

Postprint Version	1.0
Journal website	http://www.sciencedirect.com/science/journal/13538292
Pubmed link	http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=pubmed&dopt=Abstract&list_uids=15177196&query_hl=33&itool=pubmed_docsum
DOI	10.1016/j.healthplace.2003.07.002

Medical practice variations in hospital care; time trends of a spatial phenomenon

G.P. WESTERT^{A,*}, P.P. GROENEWEGEN^B, H.C. BOSHUIZEN^A, P.M.M. SPREEUWENBERG^B, M.P.M. STEULTJENS^B

^a National Institute for Public Health and the Environment (RIVM), P.O. Box 1, Bilthoven 3720 BA, Netherlands

^b Netherlands Institute for Health Services Research (Nivel), Netherlands

*Corresponding author. Tel.: +31-30-274-2470; fax: +31- 30-274-4407. E-mail address: gert.westert@rivm.nl (G.P. Westert).

ABSTRACT

A persistent finding in health services research is that health care delivery and hospital utilisation in the Western world varies widely between areas, both within and between countries. Most studies have concentrated on crosssectional variations in medical practice. The aim of this article is to investigate whether or not small area variation changed through time. We used hospital discharge rates in the Netherlands for 12 diagnostic or surgical categories to indicate medical practice patterns. The data cover a time span of almost two decades: 1980–1997. First, it was found that in most cases regions are consistently above or below the national trend in the study period. Second, the analysis revealed a statistically significant decline of regional variation in hospital discharges in general during the 1980s and the 1990s. In all but one medical category the results of the separate analyses point towards a downward trend. In one-third of the medical categories this downward trend was statistically significant. Potential parallel changes in regional disparities in need for care, e.g. morbidity or age composition of regional populations or changes in regional differences in care supply are discussed.

INTRODUCTION

A persistent finding in health services research is that health care delivery and hospital utilisation in the Western world varies widely between countries and within countries between areas (Wennberg and Gittelsohn, 1973; Ham, 1988; Folland and Stano, 1990; Keller et al., 1998). It was found that the actual use of surgical procedures is, in most cases, quite different between countries, while the small-area variation within countries for each procedure is quite consistent (e.g. appendectomy, cholecystectomy, hysterectomy) (McPherson, 1990). These differences between areas within countries do not vanish when differences in need related factors, like age, gender, and socio-economic variables between patient populations have been controlled for. In his study, McPherson concludes that at small area level the most plausible sources of medical practice variation is clinical decision making and not differences in morbidity patterns or patient demand factors (McPherson, 1990).

Most studies on small area variation have concentrated on cross-sectional analysis of this spatial phenomenon (Ashton et al., 1999). The research question whether or not small area variations changed through time received only little attention. This article focuses on this question. Studies that explicitly explored time trends of spatial variation in medical practice are rare. Via 'MEDLINE' we traced five studies that (in) directly analysed time trends of spatial variation of physician practice patterns

(Wennberg et al., 1977; Sejr et al., 1991; Lu-Yao and Greenberg, 1994; Katz et al., 1996; Chan et al., 1997). We briefly review these studies first. Wennberg et al. reported in 1977 that among 13 Vermont Hospital Service Areas, variation in tonsillectomy rates decreased over a five-year period, between 1969 and 1973 (Wennberg et al., 1977). The authors associate the changing surgical practices with feedback on discharge rates from local communities. Sejr et al. studied temporal and regional variations in the use of prostatectomy from 1977 to 1985 (Sejr et al., 1991). The authors observed that the total annual number of prostatectomies increased by 43% during the period, when the transurethral procedure (TURP) gradually replaced traditional open surgery. Substantial regional variations occurred in the recruitment areas. The variation decreased during the process of technology diffusion, but remained at a relatively high level. Lu-Yao et al. found a five-fold variation among areas in radical prostatectomy rates in 1989, with a low of 43.4 per 100,000 in Connecticut and a high of 224.4 in Seattle. The differences between the areas grew larger over time. Prostate cancer mortality rates did not increase during the period of study; there was little variation among areas in prostate-cancer mortality rates, and no apparent correlation between the incidence and mortality rates for an area (Lu-Yao and Greenberg, 1994). Katz et al. observed that the number of knee replacements increased between 1985 and 1990 corresponding to an annual increase of 18.45% and that regional variation decreased only slightly (Katz et al., 1996).

