*4* The difficulties of selecting a respondent within household in surveys applying a household sample can be easily overcome in cases when the respondent is uniquely defined (e.g. the head of the household) or all of the household members are interviewed. As the former largely restricts the possible target population, and the later inflates the estimation variances*5*, in the majority of cases selecting an individual with a quasi-random procedure is required.

Table 1*6* gives a summary of countries that have recently conducted different health interview surveys using different types of sampling designs. The information has been collected by the European Health Interview & Health Examination Surveys Database, a project financially supported by the European Commission. Examining Table 1, the conclusion can be drawn that in several countries both population register samples and address-based samples are feasible to implement; therefore it is justifiable to compare the advantages and limitations of their applicability.

3 Kish (1965)
4 Groves (2001)
5 Except for cases when there is seldom more than one member in a household or if within-household intra-class correlation of the measured variables is of negligible size.
6 https://www.iph.fgov.be/hishes/

Table 1. Examples of Countries Conducting Health Interview Surveys, and the Year of the Survey, by the Type of Sampling Design.

Sample of households - one respondent selected (Design B)	*Population register sample (Design A)*	*Sample of households - limited number of respondents*	*Sample of households – All persons interviewed (belonging to the target population)*
Albania (2002)	Croatia (2001)	Australia (2001)	Austria (1995 1999)
Croatia (2003)	Czech Republic (2002)	Belgium (1997 2001)	Bulgaria (2001)
France (1999)	Denmark (1994 2000)	Canada (2000)	Cyprus (2000 2003)
Germany (2003)	Estonia (1996 1999 2002)	Croatia (2003)	Czech Republic (2003)
Hungary (2000)	Finland (1997 2000 2001 2002)	France (2000 2001)	Finland (1996 2001)
Lithuania (1999)	France (1998 1999 2001 2002)	United Kingdom (2001)	Germany (1999)
Norway (1998)	Germany (1998)	United States (2000)	Hungary (2002)
Switzerland (1997 2001 2002)	Greece (1998)		Ireland (2000 2001)
United Kingdom (1993 1995)	Hungary (2000 2003)		Italy (1994 1999 2000 2001)
	Iceland (1989 1996 2000)		Lithuania (2002)
	Ireland (1998 2002)		Luxembourg (1996 2000 2001 2002)
	Latvia (1999 2003)		Portugal (1994 1995 1999)
	Luxembourg (1996 1999 2002)		Romania (2000)
	Macedonia (2001)		Slovakia (2002)
	Malta (2002)		U.K. (1997 1998 2000 2001 2002)
	Netherlands (1998 2001)		
	Norway (2002)		
	Slovakia (2002)		
	Spain (1995)		
	Sweden (1999 2001)		
	United Kingdom (1998)		

Source: European Health Interview & Health Examination Surveys Database

3 THEORETICAL CONSIDERATIONS ON REPRESENTATIVITY

As variables of interest often correlate with basic demographic characteristics (e.g. age, sex, settlement type), it is a standard practice to evaluate samples according to their representativness by these basic factors: the closer their sample distribution to the population, the more reliable the estimations and extrapolations of all kinds of variables are expected to be. When evaluating the representativity of samples obtained by design B, researchers often refer to the under-sampling of males and over-representation of elderly people, originating the phenomenon from the practical realisation of the interview, e.g. males being more difficult to find at home, and less willing to participate.

A survey on the performance of a Hungarian experimental health care programme, the Misszió Health Centre was conducted in 1998. The target population was limited to settlements concerned in the program. The implementation procedure included a two-stage sampling, with households selected by random walk in the first stage, interviewees selected in the second by using the Kish-grid method. The strong under-representation of young males as well as the over-representation of older women has been later explained as follows: “The questioning has been carried out by students of medicine and anthropology, within the frames of a summer camp. […] Despite the careful and detailed training and supervision, the students could not perform as professional interviewers would have; this is primarily blamed for the sample not corresponding to the target population along the main demographic

characteristics."*7* Another illustrative example is the case of the Health Behaviour Survey conducted in 1994 by the Hungarian Central Statistical Office: in the survey an address based sample has been applied, the respondent being selected by the closest-birthday method. According to the final research report "[…] the response rate differs along different demographic characteristics because the willingness to answer also differs among the different social strata, age-groups or sexes. On the other hand one must remember that the fieldwork was carried out in August, in the period of summer holidays, consequently, as it had been expected older women living in rural areas were over-represented. […] It must be mentioned however, that even after post-stratification did remain some bias in the sample: persons living in single households, widowers and the divorced, also the pensioners have been clearly over-represented."*8*

