

THEORY AND METHODS

What does self rated health measure? Results from the British Whitehall II and French Gazel cohort studies

Archana Singh-Manoux, Pekka Martikainen, Jane Ferrie, Marie Zins, Michael Marmot, Marcel Goldberg

J Epidemiol Community Health 2006;60:364–372. doi: 10.1136/jech.2005.039883

See end of article for authors' affiliations

Correspondence to:
Dr A Singh-Manoux,
INSERM, U687, HNSM,
14 rue du Val d'Osne,
94415 Saint-Maurice
Cédex, France; Archana.
Singh-Manoux@
st-maurice.inserm.fr

Accepted for publication
18 November 2005

Objectives: To investigate the determinants of self rated health (SRH) in men and women in the British Whitehall II study and the French Gazel cohort study.

Methods: The cross sectional analyses reported in this paper use data from wave 1 of the Whitehall II study (1985–88) and wave 2 of the Gazel study (1990). Determinants were either self reported or obtained through medical screening and employer's records. The Whitehall II study is based on 20 civil service departments located in London. The Gazel study is based on employees of France's national gas and electricity company (EDF-GDF). SRH data were available on 6889 men and 3403 women in Whitehall II and 13 008 men and 4688 women in Gazel.

Results: Correlation analysis was used to identify determinants of SRH from 35 measures in Whitehall II and 33 in Gazel. Stepwise multiple regressions identified five determinants (symptom score, sickness absence, longstanding illness, minor psychiatric morbidity, number of recurring health problems) in Whitehall II, explaining 34.7% of the variance in SRH. In Gazel, four measures (physical tiredness, number of health problems in the past year, physical mobility, number of prescription drugs used) explained 41.4% of the variance in SRH.

Conclusion: Measures of mental and physical health status contribute most to the SRH construct. The part played by age, early life factors, family history, sociodemographic variables, psychosocial factors, and health behaviours in these two occupational cohorts is modest.

The single item measure of self rated health (SRH) has been widely reported to be a predictor of mortality after adjusting for traditional risk factors, sociodemographics, and measures of health status.^{1–4} Reviews on a total of 46 studies in various settings have found this relation to be robust.^{5,6} Attempts to unravel this association in the past have focused on controlling for a wide variety of risk factors: family history,⁷ measures of health,^{2,8–12} chronic diseases,^{4,7,8,12–19} functional ability,^{7,10,12,14,17,19–21} sociodemographic variables,^{2–4,7,9,10,12–16,21,22} psychosocial factors,^{2,10,15,21–24} and health behaviours.^{2–4,7–9,15,16} This line of inquiry suggests that SRH is a multidimensional and a holistic measure,^{5,25} even though its correlates have rarely been examined in empirical analyses.

Understanding the relation between SRH and health or mortality requires further investigation into the criteria by which people judge their health. The aim of this paper is to identify the determinants of SRH in two different cohorts, in separate analyses for men and women. The comparative framework is aimed at assessing whether the correlates of SRH are similar in different cultural contexts. We use a range of variables, traditionally used as controls in the examination of the SRH-mortality relation, to examine the extent to which they predict SRH. Six categories of variables are used: age, early life factors and measures of family history, sociodemographic variables, psychosocial factors, health behaviours, and measures of health and disease.

METHODS

Study populations

Data for the analyses come from two European cohort studies: the British Whitehall II study and the French Gazel cohort study.

The Whitehall II study

The Whitehall II study was established in 1985 as a longitudinal study to examine the socioeconomic gradient in health and disease among 10 308 civil servants (6895 men and 3413 women).²⁶ All civil servants aged 35–55 years in 20 London based departments were invited to participate by letter, and 73% agreed. Baseline examination (phase 1) took place during 1985–1988, and entailed a clinical examination and a self administered questionnaire containing sections on demographic characteristics, health, lifestyle factors, work characteristics, social support, and life events. Clinical examination included measures of blood pressure, anthropometry, biochemical measurements, neuroendocrine function, and subclinical markers of cardiovascular disease. Subsequent phases of data collection have alternated between postal questionnaire alone and postal questionnaire accompanied by a clinical examination. Since baseline six phases of data collection have been completed, with the most recent phase (phase 7) completed in 2004. The University College London ethics committee approved the study.

The GAZEL cohort study

The GAZEL cohort was established in 1989, on employees of France's national gas and electricity company: Electricité de France-Gaz de France (EDF-GDF). Further details of this study can be found elsewhere.²⁷ At baseline, 20 625 (15 011 men and 5614 women), aged 35–50, gave consent to participate in this study. The study design consists of an annual questionnaire used to collect data on health, lifestyle, individual, familial, social and occupational factors, and life events. Various sources within EDF-GDF provide additional data about Gazel participants. Occupational and personal data are updated through human resources department files. Medical data on sick leaves and incidence of cancer and

