

RESEARCH REPORT

Socioeconomic status of very small areas and stroke incidence in the Netherlands

J Smits, G P Westert, G A M van den Bos

J Epidemiol Community Health 2002;56:637-640

See end of article for authors' affiliations

Correspondence to:
Dr J Smits, Department of
Economics, Nijmegen
School of Management,
PO Box 9108, 6500 HK
Nijmegen, Netherlands;
jeroen.smits@nsm.kun.nl

Accepted for publication
17 December 2001

Objective: To examine whether characteristics of very small living areas can be used to predict disease incidence and to use these characteristics to assess socioeconomic differences in stroke incidence in the Netherlands.

Design: Characteristics of postcode areas of stroke patients are compared with characteristics of postcode areas of all individual people in the study region, using Poisson regression analysis.

Setting: Six provinces of the Netherlands, covering about half of the country.

Patients: 760 patients who in 1991 or 1992 were consecutively admitted because of stroke to 23 Dutch hospitals.

Main results: Stroke incidence is significantly higher among people living in postcode areas with below average socioeconomic status (relative risk=1.27; 95% confidence intervals 1.08 to 1.51) and among people living in postcode areas with predominantly older inhabitants (RR=3.17; 95% CI=2.29 to 4.39). It is also significantly increased in more urbanised areas compared with the countryside, the highest incidence being found in the large cities (RR=1.78; 95% CI=1.31 to 2.44).

Conclusions: A clear socioeconomic gradient in stroke incidence in the Netherlands is observed, with people living in detailed postcode areas with below average socioeconomic status experiencing a significantly higher risk of stroke. The analysis also confirms that characteristics of detailed postcode areas can effectively be used to differentiate between areas with and areas without stroke patients.

In recent decades, the incidence and prevalence of many diseases have been found to differ among socioeconomic groups.¹⁻⁵ People with lower socioeconomic status are generally less healthy and show higher mortality rates than people with higher socioeconomic status. Such effects of socioeconomic status have also been reported for stroke. The incidence and prevalence of stroke and other manifestations of cardiovascular diseases and the risk of mortality from stroke is considerably higher for patients from lower socioeconomic groups.⁶⁻¹¹

Research into socioeconomic status effects on health is often based on small or non-representative samples or on samples of subgroups of the population. The reason for this is that in large scale registrations of disease and mortality information on the individual socioeconomic status of patients is generally lacking. There are a few countries (for example, Finland, Norway) where registration information can be enriched with individual information from population censuses or other data sources. However, for many countries this solution is not feasible and the only remaining possibility is to use socioeconomic information on the geographical area in which the patient is living. Several indices have been developed that indicate the average socioeconomic status of small areas, like postcode or enumeration districts, on the basis of available (mostly census) information on characteristics of people and households living in those areas.¹²⁻¹⁵ With regard to stroke, there are several studies that show that socioeconomic deprivation of small areas is associated with raised mortality and higher risk of hospitalisation.^{9, 12, 16, 17} In fact, most of the early studies on effects of socioeconomic status on stroke incidence and mortality were ecological studies.¹¹

A disadvantage of the deprivation indices normally used in this kind of study is that the areas on which they are based are rather large. In most cases the number of inhabitants is several thousands or more. As a consequence, the risk of making the ecological fallacy—the error made when characteristics of areas are assigned to individuals¹⁸—is rather high. If socioeconomic status effects are found using these indices, it is unclear

to what extent these effects reflect the effects of individual socioeconomic status and to what extent they reflect effects of neighbourhood deprivation. Moreover, in as far as individual socioeconomic status effects are caught by these measures, they probably underestimate the real effects heavily, because areas of this size may be rather heterogeneous with regard to socioeconomic status of their inhabitants. Given the fact that smaller areas are on average less heterogeneous, the approximation of the effects of individual socioeconomic status can be expected to be better when the area for which aggregated information is available is small.^{19, 20}

In this paper, the relation between socioeconomic status and stroke incidence is studied for the Netherlands using aggregated information on socioeconomic status for the smallest possible areas for which this information is available: the Dutch detailed (6 position) postcode areas. The average number of households living in these areas is a mere 15. This is much less than in previous studies using area based indices, and hence the approximation of individual socioeconomic status is expected to be much better.* The aim of this paper is twofold. On the one hand we would like to answer the methodological question of whether the incidence of disease can be solidly predicted on the basis of characteristics of the very small living areas individuals reside in. And on the other hand, we want to use the available information on socioeconomic status of these small areas to assess the socioeconomic differences in stroke incidence in the Netherlands.