In what follows some theoretical evidence will be given that explain the representation problems without considering these assumptions*9* Selecting households with equal probabilities at the first stage, then selecting respondents with equal probabilities *within household* at the second stage implies varying overall selection probabilities: the chance of selecting an individual will be inversely proportional to the household size. On the other hand the size of the household may also be connected to the basic demographic characteristics of the population; if this is the case than the source of representativness problems lies in the core of the sampling design itself. In these instances, i.e. the demographic characteristics being a function of the household size, the sample would not be representative even if a perfectly random sample and 100% response rate could be obtained.*10*

As mentioned in the introductory section, Design B can be implemented using the Kish-grid. Pioneering in the exploration of advantages and limitations of address-based samples, Leslie Kish has developed his tool to assure the equality of within-household selection probabilities. When evaluating the samples drawn using the grid, Kish found a close agreement between the sample and population data concerning important demographic characteristics. His results stemmed from the fact that the household structure of the USA in the 1950's showed a high concentration within a considerably small range of household sizes: over 70% of households consisted of two adults.

According to the above findings, the household structure of the population of interest and the performance of Design B are inherently connected, hence the expected performance of Design B will vary country by country. Table 2 summarises the expected performance of Design B for several countries where Design B is in use:

7 Susánszky, Szántó (2001)
8 Józan (1996)

9 Renata Németh, working paper, 2003, www.lisproject.org

10 Németh, Rudas (2002)

Table 2. The Expected Performance of the Kish-grid

Country	*Distance between the pseudopoulation and the Kish-grid sample*	*Country*	*Distance between the pseudopoulation and the Kish-grid sample*
Italy	0.0278	Russia	0.0162
Czech Republic	0.0267	France	0.0139
Hungary	0.0241	Netherlands	0.0133
Poland	0.0235	Norway	0.0132
Slovenia	0.0222	United Kingdom	0.0127
Germany	0.0217	Australia	0.0114
Ireland	0.0186	United States	0.0091
Belgium	0.0175	Finland	0.008
Austria	0.0171	Canada	0.0072

Computations have been carried out relying on surveys included in the Luxemburg Income Study*11*. Pseudopopulation has been applied instead of population data that is the survey sample has been used as a population, based on which the sampling results have been simulated. The performance of Design B (implemented using the Kish-grid) can be assessed by a distance function measuring the difference between the expected age-sex joint distribution of the sample (matrix *a*) and the pseudopopulation (matrix *A*)*12*:

$$f(a) := \sum_{i=1,2,3\ j=1,2} (a[ij] - A[ij])^2 / A[ij].$$

According to the table, Design B performs best in Finland, Canada and the United States, while the worst results in sampling representativity are expected in Italy, the Czech Republic and Hungary.

Although some authors refer to this possible source of bias*13*, in survey practice the problem is rarely taken into consideration, the research reports do not even mention it.

4 EMPIRICAL FINDINGS – EXPERIENCE OF FOUR HUNGARIAN HEALTH SURVEYS

In the current section the above mentioned problems will be investigated based on real survey data. Primarily, the overall representativity of samples according to sex and age-groups will be examined comparing four different Hungarian health surveys: the Health Behaviour Survey 1994, the Veresegyháza Health Survey 1998, and the Health Interview Surveys 2000 and 2003. As the first two employ Design B, whereas the NHIS surveys employ Design A, a direct comparison of the two designs will be feasible to carry out. Secondarily, the non-

11 Computations are based on datasets of the Luxembourg Income Study (LIS). The LIS database is a collection of household income surveys. *Microdatabase, (1994-2000); harmonization of original surveys conducted by the Luxembourg Income Study, asbl. Luxembourg, periodic updating.*

12 age has been measured with a three-category variable: the categories consisted of 18-34, 35-64, 65+ years old persons.

13 Hader, Lynn (2003)

response factors (refusal and non-availability) usually blamed for representativity problems will be explored for the NHIS2003 conducted by our research team. Finally, relying on survey results several methodological aspects of weighting will be highlighted that may also influence the choice between the two designs. The sampling design characteristics of the four health surveys concerned are summarized in Table 3.