Chan et al. conclude that after the rapid growth in spirometry utilisation between 1989 and 1995 regional variation during the same period decreased, but substantial spatial disparities remained in the study period (Chan et al., 1997). The authors point at the incomplete implementation of spirometry guidelines.

Ever since the introduction of ‘small area variation in health care delivery’ as a problem by the research community (Ham, 1988), physicians have been confronted with questions why they differ in practice style in their treatment of patients. Medical practice has the image of being based on solid scientific grounds, but nothing more underscores the reality that medical care represents a melding of art and science than the wide variation in the use and the cost of that care (Andersen and Mooney, 1990; Eddy, 1984). We know that there are local standards of care and that doctors that share the same work environment tend to conform to the *couleur locale* (Westert and Groenewegen, 1999). But the question that remains is why do local standards exist and dominate the gold standards that the public believes doctors apply. Public awareness about medical practice variation and surgical signatures has grown because researchers directly or indirectly reported in the mass media (Domenighetti et al., 1988). This situation, combined with the emphasis on cost control in the health care sector—in most Western countries from the early 1980s—pushed the medical community towards rationalisation and standardisation of medical care (Lomas et al., 1989; Ritzer and Walczak, 1988). Since then, critical pathways, consensus statements, guidelines and standards for appropriate care have been developed and have become available to medical doctors. Simultaneously, a growing number of knowledge sources have become available to the consumer. Patients are nowadays, more than ever, able to check if doctors conform to local practices or standards instead of national or global standards.

The question arises whether or not medical practice variations have changed in the past decades. Given the above we would expect a downward trend in spatial variation because contextual pressure on doctors steered them to more uniformity in decision making. However, the expected decrease could also be caused by parallel changes in regional disparities in need (e.g. morbidity and age composition of regional populations) or care supply. Furthermore, the trend towards less variation needs not to be linear. As some of the literature suggested, processes of diffusion of new treatments or new insights might cause an alternating pattern of increasing and decreasing variation (Sejr et al., 1991).

The aim of this article is to explore time trends of small area variation in medical practice. We use regional hospital discharge rates in the Netherlands for 12 diagnostic or surgical categories, as indicator of medical practice patterns. The data cover a time span of almost two decades: 1980–1997.

METHODS

Dutch hospital discharge data for the years 1980 through 1997 were analysed¹, covering 18 points in time. The study population covers all Dutch residents. ICD-9 diagnosis or procedure specific

¹Data provided by Prismant, Utrecht.

discharge data was obtained in aggregated age–gender format for 25 Dutch Health Care Regions (DHCR). Patients were counted among the region of residence. Border crossing between region of residence and region of care is not significant in the study, since only a small percentage of the cases—involved in the study—receive hospital care outside their DHCR region of residence. In 1990 the average number of hospitals per DHCR was approximately 6 and the average population size half a million. For every single year, age and gender standardised hospital discharge rates were calculated per 10,000 residents and per medical condition separately. In total 12 conditions were selected: nine medical diagnoses and three surgical procedures. The main criterion of selection is frequency of occurrence. The conditions chosen are among the high volume hospital discharge and surgical rates in the Netherlands. This criterion is important because the area variation is overestimated if the population is too small and rates of the intervention under study is low (Coyte et al., 2001). There was no selection of the conditions on variability on forehand. To avoid incompleteness in the data conditions were also selected on the criterion whether or not there is an alternative treatments setting (e.g. outpatient clinics or day surgery). Our selection is diagnoses and surgical procedures that were in the period studied solely treated as inpatient cases.

Table 1 shows the data for 1980, 1989 and 1997. In Table 1 we added the coefficient of variation (CV) as a first descriptive indicator of dispersion. CV measures the standard deviation of regional discharge rates to the mean rate (multiplied by 100). An important question is what to use as a measure for variation. It is sensible to assume that the amount of variation in an absolute sense is larger when the rate itself is larger. This often means one assumes that the amount of variation is proportional to the discharge rate. A decline in variation then is defined as a decline in the coefficient in variation. Alternatively, one could define a decline in variation as a decline in the standard deviation itself (the absolute measure of variation). In the data, all discharge rates with the exception of that for diabetes mellitus and lung cancer, remained stable or increased during our study period. Thus, using the absolute measure as our criterion for decreasing variation is stricter than using the relative criterion. We choose to be conservative in the statistical analysis and use an absolute measure, as in that case we can be sure that any effect observed does not depend on the criterion chosen. However, as descriptive results we will present the coefficient of variation also.