METHODS

Data

The data used in this study were gathered for a multicentre study on quality of care after stroke in the Netherlands.²¹⁻²³ The

*The Dutch postcode areas are probably more homogeneous than those of most other countries because natural boundaries were explicitly taken into account when they were constructed.

Table 1 Poisson regression model for the effect of selected postcode characteristics on the incidence of stroke within the postcode areas; relative risks (RR) with 95% confidence intervals (CI)

Postcode area characteristics	RR (95% CI)
Income level	
Average or above	1
Below average	1.27 (1.08 to 1.51)
Diverse	1.31 (0.88 to 1.95)
Family phase	
Young	1
Middle	1.18 (0.85 to 1.63)
Old	3.17 (2.29 to 4.39)
Diverse	2.19 (1.15 to 4.14)
Urbanisation	
Countryside (<5000)	1
Town (5000–20000)	1.43 (1.07 to 1.92)
Small city (20000–100000)	1.44 (1.08 to 1.90)
Middle city (100000–250000)	1.50 (1.05 to 2.12)
Large city (>250000)	1.78 (1.31 to 2.44)
Province*	
Groningen	2.18 (1.75 to 2.72)
Friesland	0.91 (0.67 to 1.25)
Drenthe	1.57 (1.19 to 2.08)
Utrecht	0.39 (0.28 to 0.54)
Zuid Holland	0.39 (0.32 to 0.48)
Noord Holland	2.14 (1.84 to 2.48)
Intercept	0.00
Stroke (n)	651
Total (n)	8316461

*For this variable, deviation from average coding is used.

study group consisted of 760 patients who in 1991 or 1992 were consecutively admitted because of stroke to one of 23 Dutch hospitals. Of these hospitals, 17 were located in the densely populated mid-western part of the Netherlands called the Randstad, and six in the more sparsely populated north eastern part of the Netherlands. Within these regions, the hospitals were selected randomly after stratification for hospital size. Five small hospitals (<200 beds), nine intermediate hospitals (200–400 beds), and nine large hospitals (>400 beds) participated. The patients were traced by use of the administrative records of the hospitals. Patients were considered to have had a stroke, if there was a focal neurological deficit of sudden onset that lasted at least 24 hours with no known alternative to a vascular cause. Both first and recurrent strokes were included. Patients with a transient ischaemic attack (duration of symptoms shorter than 24 hours), a subarachnoid haemorrhage, or non-stroke pathology were excluded. All patients satisfying the inclusion criteria who were admitted to the hospital during a period of half a year received a letter from the hospital in which they were asked to participate in the study. Of these patients, 760 (96%) gave permission to the researchers to use their hospital records for the study.

With only a few exceptions, the patients were living in the Randstad provinces Utrecht, Noord Holland, and Zuid Holland, and in the north eastern provinces Groningen, Friesland, and Drenthe. The analysis, therefore, is restricted to these six provinces, which together cover about half of the Netherlands. The total number of postcode areas in these provinces is 214 795 and the total number of people living in these provinces is 8 348 149. In a Dutch postcode area there live on average about 38 people (standard error=30), spread over about 15 households. We use postcode information for the year 1992.

Patients living in institutions before hospitalisation (n=45) and patients for which the place of living before hospitalisation was not known (n=38) were left out of the analysis. The reason for this is that for institutionalised patients their postcode probably is the postcode of the institution and not of their home address. Of the remaining 677 patients, 23 patients

had to be excluded because no valid postcode was obtained, and three patients because no information on socioeconomic status of their postcode area was available.

The information about the postcode areas of the patients used in this study was obtained from an existing database.† The information in this database has been collected since the mid-1980s and is continuously updated by telephone interviews with people living in the areas.²⁴ Each year about 80 000 to 100 000 interviews are held in which information about the areas is gathered using a projection method: the respondents are not asked for information about their own household, but for information about the characteristics of their postcode area (of which the interviewer tells beforehand which (parts of) streets and house numbers belong to it). To measure socioeconomic status of the areas, the respondents are asked to classify the average income level of the other inhabitants of their postcode area. For 10% to 12% of the areas, no income information could be obtained in this way. For most of these areas, extrapolation was used on the basis of information of the less detailed postcode sector to which the postcode area belonged. If the type of housing and the tenure class were comparable, the most frequent income class of the larger area was substituted. For 7940 (3.7%) of the postcode areas in our study region no income information was available. These areas, therefore, were excluded from our study. We coded the income level variable into three categories: (1) below average, (2) average or above, and (3) diverse. The category “diverse” is used for postcode areas which are heterogeneous with regard to income.†

