

eAppendix 1. Educational inequalities in smoking among Japanese adults aged 25-94 years:

nationally representative sex- and age-specific statistics

Supplementary data

Smoking

Smoking habits were surveyed based on the following four categories: (a) "I don't smoke"; (b) "I smoke every day"; (c) "I smoke occasionally but not every day"; and (d) "I have stopped smoking for more than one month". We categorized (b) and (c) as current smoker and (b) as daily smoker.

The indicators of inequality

The following explanations are largely reproduced from Harper and colleagues' work.¹⁻⁴

Scale adjustments for inequality indicators were used for convenience for data presentation.

Estimates of between-group variance and index of disparity were divided by 10 (x0.1). That of mean log deviation was multiplied by 100 (x100). Scale-adjusted values were represented in the tables and figures.

Some inequality indicators, such as concentration index, need a presumption of sequential rank of disparity variable categories like income: e.g., no income, <0-3 million yen, <3-6 million yen, and <more than 6 million yen. However, as for technical school education category, we were not able to determine an exact sequential rank for the category. Therefore, we did not rank the following

inequality indicators (assuming sequential rank): absolute concentration index, slope index of inequality, relative concentration index, and relative index of inequality.⁵

Indicators of relative inequality

Rate ratio—The rate ratio (RR) is probably the most frequently used measure of health inequality and is calculated as $RR = y_1/y_2$, where y_1 is the health status of the least healthy group and y_2 is the health status of the most healthy. In the context of social group comparisons the rate ratio is typically based on comparing, for example, the least socially advantaged group (e.g., the lowest socioeconomic group) with the most advantaged group, as a summary measure of health inequality as one would a measure of range. That is, at each time point it measures the ratio of the rates of the best and worst group (i.e., the relative range), regardless of their social group status. Thus, the specific groups with the best and worst rates may change over time. If there is no inequality, the rate ratio takes on a value of 1.

Index of disparity—The index of disparity (IDisp) summarizes the average difference between several group rates and a reference rate and expresses the summed differences as a proportion of the reference rate. This measure was introduced by Percy and Keppel⁶ and is calculated as $IDisp = (\sum_{j=1}^{J-1} |y_j - y_{ref}|/J) / y_{ref} \times 100$, where y_j indicates the measure of health status in the j th group, and y_{ref} is the health status indicator in the reference population, and J is the number of groups compared. Although in principle any reference group may be chosen, typically the best group rate is

used as the comparison, because that represents the rate desirable for all groups to achieve. In this case, it is not necessary to take the absolute value of the rate differences, since they will all be positive. If there is no inequality, the index of disparity takes on a value of 0.

Mean log deviation—The mean log deviation (MLD) is a measure of general disproportionality. It summarizes the disproportionality between shares of health and shares of population (expressed as a ratio on the natural logarithm scale). For grouped data, they may be written as $MLD = \sum p_j [-\ln r_j]$, where p_j is the proportion of the population in group j , and r_j is the ratio of the prevalence or rate of health in group j relative to the total rate, that is, $r_j = y_j/\mu$, where y_j is the rate of health in group j , and μ is the total population rate. This measure is population weighted, is more sensitive to health differences further from the average rate (because they use the logarithm of the shares of health), and may be used for both ordinal and nominal social groups. The MLD takes on a value of 0 if there is no inequality.

Indicators of absolute inequality

Rate difference—The rate difference (RD) is the absolute inequality between two health status indicators, that is the simple arithmetic difference. It is calculated as $RD = y_1 - y_2$, where y_1 and y_2 are indicators of health status in the least healthy and most healthy groups, respectively. As for the rate ratio, the rate difference is often used to compare the health of less advantaged social groups with that of more advantaged groups. However, we use rate difference as a summary measure of the absolute

gap between the best rate and the worst rate for a given outcome (i.e., the absolute range across all socioeconomic groups), regardless of which two social groups are being compared. In this case, the rate difference will be 0 if there is no inequality.

Between-group variance—The variance summarizes all squared deviations from a population average. In the case of grouped data, this is the between-group variance (BGV), and it is calculated via squaring the differences in group rates from the population average and weighting according to population size: $BGV = \sum p_j (y_j - \mu)^2$, where p_j is group j 's population size, y_j is group j 's average health status, and μ is the average health status of the population. If there is no inequality, the between-group variance is 0. One way to interpret the between-group variance is as the variance that would exist in the population if each individual had the mean health of his or her social group (i.e., if there were no within-social group variation).⁷ The between-group variance may be a useful indicator of absolute inequality for unordered social groups because it weights according to population group size and is sensitive to the magnitude of larger deviations from the population average.⁸

References for supplementary data

1. Harper S, Lynch J. Selected Comparisons of Measures of Health Disparities. Bethesda: NIH Publication No.07-6281. National Cancer Institute; 2007.
2. Harper S, Lynch J, Meersman SC, Breen N, Davis WW, Reichman MC. Trends in area-socioeconomic and race-ethnic disparities in breast cancer incidence, stage at diagnosis, screening, mortality, and survival among women ages 50 years and over (1987-2005). *Cancer Epidemiol Biomarkers Prev.* 2009;18:121-31.
3. Harper S, Lynch J. Methods for Measuring Cancer Disparities: Using Data Relevant to Healthy People 2010 Cancer-Related Objectives. Bethesda, MD: National Cancer Institute; 2005.
4. Harper S, Lynch J, Meersman SC, Breen N, Davis WW, Reichman ME. An overview of methods for monitoring social disparities in cancer with an example using trends in lung cancer incidence by area-socioeconomic position and race-ethnicity, 1992-2004. *Am J Epidemiol.* 2008;167:889-99.