We also inspected to what extent single regions produced rates consistently above or below the average national rate in the years between 1980 and 1997 and whether or not there are differences between the 12 conditions. Thereto autocorrelations were computed per region and per condition. A high autocorrelation for region R1 and condition C1 means that the discharge rate of R1 is systematically above or below the national trend.

The statistical analysis was performed with multilevel or hierarchical linear techniques (Snijders, 1996). The reason for this is that the 18 yearly discharge rates within each region can not be assumed to be independent measurements within the entire set of rates. In fact, the yearly rates are nested within regions. One can imagine that rates in region X are in general systematically higher or lower than those in region Y . Ignoring this correlation, as would be done when ordinary regression would be used, would give biased results. Theoretically, there is a third level; that of inhabitants of the regions who can be admitted or not. However, we did not have access to individual level data.

The trend analysis was performed using the following two-level model:

$$r_i(t) = \mu_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3 + \beta_4 t^4 + \beta_5 t^5 + \mu_i + \beta_i t + s_{it} \varepsilon_{it},$$

where $r_i(t)$ is the age and gender standardised hospital discharge rate for a disease in region i and year t ; t is the year number (1988=0); $\beta_1 t, \dots, \beta_5 t^5$ a polynomial function of time, included to capture the general (national) trend in discharge rate over years and s_{it} is the standard error of $r_i(t)$ as calculated from the age and sex specific numbers of discharges, assuming a Poisson distribution. Including s_{it} makes this procedure comparable to McPherson's approach to separate random variation from systematic variation (McPherson, 1990).

Finally, μ_i , β_i and ε_{it} are normally distributed with mean 0 and standard deviation $\sigma\mu$, $\sigma\beta$ and $\sigma\varepsilon$; respectively (standard assumptions in multilevel modelling).

In the first analytical step, the twelve medical categories were simultaneously used in the following operational model:

$$z_{ik}(t) = \mu_0 + \beta_{1k}t + \beta_{2k}t^2 + \beta_{3k}t^3 + \beta_{4k}t^4 + \beta_{5k}t^5 + \mu_i + \beta_i t + s_{ikt}\varepsilon_{ikt},$$

where $z_{ik}(t)$ is the z -score of the standardised hospitalisation rate for medical category k in region i and year t ; and s_{ikt} the standard error of $z_{ik}(t)$. The z score is calculated as: the difference of the observed rate and the average rate for medical category k ; divided by the standard deviation (calculated for all observations in medical category k). We used z -scores in order to include the rates from different medical categories on a comparable scale.

In these models, the time trend of spatial variation is given by the correlation of the intercept μ_i and the slope with time β_i , at level 2. If the correlation is negative and statistically significant, this means that the regional disparities have become smaller over time.

In the second analytical step a similar two level model was run for each medical category separately, in order to inspect the differences between the twelve medical categories.

RESULTS

Firstly, we inspected the extent to which particular regions are consistently above or below the national trend. The autocorrelation—averaged over the conditions—ranged from 0.54 to 0.81 between DHCRs, with an average of 0.73. This means that in general regions show a consistent pattern between 1980 and 1997: regions that are high on one time point stay high (or above the average) in neighbouring years. We also observed that the average autocorrelation per condition varied from 0.50 (stroke, breast cancer) to 0.92 (total hip).

For illustrative purposes, Fig. 1 shows the location of the 23 health care regions and indicates which parts of the country have high numbers of hospital discharges and where numbers are lower. The map shows four regions (black) that have a high average rank number (ranging from 5.1 to 9.0) calculated for the 12 conditions. These regions have the highest numbers of hospital discharges. Five regions (light grey), clustered in the western part of the country, have low average ranks, ranging from 14.8 to 21.6. In these regions relatively low numbers of hospital discharges were observed for the 12 conditions. The map describes the situation in 1989. Given the results from the autocorrelation analysis the spatial pattern in 1989 can be seen as representative for the years between 1980 and 1997.

Table 1 shows per medical category and for a selection of time points (1980, 1989 and 1997) the coefficient of variation (CV). Apart from diabetes mellitus and lung cancer, discharge rates remained stable or increased during the 1980s and 1990s. The hospital discharge rate of diabetes mellitus dropped from 13.5 to 7.6 per 10,000. In ten of the medical categories the coefficient of variation between 1980 and 1997 decreased. Only in the case of lung cancer and rheumatoid arthritis the coefficient of variation did not decrease. However, in the multilevel based approach (below) where the standard deviation is used, the case of lung cancer showed the hypothesised decline in variation.