Given the higher incidence of stroke among older persons, we have to control for age in our analysis. For this purpose, the variable “family phase” of the postcode area was used, because no individual information was available for the total “population at risk” (only for the stroke patients). Family phase was

†The postcode information used in this study was obtained from Geo-Marktprofiel BV in Nieuwegein, an organisation that gathers this information for commercial purposes.

coded into four categories: (1) young (individuals and couples who just started a household of their own), (2) middle (households with children), (3) old (older individuals and couples, without children at home), and (4) diverse. Also for this variable missing information was filled in by extrapolation from the postcode sector to which the area belongs.

The variable “urbanisation” controls for (socioeconomic) health differences between more and less urbanised areas. On the basis of the number of persons per postcode area, for each community the number of inhabitants was computed. This variable was coded into five categories: (1) less than 5000 inhabitants, (2) 5000–20 000 inhabitants, (3) 20 000 to 100 000 inhabitants, (4) 100 000 to 250 000 inhabitants, and (5) more than 250 000 inhabitants. To control for the fact that the hospitals might not be distributed equally over the six provinces, we also included dummies for the province in which the patient lived before hospitalisation in the model.

Statistical analysis

To determine the relation between socioeconomic status and stroke incidence, the characteristics of the postcode areas in which the stroke patients were living before hospitalisation were compared with the postcode area characteristics of all people living in the six provinces to which the data apply. For this comparison, we used information on the frequency distribution of the postcode characteristics in the six provinces. This frequency distribution was available for all possible combinations of the variables, which counts down to $3 \times 4 \times 5 = 60$ combinations in each of the six provinces. For each of these combinations in each of the provinces, we knew the total number of persons and also the number of stroke patients with that combination. On the basis of this information, a Poisson regression analysis was performed. The results of this analysis show to what extent the stroke patients differed with regard to socioeconomic status, family phase, and urbanisation of their postcode areas from the general pattern in the study region. The effect sizes are expressed as relative risks (RRs) taking the higher socioeconomic group as the reference category.

RESULTS

Table 1 presents the results of the Poisson regression analysis. The coefficients show that there is a significant effect of the area's socioeconomic status on stroke incidence. The effect is also in the expected direction. Among people living in a postcode area with below average socioeconomic status the incidence of stroke is significantly higher (RR=1.27) than among people living in a postcode area with average or above average socioeconomic status. The postcode areas that are heterogeneous with regard to socioeconomic status do not deviate significantly from the higher socioeconomic category.

The effects of the variable “family phase”, which controls for age differences in stroke incidence, are in the expected direction. Among people living in postcode areas with predominantly older inhabitants, stroke incidence is significantly higher (RR=3.17) than among people living in postcode areas with predominantly young inhabitants. Stroke incidence is also significantly increased in the areas that are heterogeneous in this respect (RR=2.19).

With regard to urbanisation, table 1 shows that stroke incidence is significantly lower in the countryside and that it is highest in large cities. Also among the six Dutch provinces significant differences in stroke incidence are found. Because these variables are mainly present to control for the—possible—unequal distribution of the hospitals over the provinces and over areas of different level of urbanisation, we will not draw substantial conclusions on these effects.

DISCUSSION

This study used very detailed geographical information to assess the effects of socioeconomic status on stroke incidence

Key points

- Stroke incidence tends to increase when community income levels decrease.
- In detailed postcode areas with lower income levels, stroke incidence is a quarter higher.
- Income level of the nearby living environment seems an effective substitute for individual income.

in the Netherlands. Our analysis shows a clear socioeconomic gradient with regard to stroke incidence in the early 1990s in the Netherlands: people living in detailed postcode areas with below average socioeconomic status experience a significantly higher risk of stroke. This result suggests that if socioeconomic status information at the individual level is lacking, detailed postcode area information can be used as an efficient proxy.