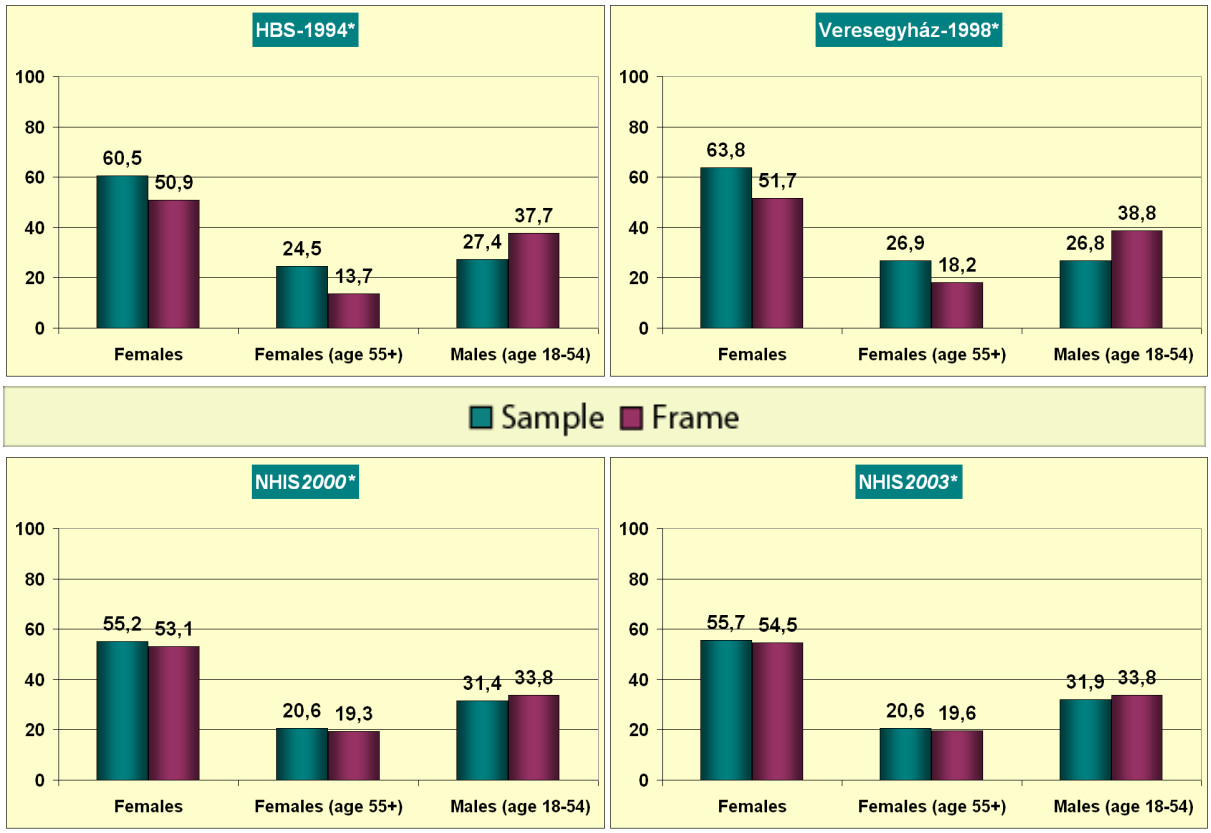
Table 3. Sampling Design Characteristics of the HBS-1994, Veresegyház-1998, NHIS*2000* and NHIS*2003* Surveys

Name of the survey, name of the institution, year of data collection	Frame	Sample design	Planned (P), realized (R) sample size	Interviewers' background
Health Behaviour Survey, Central statistical Office, 1994	Non-institutionalized Hungarian population aged 15-64	Stratification by counties Three-stage sample: 1. settlements 2. electoral districts 3. addresses Respondent selected by closest birthday-method	P.: 6411 R: 5476	Trained interviewers of the Unified System of Population Surveys
Veresegyháza Health Survey, Hungarian Academy of sciences, 1998	Population aged 18+ of the 13 villages concerned in the Mission Health Program	Number of addresses determined non-randomly in each village Addresses selected by random-walk method Respondent selected using Kish-grid	P.: 1500 R.: 1493	Trained university students of medicine and anthropology
Health Interview Survey, National Center for Epidemiology, 2000	Non-institutionalized Hungarian population aged 18 and older	Stratification by counties and settlement size Two-stage sample: 1. Settlements and the number of respondents on each determined 2. Respondents selected by simple random sampling using the electoral register	P: 7000 R.: 5503	Professional interviewers of the Hungarian Gallup Institute
Health Interview Survey, National Center for Epidemiology, 2003	Non-institutionalized Hungarian population aged 18 and older	Stratification by counties and settlement size Two-stage sample: 1. Settlements and the number of respondents on each determined 2. Respondents selected by simple random sampling using the electoral register	P.: 7000 R.: 5029	Professional interviewers of the TNS-Hungary