5. National Cancer Institute, Division of Cancer Control and Population Sciences, Surveillance Research Program and Applied Research Program. Health Disparities Calculator, Version 1.2.1 [cited 2012 30 May]. Available from: <http://seer.cancer.gov/hdcalc/>.
6. Percy JN, Keppel KG. A summary measure of health disparity. *Public Health Rep.* 2002;117:273-80.
7. Sen A, Foster J. *On economic inequality*. Expanded edition. Oxford: Clarendon Press; 1997.
8. Chakravarty S. The variance as a subgroup decomposable measure of inequality. *Social Indicators Research.* 2001;53:79-95.

Supplementary Tables

eTable 1. Subject numbers according to age, sex and education group in 2010

Age, years	25-34	35-44	45-54	55-64	65-74	75-84	85-94	Total
Men								
Junior high school	1,593 (6.3)	1,935 (6.0)	1,957 (6.4)	6,177 (17.4)	8,509 (34.0)	7,502 (47.9)	1,965 (58.6)	29,756 (17.7)
High school	10,349 (40.7)	14,492 (45.1)	14,507 (47.1)	17,669 (49.8)	11,350 (45.3)	5,780 (36.9)	916 (27.3)	75,328 (44.7)
Technical college	3,725 (14.6)	3,876 (12.1)	2,310 (7.5)	1,641 (4.6)	604 (2.4)	282 (1.8)	95 (2.8)	12,576 (7.5)
Junior college	893 (3.5)	1,161 (3.6)	1,087 (3.5)	917 (2.6)	412 (1.6)	412 (2.6)	113 (3.4)	5,013 (3.0)
University	7,765 (30.5)	9,524 (29.6)	10,092 (32.8)	8,526 (24.0)	3,972 (15.9)	1,603 (10.2)	258 (7.7)	41,883 (24.9)
Graduate school	1,122 (4.4)	1,146 (3.6)	852 (2.8)	565 (1.6)	189 (0.8)	75 (0.5)	7 (0.2)	3,970 (2.4)
Total	25,447 (100.0)	32,134 (100.0)	30,805 (100.0)	35,495 (100.0)	25,036 (100.0)	15,654 (100.0)	3354 (100.0)	168,525 (100.0)
Women								
Junior high school	1,075 (4.0)	1,146 (3.3)	1,289 (3.9)	5,911 (16.2)	11,014 (40.0)	11,662 (54.9)	4,708 (67.6)	36,927 (19.7)
High school	9,301 (34.5)	15,163 (43.8)	16,568 (50.7)	20,923 (57.2)	13,051 (47.4)	8,246 (38.8)	1,906 (27.4)	85,430 (45.6)
Technical college	4,867 (18.1)	5,312 (15.3)	4,177 (12.8)	3,140 (8.6)	1,318 (4.8)	569 (2.7)	174 (2.5)	19,619 (10.5)
Junior college	5,117 (19.0)	7,867 (22.7)	6,588 (20.1)	4,037 (11.0)	1,224 (4.4)	457 (2.2)	115 (1.7)	25,484 (13.6)
University	6,200 (23.0)	4,836 (14.0)	3,929 (12.0)	2,487 (6.8)	884 (3.2)	311 (1.5)	57 (0.8)	18,764 (10.0)
Graduate school	395 (1.5)	302 (0.9)	151 (0.5)	74 (0.2)	28 (0.1)	8 (0.0)	1 (0.0)	962 (0.5)
Total	26,955 (100.0)	34,626 (100.0)	32,702 (100.0)	36,572 (100.0)	27,519 (100.0)	21,253 (100.0)	6,961 (100.0)	187,188 (100.0)

Values are reported as n (%).

eTable 2. Subject numbers of daily smokers according to age, sex and education group in 2010