The result from the multilevel analysis is that the general model's correlation—with all medical categories included—of regional disparities and time was -0.56 ($\chi^2 = 4.04$; $p = 0.04$), meaning that the general model points at a statistical significant downward trend in regional variations.

Fig. 2—as an example—shows the time trend and course of regional variation for appendectomies in detail. The lines with 'black-boxed' markers mark the range of regional variation per year. In 1980 the range of discharge rates between the 23 regions was approximately 10 per 10,000: from 9 to 19 per 10,000 population-at-risk. The average discharge rate in 1980 was 14 per 10,000. The grey line—a linear function of the observed coefficient of variation scores—shows the development of the coefficient of variation through time. This downward trend is also shown by the diminishing range between 1980 and 1997.

Table 1, final column, shows the association of regional variation with time per medical category based on the separate multilevel models. Firstly, note that in all but one case the signs are negative. In four cases the estimated trend is negative and statistically significant ($p < 0.05$). In four cases the estimated correlation is statistically borderline ($p < 0.10$). In the rest of cases the estimated negative correlation is not statistically significant.

DISCUSSION

Our analysis revealed a significant decline of regional variation in hospital discharges in general during the eighties and the nineties. In all but one medical category the results of the separate analyses point towards a downward trend. In one-third of the medical categories this downward trend was statistical significant.

In the introduction we stated that the observed change could also be caused by parallel changes in regional disparities in need for care, e.g. morbidity or age composition of regional populations or changes in regional differences in care supply. To check this we performed additional analyses.

Firstly, we investigated whether or not spatial disparities in need, indicated by the percentage of elderly (75 years and older), showed a similar downward trend. In 1980 the average percentage of elderly was 4.5 and increased to 5.8 in 1997, but during the research period the range and standard deviation of the regional distribution per year remained the same. The yearly range was 0.3 the standard deviation 0.008.

Secondly, we also looked at a more specific indicator for need, the percentage of people in a region with at least one chronic disease, using pooled health survey data for 1989–1994 and 1995–2000.¹ Between the first and second time point the national percentage of persons with at least one chronic condition increased from 36.2 to 39.4. The variation in percentage of persons reporting at least one chronic disease between regions is small and declined slightly between 1989/1994 and 1995/2000. The coefficient of variation changed slightly: from 6.1 to 4.5 and the range between regions changed from 10.3 to 7.7.

Thirdly, the observed changes in spatial variation of hospital discharges between different points in time could also be caused by changes in health care supply. Therefore, we investigated the time trends in regional disparities in bed supply and average hospital size (beds) per DHCR in 1980, 1990 and 2000¹. The number of hospital beds in the Netherlands declined from 68.300 in 1980 to 55.229 in 2000 and the average hospital size increased from 390 beds in 1980 to 575 beds in 2000, caused by hospital merges. The number of hospitals changed from 187 in 1980 to 104 in 2000. The regional variation in average hospital size remained quite similar (coefficient of variation: 29.9, 23.7, 28.4). Differences in regional bed supply changed slightly between 1980 and 2000. The coefficient of variation decreased from 53.3 to 46.5.

A decline in small area variation in medical procedures might be interpreted within a framework of the increasing evidence base of medical practice. The international literature on changes over time in small area variation is scarce. The few studies we traced, give the following insights. Firstly, feedback of information comparing performances among neighbouring areas enhances diffusion of state-of-the-art opinion and promote greater uniformity in decision making (Wennberg et al., 1977). In later studies it is concluded that downward trends in practice variation caused by feedback to physicians is often temporary and practice style variations are difficult to change (Andersen and Mooney, 1990; Eddy, 1984). Secondly, differences in phase of diffusion of medical knowledge or new techniques lead to spatial variation in medical practice (Sejr et al., 1991; Lu-Yao and Greenberg, 1994; Katz et al., 1996; Chan et al., 1997). Thirdly, the existing studies used very limited time intervals or only a few number of time points.

In our study we found a downward trend in regional variation in diagnosis/procedure specific discharge rates over a nearly 20 year time period. It is tempting to interpret this as a result of increasing rationality of medical practice. However, two other trends—decreasing variation in the percentage of people with selfreported chronic disease and in the supply of hospitals and hospital beds—point towards a broader trend in western societies. In the course of a long term process of modernisation, regional differences tend to disappear gradually in numerous fields of social and economic activity. The unification of medical practice through scientific progress and implementation through standards and protocols is part of this broader process. It might, however, also be influenced by parallel changes in the broader context of medical practice. Where different local circumstances may lead to localised standards of what is good medical practice, diminishing differences in local circumstances also make the differences between localised standards smaller.