A few remarks should be made in interpreting the findings of our study. Firstly, there is the possibility of selection bias. Although we control for the unequal distribution of the hospitals over the provinces, the possibility remains that within the provinces not every person has the same probability of being admitted to the selected hospitals. This would be a problem if the probability of being admitted to these hospitals would be related to socioeconomic status. Such selective admittance might take place for several reasons. Firstly, it might be the result of individual choices of patients with different socioeconomic status for different hospitals. However, in the Dutch situation, this possibility does not seem very likely. In case of emergency, like acute stroke, people are generally brought to the nearest hospital. And, more generally, Dutch people tend to visit the nearest hospital for treatment.²⁵ Furthermore, because the Dutch population is almost completely covered by health insurances that cover the costs of admittance for stroke to every Dutch hospital,¹ choosing certain hospitals for financial reasons does not play much of a part in the Netherlands. Secondly, there is the possibility of socioeconomic differences in the risk of dying at home before being brought to a hospital. For this possibility, we could not control in our study. However, given the lack of barriers to calling an ambulance, these differences are probably small in the Netherlands. Moreover, because the risk of dying from stroke is higher among the lower socioeconomic groups,^{6–9} this selection effect—if it exists—most probably would lead to an underestimation of the real effect of socioeconomic status on stroke incidence. Thirdly, there may be socioeconomic differences in the regions served by the hospitals. This form of selection bias is probably ruled out to a large extent by controlling for urbanisation in the analysis. Both within large cities and within smaller cities or the countryside, Dutch hospitals tend to serve all kinds of areas. Neighbourhoods and boroughs with homogenous socioeconomic status are relatively small in the Netherlands compared with the size of the areas served by the hospitals. Still, we cannot completely rule out the possibility that after controlling for province and urbanisation some differences in socioeconomic characteristics of the serving areas remain that may bias our results in either way.

Besides the possibility that our results are biased by selection bias, there is the risk of making the ecological fallacy: because socioeconomic status of areas is used instead of individual socioeconomic status, we cannot be completely sure that the socioeconomic status effects found in this study are similar to those that would be found if individual socioeconomic status was used. Because of the small size of the postcode areas that were used, the difference will be much smaller than in previous area based studies. However, it is not reduced to zero. Even within the very small areas we studied, some variation in socioeconomic status may exist. Therefore it is not possible to make an exact estimation of the effect of the socioeconomic status of

people on the basis of this area information. However, for the study of socioeconomic status effects on health, such a lack of precision need not always be a problem. Firstly, the focus of much of this research is primarily on associations among variables and not on characteristics of people. As long as the variation in socioeconomic status within postcode areas is small compared with the variation in the total region to which the study applies, the lack of precision does not need to cause fundamental problems. Secondly, researchers of health inequalities are often interested in the total impact of socioeconomic status on health, including both individual effects and neighbourhood effects. In that case, a lack of precision at the individual level may to a certain extent be compensated by a higher precision at the neighbourhood level. Thirdly, to determine the "pure" effect of individual socioeconomic status on health, even the use of a very precise measure of individual socioeconomic status would not suffice. The coefficient of any indicator of individual socioeconomic status would to a certain extent reflect the effect of area status, as a result of the positive correlation between the socioeconomic statuses of people and of the areas in which they live. Finally, when despite the restricted measurement of socioeconomic status an effect on stroke incidence is found, this effect very probably is a real effect. When there are no systematic distortions, a poor measurement of a variable generally leads to dilution and underestimation of the true association. This means that the risk of wrongfully concluding that an effect exists is low.

There is still another reason why the effects found in this study may be underestimations of the true effects of individual socioeconomic status: the postcode areas where the patients from our study group lived are not the only postcode areas with stroke patients in the region under study. In a substantial number of the postcode areas that were used as the reference group, also stroke patients live. However, given the very large number of 206 855 postcode areas used in this study, this underestimation can be expected to be very low.

ACKNOWLEDGEMENTS

The research reported is part of the Research on Stroke Amsterdam (ROSA), a multicentre study on quality of care after stroke conducted by the Academic Medical Center / University of Amsterdam. The departments involved are Social Medicine (Prof dr G A M van den Bos), Neurology (Prof dr M Limburg), and Clinical Epidemiology and Biostatistics (Prof dr R J de Haan). ROSA was supported by grants from the Dutch Organisation for Scientific Research, the Netherlands Heart Foundation and the Dutch Fund for Investigative Medicine of the Health Insurance Council.

.....

Authors' affiliations

J Smits, G P Westert, G A M van den Bos, National Institute of Public Health and the Environment, Department for Health Services Research, Netherlands

Funding: none.

Conflicts of interest: none.