Table 1A Variables used to assess self rated health in the Whitehall II data

Variable	Description
Early life factors and family history	
Height*	Height measured at clinical screening at phase 1.
Father's social class	Assessed using registrar general's social class classification: six hierarchical categories: 1 (high) to 6 (low).
Vital status of parents	A three point scale assessing whether both parents alive (1), 1 alive or both dead (3).
Chronic diseases in parents	A six point scale assessing whether either parent suffered from angina, heart attack, stroke, high blood pressure, and diabetes; 0 = no disease and 5 = all five diseases.
Chronic diseases in siblings	A six point scale assessing any sibling suffered from angina, heart attack, stroke, high blood pressure, and diabetes; 0 = no disease and 5 = all five diseases.
Sociodemographics	
Material difficulties	Score from 0–8, composed of two five point scales assessing the extent to which person could afford food and clothing for the family and had difficulty in paying bills; 0 = not difficult and 8 = very difficult.
Education	A three point scale assessing education: 1 = university, 2 = secondary, 3 = primary level of education.
Employment grade	A six point scale representing occupational position in the British civil service; 1 (high) to 6 (low).
Marital status	Four point scale: 1 = married/cohabiting, 2 = single, 3 = divorced/separated and 4 = widowed.
Psychosocial factors	
Health locus of control	View on whether there were things one could do to reduce chance of heart attack: 1 = definitely, 2 = might, 3 = little one can do.
Affect balance scale	The 10 items of the Bradburn affect balance scale were presented as four point scales (0–3), summed up to a scale ranging from –1.5 to +1.5; high scores reflect high positive affect. ³²
Hostility score	The Cook-Medley hostility scale, high scores reflect high hostility. ⁴¹
Satisfaction with life scale	Five item scale, scores range from 0–30; high scores indicate greater life satisfaction. ⁴²
Social isolation	High scores on the scale (0–6) reflect greater social isolation; assessing lack of people to share with, no relatives, friends or work colleagues seen, and no social or religious belonging.
Network scale	Network scale (0–28) assesses the frequency and the number of friends, relatives and work colleagues seen; and participation in social and religious groups; high scores indicate bigger network.
Life events	Number of life events (0–8) in the past year: serious illness/injury, death of a close relative/friend, serious illness/injury of a close person, financial difficulty, divorce/separation, other marital/family problem, mugging/robbery, change of job/residence.
Upset with life events	Sum of four point scales (0–3) assessing the impact of the each life event (total 0–24); high scores reflect greater emotional impact.
Job demand	Karasek measure of psychological job demands; high scores reflect high demand. ⁴³
Decision latitude	Karasek measure: 15 items dealing with decision authority and skill discretion, combined into an index of decision latitude (or job control); high scores reflect greater control. ³⁹
Work satisfaction	A one item 4-point measure from the Karasek questionnaire asking respondents to judge how satisfied they were with their job overall; high scores reflect greater dissatisfaction. ³⁹
Health behaviour	
Smoking	Five point composite index representing non/former/light/medium/heavy smoker; 1 = non-smoker and 5 = heavy smoker.
Alcohol consumption	Composite index assessing the units of alcohol consumed in the past week.
Level of physical activity	Composed of a three point scale assessing low/medium or high level of physical activity; 1 = high and 3 = low level.
Frequency of eating fresh fruit and vegetables	Frequency of eating fresh fruit and vegetables on an eight point scale; 1 = consumed 2 or more times a day and 8 = seldom or never consumed.
Health	
Sickness absence	Self reported number of sick days in the past year.
Longstanding illness	Response to a question asking the person if they had a longstanding illness: 0 = no and 1 = yes.
Number of longstanding illnesses	The sum of the number of longstanding illnesses indicated by participants; 0 = no illness.
Number of recurring health problems	Number of health problems out of a list of 13: bronchitis, arthritis, sciatica, persistent skin trouble, asthma, hay fever, persistent stomach trouble, constipation, piles, persistent foot trouble, trouble with varicose veins, depression, persistent trouble with gums or mouth. 0 = no problems and 13 = all of them.
Symptom score	Number of symptoms in the last 14 days: phlegm, diarrhoea, indigestion, shortness of breath, dizziness, earache, swollen ankles, nervous/depressed, cold/flu, sore throat, sleep problems, pains in chest, backache, nausea/vomiting, tired, skin trouble, toothache. 0 = no problems and 17 = all of them.
Number of medicines used in the past 14 days	List, indicating the number of all prescribed drugs used by the participant in the past 14 days; 0 = no medicines.
Sleep	Five categories of number of hours slept: 1 = <9; 2 = <8; 3 = <7; 4 = <6; 5 = <5.
Body mass index (BMI)*	Weight in kilograms divided by the square of the height in metres.
Physical illness indicator	Composite indicator of three doctor diagnosed illnesses (diabetes, heart trouble, hypertension) and respiratory illness measured by MRC chronic bronchitis questionnaire; 0 = no illness and 4 = all 4 illnesses.
Minor psychiatric morbidity	Mental health was assessed using the general health questionnaire; 0 = good mental health and 90 = poor mental health. ³⁴

*Not self reported.

coronary heart diseases come from the company's health insurance department. Occupational physicians collect data on working conditions and occupational exposures. Sources outside EDF-GDF provide data on causes of death, health care use, and admissions to hospital. Every five years, volunteers are invited to visit a health centre run by the national health insurance fund, where they undergo a full standardised medical examination.

Measures

All the variables used in the analyses come from wave 1 (1985–89) of the Whitehall II study and wave 2 of the Gazel study (1990) as the wave 1 data did not have several critical variables. Earliest waves of the study have been used to reduce bias related to loss of follow up.

Self rated health

The Whitehall II measure of SRH asked respondents: "Over the past 12 months would you say your health has been—very good, good, average, poor or very poor". The Gazel participants responded on an eight point scale, anchored by 1 = "very good" and 8 = "very bad", to the following question: "How would you judge the state of your general health?"

Explanatory variables

Tables 1A and 1B show the explanatory variables, definitions, and analysis categories for both studies. These variables have been grouped into six theoretical categories: early life factors and measures of family history, sociodemographic variables, psychosocial factors, health behaviour, and measures of