What are the broader implications of this line of reasoning for the study of medical practice variations? Firstly, we would hypothesise that the international differences in medical practice also have decreased over the past decades. Medical science is increasingly international and medical technology is part of the globalised economy. This is not to say that international differences have

disappeared. Comparative analysis of guidelines used in different countries still shows large differences (Vader, 2001; Tuut et al., 2002). Medical guidelines are still ‘localised’ but differences are only visible on a bigger scale. Secondly, besides the broader societal process of unification, new sources of diversity do appear, that might have influence on localised standards. One could think of the inflow of migrants and the tendency of segregation within cities. Seen from the user side, this might lead to variations in utilisation patterns of health care, and from the side of medical practice to new practice variations.

TABLES AND FIGURES

Table 1

Hospital discharges per 10.000 population-at-risk in DHCR of residence and coefficient of variation (between DHCR) per diagnostic category (ICD9) or surgical procedure in 1980, 1989 and 1997 and (multilevel model based) correlation with time between 1980 and 1997

Medical discharge diagnosis (ICD9 code) or surgical procedure	1980	1989	1997	1980–1997
	Discharge rate/ coefficient of variation	Discharge rate/ coefficient of variation	Discharge rate/ coefficient of variation	Correlation
Breast cancer (ICD 174)	12.2/19.4	14.1/16.9	14.5/10.0	-0.92**
COPD (ICD 490–492, 495 and 496)	8.6/32.9	12.2/26.0	12.7/24.5	-0.48*
Coronary heart disease (ICD410–414)	35.6/19.4	47.0/14.7	50.0/13.3	-0.49*
Diabetes mellitus (ICD 250)	13.5/22.0	11.4/16.8	7.6/20.2	-0.87**
Lung cancer (ICD 162)	15.6/16.9	16.0/17.0	12.4/19.7	-0.54*
Osteoarthritis (ICD 715)	7.3/15.2	11.1/13.5	14.8/9.7	-0.51*
Pneumonia (ICD 480–486)	6.2/24.6	7.7/25.7	11.4/14.8	-0.30
Rheumatoid arthritis (ICD 714)	2.6/22.2	3.0/35.2	2.6/48.1	0.00
Stroke (ICD430–438)	15.5/17.4	15.7/13.6	16.6/11.0	-0.75**
Appendectomy	13.9/17.6	11.7/16.1	10.4/12.8	-0.88**
Caesarean section	24.1/17.7	36.4/19.9	52.8/11.3	-0.37
Hip replacement	5.0/17.3	7.9/13.2	10.8/8.8	-0.32

** $p < 0.05$, * $p < 0.10$.

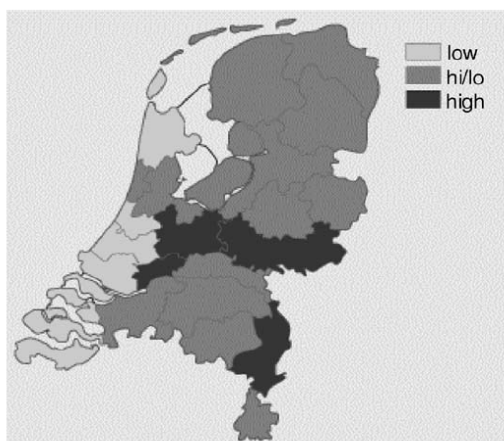


Fig. 1. The average ranking of Dutch Health Care regions for hospital discharge rates of 12 medical conditions (1989).

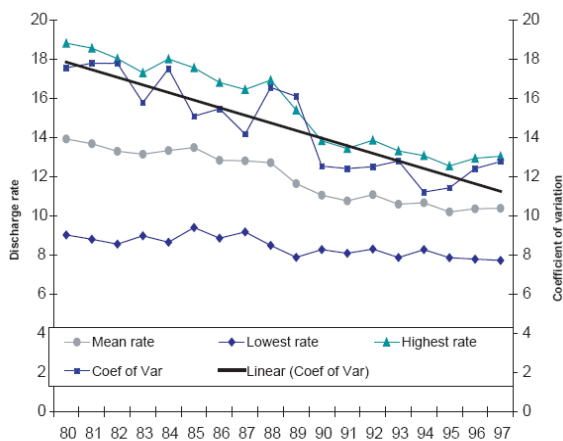


Fig. 2. Regional variation of hospital discharges per 10,000 population-at-risk for appendectomy between 1980 and 1997.