4.1 Comparing the Frame and the Realized Sample

Relying on the data-bases of the above described surveys representativity according to age and sex is feasible to examine. As mentioned earlier, according to research reports the over-representation of females, especially older females and the under-representation of younger males can be observed in the first two surveys. Therefore, comparison of frames and realized sample has been carried out in three subgroups of the population: females, females aged 55 and over, and males aged 18-45. (For the Health Behaviour Survey age groups of 15-49 and 50-64 were only available; the comparison, however, is still relevant.) The results illustrated on Chart 1 clearly support our theoretical considerations: using Design B, that does not take into consideration the Hungarian current household structure perform poorly regarding the subpopulation of older women and young males. However, the bias has been nearly entirely eliminated in both NHIS2000 and NHIS2003 that were implemented using a population register based sample. These findings suggest that the practical realization problems (refusals or availability) are not the only one to be blamed for sample representativity problems. The role of Design A in avoiding them may be greater then usually considered.

Chart 1. Frame and Realized Sample of the HBS-1994, Veresegyház-1998, NHIS*2000* and NHIS*2003* Surveys



* Source: NHIS*2000* Final Report, Hungarian National Centre for Epidemiology, 2002

4.2 Contrasting the Planned and the Realized Sample

The comparison of the planned and realized sample composition of the NHIS 2003 allows for analyzing non-response causes (refusal, non-availability, register error) separately. Our major findings can be summarized as follows:

Primarily, patterns of non-response observed by other researchers can be clearly recognized in case of the NHIS 2003 as well (Table 4 and Table 5): compared to the planned sample the young, especially the young males are under-represented (65% vs. 71% for young females), whereas middle-aged and older females are excessively over-represented compared to males (77% vs. 71%, and 74% vs. 72% respectively). Concerning the non-response causes separately, refusal rates for different age groups and sexes are lower for females, especially older females (8.81 percent compared to 11.01 for males); refusal rates are also lower for settlements with less than 10.000 inhabitants (5%), slightly greater (7%) for settlements with more than 10.000 inhabitants except for Budapest, while the refusal rates are the highest for the capital. (18%). Having an overall rate of 11%, non-availability referred to as "other" causes of non-response are the highest for the subgroups of the young (males: 18%, females: 12%), and are the lowest for the older males (6%) and middle aged and older females (8% both) which may also contribute to their overall under-representation in the realized sample. With respect to the settlement size, non-availability rates were the most elevated in settlements having more than 50.000 inhabitants.

Secondarily, the use of electoral registers as a list of individuals is often counter-advised referring to the increased level of errors due to under-registration and under-recorded mobility. The experience of Health Interview Survey 2000 and 2003 do not confirm these ideas: although it may be less reliable in several subgroups (e.g. error rates around 10% for the young for both sexes, and in settlements with more than 50.000 inhabitants) the overall error rate is smaller than 8%. Of course, in case of applying Design B, as there is no pre-established list of respondents, this error would not be possible to occur. However, the planned sample missing at Design B can cause other, even greater problems; the next part of the section will be dedicated to this topic.

Table 4. Non-Response Causes by Sex and Age-Groups, NHIS*2003*, row and col. percentages

	Males						***Females***						
Agegroup	***18-34***		***35-64***		***65+***		***18-34***		***35-64***		***65+***		***Total***	
succesful interview	14.62	64.59	23.32	71.35	6.4	72.4	16.37	70.89	27.64	76.5	11.65	73.71	100	71.84
register error	22.22	10.37	22.79	7.36	5.65	6.74	22.03	10.08	19.02	5.56	8.29	5.53	100	7.59
refused to answer	12.71	6.68	25.25	9.18	8.19	11	12.54	6.46	29.6	9.74	11.71	8.81	100	8.54
incapable to answer	6.9	0.53	20.69	1.09	18.39	3.6	6.9	0.52	10.34	0.5	36.78	4.03	100	1.24
other	26.89	17.84	23.97	11.01	3.71	6.29	18.54	12.06	18.54	7.71	8.34	7.92	100	10.79
Total	16.26	100	23.49	100	6.36	100	16.59	100	25.96	100	11.36	100	100	100