Table 1B Variables used to assess self rated health in Gazel data

Variable	Description
Early life factors and family history	
Height	Self reported height
Father's social class	INSEE categories to assess socioeconomic position of father; 1 (high) to 8 (low). (http://www.insee.fr/fr/nom_def_met/nomenclatures/prof_cat_soc/html/LO3_N1.HTM);
Myocardial infarction in parents	A three point (0–2) scale assessing whether both parents had myocardial infarction (2), only one (1) or neither (0).
Cancer in parents	A three point (0–2) scale assessing whether both parents were diagnosed with a cancer (2), only one (1) or neither (0).
Chronic disease in siblings	A three point scale (0–2) assessing whether both brother and sister were diagnosed with a cancer (2), only one (1) or neither (0).
Sociodemographics	
Income	Seven level hierarchic measure assessing the net available monthly household income; 1 => >25000FF and 7 = <7500FF.
Education	Seven level hierarchic measure of educational qualification; 1 = university degree and 7 = primary school level.
Employment grade	INSEE categories to assess own socioeconomic position; 1 (high) to 8 (low). (http://www.insee.fr/fr/nom_def_met/nomenclatures/prof_cat_soc/html/LO3_N1.HTM).
Marital status	Four point scale: 1 = married/cohabiting, 2 = single, 3 = divorced/separated and 4 = widowed.
Psychosocial factors	
Emotional reactions (NHP)	Emotional reactions were measured using the Nottingham health profile (NHP) emotional reactions scale, composed of nine items; higher scores represent greater emotional reactions. ³²
Social isolation (NHP)	Social isolation was measured using the NHP social isolation scale, composed of five items; higher scores represent greater social isolation. ³²
Network scale	A 1.5 point scale composed of three items: a measure of the number of friends and relatives seen regularly and the number seen at least once a month; high scores indicate bigger network.
Life events	Number of life events in the past year: marriage, divorce/separation, birth/adoption, miscarriage, death of a close relative/friend, hospitalisation, house move, retirement, unemployment of partner/relative/close friend, important change at work, loss or theft of goods, important purchase, change in financial condition.
Support during life events	An eight point scale assessing whether the respondent received support in relation to the three major life events; scale from 0 to 21; high scores indicate low support.
Satisfaction with work	An eight point scale assessing whether the respondent was "satisfied with their work overall"; 1 = very satisfied and 8 = not at all.
Work physically tiring	An eight point scale assessing whether the respondent found their work "physically tiring"; 1 = not at all and 8 = very tiring.
Work mentally tiring	An eight point scale assessing whether the respondent found their work "mentally tiring"; 1 = not at all and 8 = very tiring.
Health behaviour	
Smoking	Five point composed index representing non/former/light/medium/heavy smoker; 1 = non-smoker and 5 = heavy smoker.
Alcohol consumption	Composed index assessing the units of alcohol consumed in the past week.
Level of physical activity	Three point scale assessing low/medium or high level of physical activity; 1 = high and 3 = low level.
Frequency of eating fresh fruit	Frequency of eating fresh fruit on a four point scale; 1 = consumed most days and 4 = seldom or never consumed.
Frequency of eating fresh vegetables	Frequency of eating fresh vegetables on a four point scale; 1 = consumed most days and 4 = seldom or never consumed.
Health	
Sickness absence*	Physician certified number of sick days in the past year recorded for the person by the employer.
Off sick for over a month	Response to a question asking the individual whether they were off sick from work for over a month at a time; 0 = no and 1 = yes.
Number of health problems in the past year	Number of health problems in the last year out of a list of 53 possible ailments, categorised in nine subgroups: respiratory diseases cardiovascular diseases, digestive ailments, musculoskeletal diseases, urinary and genital problems, endocrine and metabolic ailments, trouble with sensory organs, psychiatric troubles, and other diseases.
Number of new health problems in the past year	Respondents were asked to indicate the number of new health problems in the past year from a list of 53 possible ailments described above.
Physical mobility	Physical mobility was measured using the NHP, physical mobility scale, composed of eight items; higher scores indicate greater trouble with physical mobility. ³²
Number of prescription drugs	List, indicating the number of all prescribed drugs currently being used by the participant.
Sleep (NHP)	Sleep was measured using the NHP sleep scale, composed of five items; high scores indicate poor sleep quality. ³²
BMI	Weight in kilograms divided by the square of the height in metres.
Physical illness indicator	Composite indicator of the three diagnosed illnesses (diabetes, heart trouble, hypertension) and respiratory illness measured by MRC chronic bronchitis questionnaire; 0 = no illness and 4 = all 4 illnesses as before.
Physical tiredness	An eight point scale indicating the level of current physical tiredness; 0 = not at all tired and 7 = very tired.

*Not self reported.

health and disease; further details of these measures can be found in table 1A for the Whitehall II and table 1B for Gazel.

Statistical analysis

Analysis was carried out separately, in men and women and in the two cohorts. The first step consisted of examining correlations to decide the variables that would be used to predict SRH. All variables that were significantly correlated with SRH at $p < 0.05$ were then used as determinants of SRH. We used multiple linear regressions to build models to predict SRH.²⁸ Linearity of association between the determinants and SRH was ensured using the one way analysis of variance test of linearity; further examination of residuals scatterplots found no important deviations from the assumptions of normality and linearity. In the first set of analyses, the variables were entered in a hierarchical multiple regression; the order of entry of variables (or a set of variables) here was theoretically determined. We chose the order in an attempt to reflect temporality over the life course; early life measures were entered before later life measures, health behaviours

before health itself, and so on. The order of entry of variables was as follows: age, early life factors and measures of family history, sociodemographic variables, psychosocial factors, health behaviours, and measures of health and disease. *F* tests were used to compute the significance of each set of variables; the resulting change in r^2 reflected the variance explained in the outcome. The change in r^2 was tested at $p < 0.05$ to check if it was significantly different from zero.

The next step involved stepwise method of multiple regression to identify a small subset of determinants of SRH. This analysis is appropriate here as the cases (number) to determinants ratio is much larger than the recommended 50:1 ratio.²⁹ The basic procedures involved are: identifying an initial model with one determinant, one that is most highly correlated with SRH; iteratively "stepping", that is repeatedly changing the model at the previous step by adding or removing determinants in accordance with the "stepping criteria" defined; and terminating the search when stepping is no longer possible given the stepping criteria. Given the large number of cases in the analysis the stepping criteria was

Table 2A Whitehall II: correlations between self rated health† and determinants

	Men		Women	
	r	Number	r	Number
Age	NS (6869)	6869	0.06***	3403
Early life factors and family history				
Height‡§	-0.04**	6860	-0.07***	3402
Father's social class	NS	4764	0.06**	2204
Vital status of parents	0.02*	6827	0.05**	3379
Chronic diseases in parents	0.04***	6812	0.05**	3358
Chronic diseases in siblings	0.06***	5539	0.08***	2857
Sociodemographics				
Material difficulties	0.12***	5158	0.16***	2491
Education	0.04**	5168	0.06**	2499
Employment grade	0.13***	6869	0.13***	3403
Marital status	0.10***	6850	NS	3385
Psychosocial factors				
Health locus of control	0.08***	5123	0.07***	2482
Affect balance scales	-0.30***	5034	-0.30***	2348
Hostility score	0.17***	4255	0.17***	1769
Satisfaction with life scales	-0.27***	5117	-0.23***	2429
Isolation score	0.10***	6750	0.04*	3295
Network scales	-0.08***	6802	-0.04*	3346
Life events	0.16***	6731	0.13***	3237
Upset with life events	0.20***	6697	0.18***	3201
Job demand	NS	6855	0.05**	3361
Decision latitude§	-0.18***	6836	-0.14***	3330
Work satisfaction	0.19***	6832	0.11***	3359
Health behaviour				
Smoking	0.12***	6682	0.08***	3352
Alcohol consumption	NS	6816	-0.08***	3364
Level of physical activity	0.13***	6869	0.07***	3403
Frequency of eating fresh fruit and vegetables	0.13***	6852	0.08***	3389
Health				
Sickness absence	0.35***	6736	0.36***	3273
Longstanding illnesses	0.31***	5156	0.30***	2487
Number of long standing illnesses	0.28***	5093	0.28***	2456
Number of recurring health problems	0.33***	5134	0.38**	2459
Symptom score	0.40***	5137	0.44***	2437
Number of medicines used in the past 14 days	-0.07***	6869	NS	3403
Sleep	0.05***	6839	0.11***	3391
BMI‡	0.06***	6857	0.12***	3401
Physical illness indicator‡	0.19***	6869	0.22**	3403
Minor psychiatric morbidity: GHQ	0.32***	6665	0.33***	3240