REFERENCES

1. Andersen, T.F., Mooney, G., 1990. *The Challenges of Medical Practice Variations*. MacMillan Press, London.
2. Ashton, C.M., Petersen, N.J., Soucek, J., et al., 1999. Geographic variation in utilisation rates in veterans affairs hospitals and clinics. *New England Journal of Medicine* 340, 32–39.
3. Chan, B., Anderson, G., Dales, R.E., 1997. Spirometry utilization in Ontario: practice patterns and policy implications. *Canadian Medical Association Journal* 156 (2), 169–176.
4. Coyte, P.C., Croxford, R., Asche, C.V., To, T., Feldman, W., Friedberg, J., 2001. Physician and population determinants of rates of middle-ear surgery in Ontario. *Journal of the American Medical Association* 286, 2128–2135.
5. Domenighetti, G., Luraschi, P., Casabianca, A., Gutzwiller, F., Spinelli, A., Pedrinis, E., Repetto, F., 1988. Effect of information campaign by the mass media on hysterectomy rates. *Lancet* 2 (8626–8627), 1470–1473.
6. Eddy, D.M., 1984. Variation in physician practice: the role of uncertainty. *Health Affairs* 2, 74–89.
7. Folland, S., Stano, M., 1990. Small area variations: a critical review of propositions, methods, and evidence. *Medical Care Review* 47, 419–465.
8. Ham, C., 1988. A review of the literature. In: Ham, C. (Ed.), *Health Care Variations; Assessing the Evidence Research Report No 2*. King's Fund Institute, London.
9. Katz, B.P., Freund, D.A., Heck, D.A., Dittus, R.S., Paul, J.E., Wright, J., Coyte, P., Holleman, E., Hawker, G., 1996. Demographic variation in the rate of knee replacement: A multi-year analysis. *Health Service Research* 31 (2), 125–140.
10. Keller, R.B., Largay, A.M., Soule, D.N., Katz, J.N., 1998. Maine Carpal tunnel study: small area variations. *Journal of Hand Surgery* 23, 692–696.
11. Lomas, J., Anderson, G.M., Domnick-Pierre, K., Vayda, E., Enkin, M.W., Hannah, W.J., 1989. Do practice guidelines guide practice The effect of a consensus statement on the practice of physicians. *New England Journal of Medicine* 321, 1306–1311.
12. Lu-Yao, G.L., Greenberg, E.R., 1994. Changes in prostate cancer incidence and treatment in USA. *Lancet* 343 (8892), 251–254.
13. McPherson, K., 1990. Why do variations occur? In: Andersen, T.F., Mooney, G. (Eds.), *The Challenges of Medical Practice Variations*. MacMillan Press, London, pp. 16–35 (Chapter 2).
14. Ritzer, G., Walczak, D., 1988. Rationalization and the deprofessionalization of physicians. *Social Forces* 67, 1–22.
15. Sejr, T., Andersen, T.F., Madsen, M., Roepstorff, C., Bilde, T., Bay-Nielsen, H., Blais, R., Holst, E., 1991. Prostatectomy in Denmark. Regional variation and the diffusion of medical technology 1977–1985. *Scandinavia Journal of Urology and Nephrology* 25 (2), 101–106.
16. Snijders, T., 1996. Analysis of longitudinal data using the hierarchical linear model. *Quality and Quantity* 30, 405–426.
17. Tuut, M.K., Van Tulder, M.W., Van Everdingen, J.J.E., Sanders, R., Assendelft, W.J.J., 2002. Aspecifieke lage rugklachten. *Medisch Contact* 57, 1016–1019.
18. Vader, J.P., 2001. Evaluating the appropriateness of low back pain care. *Schweizerische Rundschau für Medizin Praxis* 43, 1863–1868.
19. Wennberg, J.E., Gittelsohn, A., 1973. Small area variations in health care delivery. *Science* 182, 1102–1108.
20. Wennberg, J.E., Blowers, L., Parker, R., Gittelsohn, A.M., 1977. Changes in tonsillectomy rates associated with feedback and review. *Pediatrics* 59 (6), 821–826.
21. Westert, G.P., Groenewegen, P.P., 1999. Medical practice variations: changing the theoretical approach. *Scandinavia Journal of Public Health* 27, 173–180.