Source: NHIS2003, Hungarian National Centre for Epidemiology

Table 5. Non-Response Causes by Settlement Size, *NHIS2003,* row and col. percentages

Settlement size	*0-1000*		*1000-10000*		*10000-50000*		*50000+*		*Budapest*		*Total*	
succesful interview	8.05	79.41	37.24	79.7	23.13	73.6	17.58	65.97	14	57.7	100	71.84
register error	4.9	5.1	27.12	6.13	20.15	6.77	23.92	9.48	23.92	10.41	100	7.59
refused to answer	4.35	5.1	20.57	5.23	20.9	7.91	17.89	7.99	36.29	17.79	100	8.54
incapable to answer	9.2	1.57	25.29	0.94	22.99	1.27	22.99	1.49	19.54	1.39	100	1.24
other	5.96	8.82	24.9	8	21.85	10.4	26.75	15.07	20.53	12.7	100	10.79
Total	7.29	100	33.57	100	22.57	100	19.14	100	17.43	100	100	100

Source: NHIS2003, Hungarian National Centre for Epidemiology

Comparing Design A and Design B, a major aspect to discuss is the problem of respondent selection in case of the later. When population register based sampling is carried out, the precise list of respondents is set in advance, thus information on the sex, age-group, and settlement size in all planned interviews – whether successful or failed – is provided. Consequently, the planned and the realized sample are clearly separable, allowing for the weighting to correct separately for two types of errors: the sampling error and the non-response error*14*. The previous is the difference between the distribution of the planned sample and the frame, and stems from the variability caused by the random sampling method. Sampling error can be overcome by posterior stratification. On the other hand, the non-response error that is the deviation of the planned and realized sample distribution is one of the sampling factors most difficult to control; its effect can be counterbalanced by non-response correction. Table 6 shows the joint distribution of sex, age-groups and settlement types for the NHIS2003 data for the three distributions concerned: the frame, the planned and the realized sample.

In contrast to Design A, applying Design B no such list of respondents is available in advance. The addresses are previously selected, while the individuals are chosen on the spot. Therefore, the exact process of selection can not be controlled, which may formulate a potential source of error. One of the most widely known problems of this type – both in telephone and personal interview surveys - is under-coverage within household*15*: persons listing the household members often do not provide complete information. As being interested in as many successful interviews as possible, the interviewer also may manipulate the household roster. Therefore in these cases a bias may be caused regarding the representativity of the realized sample; moreover, as the mistakes of the household rosters will never be detected, the planned and the realized sample will be impossible to separate, hence no exact weighting corrections will be feasible to implement.

14 Botman (2000)
15 Groves (2001)

Table 6. Realized Sample, Planned Sample and the Frame, NHIS*2003*, percentages

	Sex	*Males*				*Females*				
	settlement size	*0-10000*	*10000-50000*	*50000+*	*Budapest*	*0-10000*	*10000-50000*	*50000+*	*Budapest*	*Total*
Age 18-34	*realized*	6.74	3.5	2.66	1.71	7.32	3.9	3.06	2.09	*30.98*
	planned	6.44	3.79	3.43	2.6	6.66	3.76	3.49	2.69	*32.84*
	frame	6.59	3.76	3.16	2.62	6.17	3.63	3.13	2.66	*31.73*
Age 35-64	*realized*	11.08	5.37	3.78	3.1	11.87	6.34	5.07	4.35	*50.96*
	planned	10.01	5.33	4.19	3.96	10.24	5.89	5.09	4.74	*49.44*
	frame	10.02	5.47	4.44	3.88	10.03	5.93	5.11	4.65	*49.54*
Age 65+	*realized*	2.82	1.59	1.11	0.87	5.47	2.43	1.89	1.87	*18.06*
	planned	2.56	1.44	1.1	1.26	4.94	2.37	1.86	2.19	*17.71*
	frame	2.89	1.47	1.22	1.32	4.95	2.45	2.05	2.38	*18.73*
Total	*realized*	*20.64*	*10.46*	*7.56*	*5.69*	*24.66*	*12.67*	*10.02*	*8.31*	*100*
	planned	*19.01*	*10.56*	*8.71*	*7.81*	*21.84*	*12.01*	*10.43*	*9.61*	*100*
	frame	*19.5*	*10.7*	*8.82*	*7.82*	*21.15*	*12.02*	*10.29*	*9.69*	*100*

Source: NHIS2003, Hungarian National Centre for Epidemiology

4.3 Weighting effects

It is widely accepted that the smaller the differences between the estimations with and without weighting, the more reliable are the survey results. The stability of estimations indicates a balanced sample and a low level of „artificial intervention" on the part of researchers. In what follows, we will compare these differences for a few health indicators for the HBS1994 and NHIS2000 surveys. (The reason of our choice was the nearly equal sample sizes of these two surveys.) Considering the above findings, HBS1994 that uses design B is expected to have weights of greater variability compared to NHIS2000; our empirical findings summarized in Table 7 and Table 8 are consistent with these presumptions.