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$. †SRH measured on a five point scale; 1 = very good and 5 = very poor. ‡Not self reported. §Negative correlations explained by the fact that high scores on these measures denote positive outcomes for the person.

set at $p = 0.001$ for entry and $p = 0.01$ at exit.³⁰ Furthermore, only variables that explained more than 1% of the variance ($r^2 > 0.01$) in SRH were retained as determinants. Stepwise regression provides an objective method for selecting variables that maximise r^2 with the smallest number of variables used. As a large number of determinants were used, multicollinearity statistics were analysed throughout using tolerance statistics, set here at 0.50. All analyses were carried out using SPSS 11.5.1 (SPSS for Windows, rel 11.0.1, Chicago, SPSS, 2001).

RESULTS

SRH data were available on 6889 men and 3403 women (10 292 overall; 99.85% of total number at baseline) in the Whitehall II study and 13 008 men and 4688 women in the Gazel study (17 696 overall; 88.85% of total number at baseline). Because of missing values, our analyses, at best, include 7043 (68.43% of those with SRH data) people in Whitehall II and 17 438 (98.54% of those with SRH data) in the Gazel study. Lower numbers in the Whitehall II data are related to four different versions of a questionnaire being used at phase 1 of the study; the first 2913 participants answered versions 1–3 of the questionnaire that did not have several key variables used here (for example, longstanding illness). Thus, many of the bivariate analyses are based on smaller numbers. The key variable associated with missing data in both cohorts

is employment grade, with data more likely to be missing among people occupying low employment grade.

Tables 2A and 2B present bivariate analyses, correlation coefficients here, to show the association between SRH and its correlates in the Whitehall II and the Gazel data, respectively. The correlations were deemed to be significant if $p < 0.05$. Table 2A (Whitehall II results) shows that in men four (age, father's social class, job demand, and alcohol consumption) of the 35 determinants considered and two among women (father's social class and number of medicines used in the past 14 days) did not have a significant association with SRH. These variables were dropped in subsequent analyses. Symptom score, sickness absence, and the number of recurring health problems were the three variables most strongly associated with SRH in the Whitehall II data (table 2A).

Table 2B presents the correlations between 33 determinants and SRH in the Gazel data. Three of these variables in men (height, father's social class, support received during life events) and nine in women (height, father's social class, myocardial infarction in parents, cancer in parents, cancer in siblings, smoking, alcohol consumption, frequency of eating fresh fruit, and frequency of eating fresh vegetables) were not associated with SRH and are not considered in further analyses. The measures most strongly related to SRH in the Gazel data in bivariate analysis were: a measure of physical

Table 2B Gazel: correlations between self rated health† and determinants

	Men		Women	
	r	Number	r	Number
Age	0.07***	13008	0.11	4688
Early life factors and family history				
Height‡	NS	12773	NS	4535
Father's social class	NS	12757	NS	4538
Myocardial infarction in parents	0.03***	10409	NS	3549
Cancer in parents	0.04***	13008	NS	4688
Cancer in siblings	0.03***	13008	NS	4688
Sociodemographics				
Income	0.10***	12617	0.08***	4430
Education	0.06***	12788	0.05***	4572
Employment grade‡	0.04***	12940	0.05**	4656
Marital status	0.04***	12926	0.05***	4651
Psychosocial factors				
Emotional reactions (NHP)	0.35***	12856	0.43***	4637
Social isolation (NHP)	0.22***	12856	0.27***	4637
Network scale§	-0.09***	11645	-0.10***	4034
Life events	0.03**	13008	0.07***	4688
Support during life events	NS	13008	0.03*	4688
Satisfaction with work	0.23***	11720	0.19***	4003
Work physically tiring	0.21***	12556	0.28***	4440
Work mentally tiring	0.20***	12854	0.28***	4588
Health behaviour				
Smoking	0.07***	12668	NS	4499
Alcohol consumption	0.02*	13008	NS	4688
Level of physical activity	0.20***	11870	0.17***	4127
Frequency of eating fresh fruit	0.04***	12831	NS	4621
Frequency of eating fresh vegetables	0.06***	12927	NS	4653
Health				
Sickness absence‡	0.13***	13008	0.16***	4668
Whether off sick for over a month	0.18***	12969	0.22***	4649
Number of health problems in the past year	0.43***	13008	0.48***	4688
Number of new health problems in the past year	0.17***	13008	0.20***	4688
Physical mobility (NHP)	0.34***	12856	0.35***	4637
Number of prescription medication	0.30***	13008	0.32***	4688
Sleep	0.28***	12856	0.33***	4637
BMI	0.09***	12756	0.08***	4522
Physical illness indicator	0.24***	13008	0.18***	4688
Physical tiredness	0.58***	12977	0.59***	4665

*p<0.05; **p<0.01; ***p<0.001. †SRH measured on an eight point scale; 1 = very good and 8 = very poor.
‡Not self reported. §Negative correlations explained by the fact that high scores on these measures denote positive outcomes for the person.

tiredness, number of health problems in the past year and emotional reactions, as measured by the Nottingham health profile (NHP).³² The principal differences between the bivariate associations in Gazel and Whitehall II stem from the lack of association between SRH and measures of early life factors, family history, and health behaviours among the Gazel women.