REFERENCES

- 1 **Ruwaard D**, Kramers PGN, eds. *Public health status and forecast 1997; Health, prevention and health care in the Netherlands until 2015*. Maarsse: Elsevier/De Tijdstroom, 1997.
- 2 **Townsend P**, Davidson N, Whitehead M. *Inequalities in health (The Black report and the health divide)*. London: Penguin Books, 1988.
- 3 **Kaplan GA**, Keil JE. Socioeconomic factors and cardiovascular disease: a review of the literature. *Circulation* 1993;**88**:1973-98.
- 4 **Gordon D**, Shaw M, Dorling D, et al, eds. *Inequalities in health*. Bristol: The Policy Press, 1999.
- 5 **Keskimäki I**. *Social equity in the use of hospital inpatient care in Finland*. [Thesis]. STAKES, 1997.
- 6 **Siegel PZ**, Deeb LC, Wolfe LE, et al. Stroke mortality and its socioeconomic, racial and behavioral correlates in Florida. *Public Health Rep* 1993;**108**:454-8.
- 7 **Bennett S**. Socioeconomic inequalities in coronary heart disease and stroke mortality among Australian men, 1979-1993. *Int J Epidemiol* 1996;**25**:266-75.
- 8 **Kunst AE**, del Rios M, Groenhouf F, et al. Socioeconomic inequalities in stroke mortality among middle-aged men: an international overview. European Union Working Group on Socioeconomic Inequalities in Health. *Stroke* 1998;**29**:2285-91.
- 9 **Maheswaran R**, Elliot P, Strachan DP. Socioeconomic deprivation, ethnicity and stroke mortality in Greater London and south east England. *J Epidemiol Community Health* 1997;**51**:127-31.
- 10 **Reeder BA**, Liu L, Horlick L. Sociodemographic variation in the prevalence of cardiovascular disease. *Can J Cardiol* 1996;**12**:271-7.
- 11 **Van Rossum CT**, van de Mheen H, Breteler MMB, et al. Socioeconomic differences in stroke among Dutch elderly women: the Rotterdam study. *Stroke* 1999;**30**:357-62.
- 12 **Carstairs V**, Morris R. Deprivation and health in Scotland. *Health Bull* 1990;**48**:162-75.
- 13 **Townsend P**, Phillimore P, Beattie A. *Health and deprivation: inequality and the North*. London: Croom Helm, 1988.
- 14 **Krieger N**. Overcoming the absence of socioeconomic data in medical records: validation and application of a census-based methodology. *Am J Public Health* 1992;**82**:703-10.
- 15 **Reijneveld SA**, Verheij RA, de Bakker DH. The impact of area deprivation on differences in health: does the choice of the geographical classification matter? *J Epidemiol Community Health* 2000;**54**:306-13.
- 16 **Huff N**, Macleod C, Ebdon D, et al. Inequalities in mortality and illness in Trent NHS region. *J Public Health Med* 1999;**21**:81-7.
- 17 **Starr JM**, Thomas B, Whalley LJ. Population risk factors for hospitalisation for stroke in Scotland. *Int J Epidemiol* 1996;**25**:276-81.
- 18 **Piantadosi S**, Byar DP, Green SB. The ecological fallacy. *Am J Epidemiol* 1988;**127**:893-904.
- 19 **Hyndman JCG**, d'Arcy C, Holman J, et al. Misclassification of social disadvantage based on geographical areas: comparison of postcode and collector's district analyses. *Int J Epidemiol* 1995;**24**:165-76.
- 20 **Jarman B**. Scores should be based on enumeration districts and payments should be phased in gradually. *BMJ* 1997;**314**:228-9.
- 21 **De Haan R**, Limburg M, Van der Meulen J, et al. Quality of life after stroke: impact of stroke type and lesion location. *Stroke* 1995;**26**:402-8.
- 22 **Van Straten A**, van der Meulen JHP, van den Bos GAM, et al. Length of hospital stay and discharge delays in stroke patients. *Stroke* 1997;**28**:137-40.
- 23 **Scholte op Reimer WJM**, de Haan RJ, Rijnders PT, et al. Unmet care demands after stroke: deficits in health care? *Qual Health Care* 1999;**8**:30-5.
- 24 **Van der Kooij J**. Geo-demografische informatie op laag schaalniveau: De postcode als ruimtelijke ingang. *Planning, Methodiek en Toepassing* 1992;**41**:26-33.
- 25 **Westert GP**. *Variations in use of hospital care*. Assen: Van Gorcum, 1992.