According to the NHIS2000 table, weighting have changed the category frequencies only by decimals, and the standard errors have also stayed almost unaltered. An opposite pattern can be observed in Table 8, differences were notably greater between the estimates with and without weights: the proportion of smokers has increased by 4 percent, while the occurrence of long-term disease has dropped with 5 percent. The standard errors have also been raised with a greater amount than in the case of NHIS2000.

Table 7. Changes in Estimates Due to Weighting, NHIS*2000*

	Without weights		***Weighted***	
Smoking	***Estim. proportion***	***Std. Error***	***Estim. proportion***	***Std. Error***
regularly, at least one pack per day	16.3	0.5	16.76	0.52
regularly. less than one pack per day	13.39	0.46	13.4	0.47
occasionally	2.8	0.22	2.84	0.23
gave up smoking	20.24	0.54	20.24	0.55
never smoked	47.27	0.67	46.77	0.68
Activity restriction caused by mental health problem	13.4	0.46	13.12	0.46

Source: NHIS2000, Hungarian National Centre for Epidemiology

Table 8. Change in Estimates Due to Weighting. HBS-1994

	Without weights		*Weighted*	
Smoking	***Estim. proportion***	***Std. Error***	***Estim. proportion***	***Std. Error***
currently smoking	0.313	0.0063	0.3501	0.0077
gave up smoking within a year	0.0143	0.0016	0.0179	0.0025
gave up smoking more than a year ago	0.0963	0.004	0.0932	0.0044
never smoked	0.5764	0.0067	0.5388	0.008
Having a long term disease	0.3201	0.0063	0.2722	0.0068

Source: HBS-1994, Central Statistical Office, Hungary

5 CONCLUSIONS

In cases when the implementation of both design A and design B is feasible, it may be worth considering the above mentioned aspects. We did the same in 2002 when the HNCE joined the WHO World Health Survey covering 70 countries. Although the WHO sampling plan suggested design B with the Leslie Kish grid as a sampling design, referring to the above findings and also to the existence of a proper quality list of the Hungarian adult population we suggested the implementation of design A instead; the proposal has been accepted by the WHO.

REFERENCES

Boros, J., Csizmadia, P., Hermann, D., Kéki, Zs., Kiss, G., Németh, R., Országh, S., Vitrat, J., Vizi, J., Zakariás, I. (2004) *National Health Interview Survey 2003* – Hungary, National Centre for Epidemiology.

Botman S.L., Moore T.F., Moriarity C.L., and Parsons V.L. (2000) Design and Estimation for the National Health Interview Survey, 1995-2004. National Center for Health Statistics. *Vital and Health Statistics 2*(130).

Groves, R.M., Biemer, P.P., Lyberg, L.E., Massey, J.T., Nicholls, W.L., Waksberg, J. (2001) *Telephone survey Methodology*. New York: John Wiley and sons, Inc.

Józan, P., Gárdos, É., Juhász, J., Éltető, Ö. (1996) *Health Behaviour Survey, 1994*. Central Statistical Office, Budapest,

Kish,L. (1965) *Survey Sampling*. New York: John Wiley and Sons, Inc.

Németh, R., Rudas, T. (2002) Sample selection with the Kish grid. *Statisztikai Szemle, 80*, 309-327.

Németh, R.: (2003) Respondent Selection within Household – A Modification of the Kish-grid, working paper, www.lisproject.org

Susánszky, É., Szántó Zs. (2001) *Táblaképek az egészségről – A veresegyházi példa*, Budapest, Magyar Tudományos Akadémia.

Vokó, Z., Vitrai, J., Boros, J., Borsos, K., Csizmadia, P., Lobmayer, P., Németh, R., Országh, S., Széles, Gy., Szende, Á., Vizi, J. (2002) *Országos Lakossági Egészségfelmérés OLEF2000*, Budapest: Országos Epidemiológiai Központ.