Table 3 presents the results of the hierarchical multiple regression, where determinants were entered in theoretically defined blocks. The top half of table 3 presents results for Whitehall II data; 31 correlates in men and 33 in women, grouped in five blocks for men (as age was not significantly correlated with SRH in men) and six blocks for women. In both men and women, all the blocks considered for analysis

Table 3 Predicting self rated health using categories of determinants. Whitehall II and Gazel

Model	r ² change	Change statistics (F)	r ² change	Change statistics (F)
Whitehall II				
		Men		Women
1		Age	0.006	6.67**
2		Early life factors and family history	0.017	3.72**
3	0.008	Sociodemographics	0.017	6.12***
4	0.030	Psychosocial factors	0.097	10.59***
5	0.105	Health behaviour	0.011	3.22**
6	0.021	Health	0.240	44.60***
	0.204	Overall model (Whitehall II)	r ² _{adjusted} = 0.369	F _{31, 1021} = 19.67***
	r ² _{adjusted} = 0.361	F _{31, 2820} = 52.97***		
Gazel				
		Men		Women
1		Age	0.004	32.33***
2		Early life factors and family history	0.001	3.10*
3	0.011	Sociodemographics	0.005	4.16**
4	0.171	Psychosocial factors	0.231	124.30***
5	0.020	Health behaviour	0.010	43.28***
6	0.221	Health	0.191	111.91***
	0.221	Overall model (Gazel)	r ² _{adjusted} = 0.443	F _{24, 3239} = 108.97***
	r ² _{adjusted} = 0.427	F _{30, 8250} = 206.34***		

*p<0.05; **p<0.01; ***p<0.001.

Table 4A Whitehall II: predicting SRH using stepwise regression†

Men		Overall model	$r^2_{\text{adjusted}}=0.330$	$F_{5, 4846}=478.86; p<0.0001$
			r^2 change	Change statistics (F)
1	Symptom score		0.154	884.72; $p<0.0001$
2	Sickness absence		0.096	621.88; $p<0.0001$
3	Longstanding illness		0.040	272.73; $p<0.0001$
4	Affect balance score		0.029	205.18; $p<0.0001$
5	Number of recurring health problems		0.012	83.53; $p<0.0001$
Women		Overall model	$r^2_{\text{adjusted}}=0.350$	$F_{5, 2190}=237.59; p<0.0001$
			r^2 change	Change statistics (F)
1	Symptom score		0.197	537.29; $p<0.0001$
2	Sickness absence		0.081	244.29; $p<0.0001$
3	Longstanding illness		0.038	120.38; $p<0.0001$
4	Minor psychiatric morbidity: GHQ		0.025	81.31; $p<0.0001$
5	Number of recurring health problems		0.012	41.55; $p<0.0001$
Men and women		Overall model	$r^2_{\text{adjusted}}=0.347$	$F_{5, 7038}=747.94; p<0.0001$
			r^2 change	Change statistics (F)
1	Symptom score		0.181	1554.21; $p<0.0001$
2	Sickness absence		0.091	883.45; $p<0.0001$
3	Longstanding illness		0.039	398.02; $p<0.0001$
4	Minor psychiatric morbidity: GHQ		0.024	250.39; $p<0.0001$
5	Number of recurring health problems		0.012	131.94; $p<0.0001$

†Stepping criteria: $p=0.001$ for entry, $p=0.01$ for exit and r^2 change associated with the variable to be >0.01 .

made a significant contribution to the prediction of SRH. The largest improvement in r^2 , in both men and women, was from model 5 to 6, reflecting the importance of health variables in predicting SRH. All the variables considered together explained 36.1 % of the variance in SRH in men and 36.9% in women.

The lower half of table 3 shows results for the Gazel data. All the blocks considered for analyses (none of the measures of early life factors and family history were considered for further analyses in women) predicted SRH significantly. The largest improvement in r^2 in men was from model 5 to 6 ($r^2 = 0.221$), followed by model 3 to 4 ($r^2 = 0.171$). Among women, the largest improvement was from model 3 to 4 ($r^2 = 0.231$), followed by model 5 to 6 ($r^2 = 0.191$). Thus measures of health and psychosocial factors were major determinants of SRH in the Gazel data. Overall, the variables considered explained 42.7% of the variance in SRH in men and 44.3% in women.

Table 4A presents the results from stepwise regression in the Whitehall II data. Five variables explained a total of 33% of the variance in SRH in men and 35% in women. Symptom

What this paper adds

- Data from two large European cohorts were used to examine the determinants of self rated health from six categories of variables: age, early life factors and measures of family history, sociodemographic variables, measures of health and disease, health behaviour, and psychosocial factors.
- Our results show that it is mostly measures of health, mental and physical, that are independent determinants of self rated health.

score was the strongest correlate of SRH in both men and women, accounting for 15.4% of the variance in SRH in men and 19.7% in women. Sickness absence was the second strongest correlate, accounting for 9.6% of the variance in SRH in men and 8.1% in women. The difference between men and women relates to the fourth correlate: affect balance

Table 4B Gazel: predicting SRH using stepwise regression†

Men		Overall model	$r^2_{\text{adjusted}}=0.415$	$F_{5, 11550}=1638.02; p<0.0001$
			r^2 change	Change statistics
1	Physical tiredness		0.321	5457.05; $p<0.0001$
2	Number of health problems in the year		0.056	1037.48; $p<0.0001$
3	Number of prescription drugs		0.015	286.14; $p<0.0001$
4	Physical mobility (NHP)		0.013	260.64; $p<0.0001$
5	Satisfaction with work		0.010	190.50; $p<0.0001$
Women		Overall model	$r^2_{\text{adjusted}}=0.439$	$F_{4, 4609}=904.13; p<0.0001$
			r^2 change	Change statistics
1	Physical tiredness		0.346	2439.88; $p<0.0001$
2	Number of health problems in the year		0.062	487.04; $p<0.0001$
3	Emotional reactions (NHP)		0.017	134.46; $p<0.0001$
4	Physical mobility (NHP)		0.014	118.74; $p<0.0001$
Men and women		Overall model	$r^2_{\text{adjusted}}=0.414$	$F_{4, 17434}=3083.51; p<0.0001$
			r^2 change	Change statistics (F)
1	Physical tiredness		0.336	8818.74; $p<0.0001$
2	Number of health problems in the year		0.051	1441.998; $p<0.0001$
3	Physical mobility (NHP)		0.017	492.95; $p<0.0001$
4	Number of prescription drugs		0.011	323.43; $p<0.0001$

†Stepping criteria: $p=0.001$ for entry, $p=0.01$ for exit and r^2 change associated with the variable to be >0.01 .

score in men and GHQ in women. Affect balance score³³ is a measure of psychological wellbeing whereas GHQ³⁴ measures minor psychiatric morbidity. These two measures are highly correlated in men ($r = -0.62$, $p < 0.0001$) and women ($r = -0.65$, $p < 0.0001$) in the Whitehall II study. In analyses of men and women together (table 4A), it is the GHQ that comes out a correlate, the other four remain the same. Longstanding illness and the number of recurring health problems were the other two predictors, explaining about 4% and 1.2 % of the variance respectively.

Table 4B presents the results from a similar stepwise regression in the Gazel data. In men, five variables met the stepping criteria and explained 41.5% of the variance in SRH. Feeling “physically tired” explained most of the variance in SRH in men (32.1%) and women (34.6%). One of the determinants in men is a non-health measure: work satisfaction, explaining 1% of the variance in SRH. Among women four determinants (physical tiredness, number of health problems in the past year, emotional reactions (NHP),³¹ physical mobility (NHP)³¹ explained 43.9% of the variance in SRH. In analyses of men and women together four measures of health (physical tiredness, number of health problems in the past year, physical mobility (NHP),³¹ number of prescription drugs) explained 41.4% of the variance in SRH.

The solution from the stepwise regression (tables 4A and 4B) was checked with other forms of statistical regression (forward selection and backward deletion) and no differences were found. There were no problems of collinearity in the final stepwise models, all tolerance levels were over 0.50.

DISCUSSION

SRH has two key features: it is subjective and evaluative. However, little is known about its determinants. The purpose of these analyses was to identify its determinants, using multivariate regression analyses, in two prospective cohort studies. Six categories of determinants were considered: age, early life factors and family history, sociodemographic variables, psychosocial factors, health behaviour, and health. These blocks explained 36.1% of the variance in men (with 31 determinants) in the Whitehall II data and 36.9% in women (with 33 determinants). In the Gazel data the variance explained was 42.7% in men (with 30 determinants) and 44.3% in women (with 24 determinants). A search for a smaller set of correlates using stepwise regression found five determinants overall in the Whitehall II data and four in the Gazel data explaining 34.7% and 41.4% of the variance in the two studies, respectively. In both studies, it is essentially measures of health that were independent determinants of SRH.

Before further discussion of the results, the similarities and differences in the two cohorts need to be highlighted. Besides cultural differences between the two, there are also other differences. Although they are both occupation based cohorts, the work content of the employees is quite different. The Whitehall II study is based on employees of the British civil service, all of whom have office based jobs. The Gazel study is based on the employees of the gas and electricity public utilities in France (EDF-GDF), with many manual workers, both skilled and unskilled. At baseline (1989), 14.8% of male and 29.1% of female Gazel participants were classified as being “unskilled” workers by their employer in a three level classification: unskilled workers, skilled workers, and managers.²⁷ The essential similarity between the two cohorts is that participants in both studies had comparatively stable employment. Earlier comparative work shows similar social gradients in sickness absence and SRH in the two cohorts.³⁴

An explicit examination of the determinants of SRH is important as it provides an insight into the interpretation of this concept. The main question relates to whether people use conventional indicators of health or if other elements like family history, psychosocial factors, and socioeconomic position influence the assessment of SRH. In other words, to what extent is the perception of one’s health status determined by traditional measures of health and to what extent is it determined by other factors, which may be correlated with SRH at the bivariate level. Our results show that the statistically significant correlations at the bivariate level (tables 2A and 2B) do not translate to independent associations in multivariate analyses. The salient if unsurprising finding of this study is that SRH is mostly a measure of health; measures of early life factors, family history, sociodemographics, and health behaviours make only minor contributions to SRH whereas measures of health and in Gazel also psychosocial factors are important correlates of SRH. It has been suggested that SRH is a multidimensional phenomenon^{5 25}; our results show that it is mostly dimensions of health that are independent determinants of SRH.

Results from the stepwise regression (men and women combined) in the Whitehall II data (table 4A) show that SRH is predicted by measures of current health status (symptom score), health over the past year (sickness absence and recurring health problems over the year), longer term health problems (longstanding illness), along with measure of minor psychiatric morbidity (GHQ). Sex differences in the determinants of SRH in Whitehall II are minor; in men it is the affect balance score that is a correlate rather than the GHQ. Sex differences in the Gazel data are more pronounced. The nature of the psychosocial determinant differs by sex, it is “satisfaction with work” among men and “emotional reactions” among women.

The striking aspect of the Gazel results is the importance of the measure “physically tired” in predicting SRH; it explains about 81.2% of the total variance explained (33.6% of 41.4%) by the full model. This result can be explained in two ways. Firstly, it reflects the essential difference between the two cohorts, the results from the Gazel participants reflecting the fact that their work is more physically demanding. This interpretation is supported by results (not shown but available from the authors) of a stepwise regression that excluded the “physical tiredness” variable, which resulted in the variable “work physically tiring” (table 1B) becoming an independent determinant of SRH. The second interpretation is that the measure of “physical tiredness” is a good comprehensive measure of health, as reflected in the strong correlation with SRH.

In both studies only part the variance was explained by the determinants examined, 34.7% in Whitehall II and 41.4% in the Gazel data. In the Whitehall II data we examined (results not shown) whether other physiological measures (cholesterol, blood pressure) added to the prediction equation and found it not to be the case. However, previous analysis on Gazel data has shown SRH to be linked to diseases, in both cross sectional and longitudinal analysis.³⁵ One could

Policy implications

Self rated health is a widely used measure of health status, mainly used because of its simplicity and its strong relation with outcomes like mortality. Self rated health is believed to be a multidimensional phenomenon even though the determinants of self rated health have rarely been investigated in empirical analysis. Our analysis shows that this simple measure is in effect a valid measure of health.

speculate about the unexplained variance, traditionally seen to be made up of determinants not included in the regression equation, measurement error and individual differences. It is probable that individual history of past health and expected or aspired future health may be potential determinants of SRH.

Comparison with other studies

SRH is a widely used measure in epidemiological studies; a search on Medline (combination self rated health OR self assessed health OR self reported health OR self perceived health AND 2002 as year of publication) found 1991 papers.³⁶ However, an explicit quantitative analysis of the independent determinants of SRH remains elusive in the research literature. An exception is a study by Jylhä and colleagues that examined the correlational structure of SRH in an Italian and Finnish sample.¹⁹ This study used logistic regression to examine the relation between good (very good and fairly good) and poor (average, fairly poor, and poor) health using a smaller range of determinants (number of diseases, functional ability, number of symptoms, problems with vision and hearing, number of medical drugs being used, and education). All determinants were significantly related to SRH in multivariate analyses, the odds ratios associated with “functional ability” and “number of symptoms” were the highest. We were able to extend the analyses by including a wider range of correlates and quantifying the unique variance in SRH explained by the correlates.

Study strengths and limitations

To our knowledge, this is the first study to investigate explicitly the determinants of SRH by quantifying the amount of variance in SRH explained by independent determinants. Data come from two large European cohorts, providing sufficient numbers in each cohort to permit “stepwise regression” without the risk of “overfitting”.^{29–30} We were able to make an objective assessment of the predictive power of several variables. Several categories of determinants were examined in this study. There is evidence to suggest that SRH forms a continuum from poor through average to good health,^{37–38} justifying our use of linear regression models and further permitting comparability of a five point and an eight point measure of SRH. Furthermore, it has been shown that different ways of treating the SRH variable (dichotomous, ordinal, etc) yields similar results.³⁹ A number of limitations should also be noted. Firstly, data here are from employees with stable jobs and cannot be assumed to represent general populations. However, similarity of results in the two contexts provides support for the generalisability of our findings, at least to people in this age group. A recent study examined the 59 year longitudinal trajectory of SRH and concluded that it was comparatively stable until age 50⁴⁰; the predictors of SRH beyond this age might be different. Secondly, our analyses are exploratory in nature; further work with other measures of health is required to delineate the determinants of SRH and possible sex differences. Furthermore, the determinants in the two studies are not strictly comparable; this is rarely possible when international comparisons are being made. Rather than restrict analyses to measures where some degree of equivalence could be established, we chose to use multiple measures from the two studies to permit an assessment of the principal determinants of SRH. Finally, data are not missing at random; the non-response is higher among persons who are socially disadvantaged and likely to be in poorer health. This is likely to lead to some underestimation of the association between SRH and its determinants.

In conclusion, our results show the determinants of SRH to be measures of physical and mental health. The

“multidimensional” nature attributed to SRH mainly translates to it being a holistic measure of health rather than a measure of non-health circumstances as well.

ACKNOWLEDGEMENTS

AS-M is supported by a “Chaire d’excellence” award from the French Ministry of Research. PM is supported by the Academy of Finland. JEF is supported by the MRC (grant number G8802774). The Whitehall II study is supported by grants from the Medical Research Council; British Heart Foundation; Health and Safety Executive; Department of Health; National Heart Lung and Blood Institute (HL36310), US, NIH; National Institute on Aging, US, NIH; Agency for Health Care Policy Research (HS06516); and the JD and CT MacArthur Foundation Research Networks on Successful Midlife Development and Socio-economic Status and Health. We thank all participating civil service departments and their welfare, personnel, and establishment officers; the Occupational Health and Safety Agency; the Council of Civil Service Unions; all participating civil servants in the Whitehall II study; and all members of the Whitehall II study team. The Gazel cohort is supported by Electricité de France-Gaz de France (EDF-GDF). We would like to thank all the staff at INSERM 687; special thanks to Sébastien Bonenfant for his help with the Gazel data.

Authors’ affiliations

A Singh-Manoux, M Zins, M Goldberg, INSERM, U687, HNSM, Saint-Maurice Cédex, France

A Singh-Manoux, P Martikainen, J Ferrie, M Marmot, Department of Epidemiology and Public Health, University College London, UK

P Martikainen, Population Research Unit, Department of Sociology, University of Helsinki, Finland

Conflicts of interest: none declared.

REFERENCES

- 1 **Mossey JM**, Shapiro E. Self-rated health: a predictor of mortality among the elderly. *Am J Public Health* 1982;**72**:800–8.
- 2 **Kaplan GA**, Camacho T. Perceived health and mortality: a nine-year follow-up of the human population laboratory cohort. *Am J Epidemiol* 1983;**117**:292–304.
- 3 **Idler EL**, Angel RJ. Self-rated health and mortality in the NHANES-1 epidemiologic follow-up study. *Am J Public Health* 1990;**80**:446–52.
- 4 **Idler EL**, Kasl S, Lemke JH. Self-evaluated health and mortality among the elderly in New Haven, Connecticut, and Iowa and Washington Counties, Iowa, 1982–1986. *Am J Epidemiol* 1990;**131**:91–103.
- 5 **Idler EL**, Benyamini Y. Self-rated health and mortality: a review of twenty-seven community studies. *J Health Soc Behav* 1997;**38**:21–37.
- 6 **Benyamini Y**, Idler EL. Community studies reporting association between self-rated health and mortality. *Res Aging* 1999;**21**:392–401.
- 7 **Benjamins MR**, Hummer RA, Eberstein IW, et al. Self-reported health and adult mortality risk: an analysis of cause-specific mortality. *Soc Sci Med* 2004;**59**:1297–306.
- 8 **Appels A**, Bosma H, Grabauskas V, et al. Self-rated health and mortality in a Lithuanian and a Dutch population. *Soc Sci Med* 1996;**42**:681–9.
- 9 **Heidrich J**, Liese AD, Lowel H, et al. Self-rated health and its relation to all-cause and cardiovascular mortality in southern Germany. Results from the MONICA Augsburg cohort study 1984–1995. *Ann Epidemiol* 2002;**12**:338–45.
- 10 **Helmer C**, Barberger-Gateau P, Letenneur L, et al. Subjective health and mortality in French elderly women and men. *J Gerontol B Psychol Sci Soc Sci* 1999;**54**:584–92.
- 11 **Idler EL**, Angel RJ. Self-rated health and mortality in the NHANES-1 epidemiologic follow-up study. *Am J Public Health* 1990;**80**:446–52.
- 12 **Walker JD**, Maxwell CJ, Hogan DB, et al. Does self-rated health predict survival in older persons with cognitive impairment? *J Am Geriatr Soc* 2004;**52**:1895–900.
- 13 **Burström B**, Fredlund P. Self-rated health: is it as good a predictor of subsequent mortality among adults in lower as well as in higher social classes? *J Epidemiol Community Health* 2001;**55**:836–40.
- 14 **Grant MD**, Piotrowski ZH, Chappell R. Self-reported health and survival in the Longitudinal Study of Aging, 1984–1986. *J Clin Epidemiol* 1995;**48**:375–87.
- 15 **Hays JC**, Schoenfeld D, Blazer DG, et al. Global self-ratings of health and mortality: hazard in the North Carolina Piedmont. *J Clin Epidemiol* 1996;**49**:969–79.
- 16 **Helweg-Larsen M**, Kjoller M, Thøning H. Do age and social relations moderate the relationship between self-rated health and mortality among adult Danes? *Soc Sci Med* 2003;**57**:1237–47.
- 17 **Idler EL**. Age differences in self-assessments of health: age changes, cohort differences, or survivorship? *J Gerontol* 1993;**48**:S289–300.
- 18 **Idler E**, Leventhal H, McLaughlin J, et al. In sickness but not in health: self-ratings, identity, and mortality. *J Health Soc Behav* 2004;**45**:336–56.

- 19 Jylhä M, Guralnik JM, Ferrucci L, et al. Is self-rated health comparable across cultures and genders? *J Gerontol B Psychol Sci Soc Sci* 1998;**53**:S144–52.
- 20 Benyamini Y, Blumstein T, Lusky A, et al. Gender differences in the self-rated health-mortality association: is it poor self-rated health that predicts mortality or excellent self-rated health that predicts survival? *Gerontologist* 2003;**43**:396–405.
- 21 Deeg DJ, Kriegsman DM. Concepts of self-rated health: specifying the gender difference in mortality risk. *Gerontologist* 2003;**43**:376–86.
- 22 Mackenbach JP, Simon JG, Looman CWN, et al. Self-assessed health and mortality: could psychosocial factors explain the association? *In J Epidemiol* 2002;**31**:1162–8.
- 23 Carlson P. The European health divide: a matter of financial or social capital? *Soc Sci Med* 2004;**59**:1985–92.
- 24 Larsson D, Hemmingsson T, Allebeck P, et al. Self-rated health and mortality among young men: what is the relation and how may it be explained? *Scand J Public Health* 2002;**30**:259–66.
- 25 Borg V, Kristensen TS. Social class and self-rated health: can the gradient be explained by differences in life style or work environment? *Soc Sci Med* 2000;**51**:1019–30.
- 26 Marmot MG, Davey Smith G, Stansfeld S, et al. Health inequalities among British civil servants: the Whitehall II study. *Lancet* 1991;**337**:1387–93.
- 27 Goldberg M, Chastang JF, Leclerc A et al. Socioeconomic, demographic, occupational, and health factors associated with participation in a long-term epidemiologic survey: a prospective study of the French GAZEL cohort and its target population. *Am J Epidemiol* 2001;**154**:373–84.
- 28 Cohen J, Cohen P. *Applied multiple regression/correlation analysis for the behavioural sciences*. 2nd ed. Hillsdale, NJ: Lawrence Erlbaum Associates, 1983.
- 29 Tabachnik BG, Fidell LS. *Using multivariate statistics*. 2nd ed. New York: Harper Collins, 1989.
- 30 Hair JF, Anderson RE, Tatham RL, et al. *Multivariate data analysis*. New Jersey, USA: Prentice-Hall International, 1998.
- 31 Hunt SM, McEwen J. The development of a subjective indicator. *Soc Health Illness* 1980;**2**:231–46.
- 32 Bradburn NM. *The structure of psychological well-being*. Chicago: Aldine, 1969.
- 33 Goldberg DP. *The detection of psychiatric illness by questionnaire*. (Institute of Psychiatry Maudsley monographs no 21). London: Oxford University Press, 1972.
- 34 Fuhrer R, Shipley MJ, Chastang JF, et al. Socioeconomic position, health, and possible explanations: a tale of two cohorts. *Am J Public Health* 2002;**92**:1290–4.
- 35 Goldberg P, Gueguen A, Schmaus A, et al. Longitudinal study of associations between perceived health status and self reported diseases in the French Gazel cohort. *J Epidemiol Community Health* 2001;**55**:233–8.
- 36 Mantzavinis GD, Pappas N, Dimoliatis ID, et al. Multivariate models of self-reported health often neglected essential candidate determinants and methodological issues. *J Clin Epidemiol* 2005;**58**:436–43.
- 37 Manderbacka K, Lahelma E, Martikainen P. Examining the continuity of self-rated health. *Int J Epidemiol* 1998;**27**:208–13.
- 38 Mackenbach JP, Van Den Bos J, Jaung IMA et al. The determinants of excellent health: different from the determinants of ill-health? *Int J Epidemiol* 1994;**23**:1273–81.
- 39 Manor O, Matthews S, Power C. Dichotomous or categorical response? Analysing self-rated health and lifetime social class. *Int J Epidemiol* 2000;**29**:149–57.
- 40 McCullough ME, Laurenceau JP. Gender and the natural history of self-rated health: a 59-year longitudinal study. *Health Psychol* 2004;**23**:651–5.
- 41 Cook WW, Medley DM. Proposed hostility and Pharisac values for the MMPI. *J Appl Psychol* 1954;**38**:414–18.
- 42 Diener E, Emmons RA, Larsen RJ, et al. The satisfaction with life scale. *J Pers Assess* 1985;**49**:71–6.
- 43 Karasek RA, Theorell T. *Healthy work: stress, productivity and the reconstruction of working life*. New York: Basic Books, 1990.

bmjupdates+

bmjupdates+ is a unique and free alerting service, designed to keep you up to date with the medical literature that is truly important to your practice.

bmjupdates+ will alert you to important new research and will provide you with the best new evidence concerning important advances in health care, tailored to your medical interests and time demands.

Where does the information come from?

bmjupdates+ applies an expert critical appraisal filter to over 100 top medical journals. A panel of over 2000 physicians find the few 'must read' studies for each area of clinical interest.

Sign up to receive your tailored email alerts, searching access and more...

www.bmjupdates.com