Age, years	25-34	35-44	45-54	55-64	65-74	75-84	85-94	Total
Men								
Junior high school	1,053 (9.2)	1,247 (8.6)	1,093 (8.4)	2,640 (20.4)	2,186 (38.2)	1,121 (51.7)	198 (65.3)	9,675 (16.0)
High school	5,492 (48.2)	7,610 (52.7)	7,051 (54.0)	6,787 (52.6)	2,562 (44.8)	809 (37.3)	72 (23.8)	30,673 (50.6)
Technical college	1,724 (15.1)	1,753 (12.1)	1,005 (7.7)	600 (4.6)	133 (2.3)	30 (1.4)	6 (2.0)	5,294 (8.7)
2-year college	391 (3.4)	499 (3.5)	438 (3.4)	283 (2.2)	74 (1.3)	45 (2.1)	8 (2.6)	1,754 (2.9)
University (4-year)	2,546 (22.4)	3,151 (21.8)	3,334 (25.5)	2,520 (19.5)	744 (13.0)	156 (7.2)	19 (6.3)	12,579 (20.8)
Graduate school	184 (1.6)	173 (1.2)	128 (1.0)	85 (0.7)	21 (0.4)	9 (0.4)	0 (0.0)	605 (1.0)
Total	11,390 (100.0)	14,433 (100.0)	13,049 (100.0)	12,915 (100.0)	5,720 (100.0)	2,170 (100.0)	303 (100.0)	60,580 (100.0)
Women								
Junior high school	493 (12.3)	512 (10.0)	330 (8.2)	728 (25.2)	555 (49.1)	309 (64.2)	71 (74.7)	3,167 (17.2)
High school	1,999 (49.9)	3,054 (59.6)	2,509 (62.1)	1,558 (53.9)	467 (41.3)	145 (30.1)	21 (22.1)	10,050 (54.7)
Technical college	731 (18.2)	800 (15.6)	528 (13.1)	272 (9.4)	49 (4.3)	14 (2.9)	2 (2.1)	2,460 (13.4)
2-year college	441 (11.0)	520 (10.2)	453 (11.2)	211 (7.3)	29 (2.6)	3 (0.6)	1 (1.1)	1,701 (9.3)
University (4-year)	330 (8.2)	233 (4.5)	211 (5.2)	120 (4.1)	28 (2.5)	10 (2.1)	0 (0.0)	959 (5.2)
Graduate school	13 (0.3)	4 (0.1)	7 (0.2)	3 (0.1)	2 (0.2)	0 (0.0)	0 (0.0)	30 (0.2)
Total	4,007 (100.0)	5,123 (100.0)	4,038 (100.0)	2,892 (100.0)	1,130 (100.0)	481 (100.0)	95 (100.0)	18,366 (100.0)

Values are reported as n (%).

Table S3. Sex- and age-specific estimates (95% confidence interval) of inequality indicators for current smoking rates by education levels.

Inequality indicators	age group, years							Total
	25-34	35-44	45-54	55-64	65-74	75-84	85-94	
Men								
Rate Difference*	48.9 (45.7, 52.2)	49.3 (46.2, 52.3)	41.1 (37.7, 44.4)	28.1 (24.8, 31.5)	15.5 (10.7, 20.2)	5.5 (1.7, 9.2)	2.8 (-0.9, 6.5)	27.1 (25.7, 28.6)
Between-Group Variance*	12.8 (11.5, 14.0)	12.0 (10.9, 13.1)	8.0 (7.0, 8.9)	2.9 (2.4, 3.4)	0.9 (0.6, 1.2)	0.3 (0.1, 0.5)	0.1 (-0.1, 0.4)	3.4 (3.1, 3.6)
Rate Ratio	3.5 (3.1, 4.0)	3.9 (3.4, 4.4)	3.4 (3.0, 4.0)	2.7 (2.2, 3.2)	2.3 (1.5, 3.3)	1.5 (1.1, 2.1)	1.3 (0.9, 2.0)	2.6 (2.4, 2.8)
Index of Disparity	16.5 (13.3, 19.6)	19.3 (15.5, 23.0)	17.6 (13.4, 21.8)	12.4 (8.2, 16.5)	8.9 (1.6, 16.3)	2.7 (-1.6, 7.1)	1.5 (-3.6, 6.6)	12.0 (10.5, 13.5)
Mean Log Deviation	3.4 (2.9, 3.8)	3.3 (3.0, 3.7)	2.5 (2.1, 2.8)	1.2 (1.0, 1.4)	0.8 (0.5, 1.1)	0.7 (0.3, 1.2)	0.6 (-0.5, 1.8)	1.4 (1.3, 1.5)
Women								
Rate Difference*	44.5 (40.8, 48.2)	45.5 (42.2, 48.8)	22.1 (17.8, 26.4)	7.4 (6.1, 8.7)	2.9 (1.9, 3.9)	3.0 (0.7, 5.3)	1.1 (-0.7, 2.8)	9.7 (8.4, 11.0)
Between-Group Variance*	9.4 (8.4, 10.3)	8.0 (7.2, 8.7)	2.6 (2.2, 3.0)	0.5 (0.4, 0.6)	0.08 (0.03, 0.12)	0.03 (0.01, 0.05)	0.01 (-0.01, 0.02)	0.7 (0.6, 0.8)
Rate Ratio	10.2 (6.6, 16.0)	23.9 (10.8, 52.9)	5.2 (2.6, 10.2)	2.3 (1.9, 2.7)	2.0 (1.5, 2.9)	4.4 (1.4, 13.5)	2.2 (0.3, 15.8)	3.4 (2.5, 4.7)
Index of Disparity	34.7 (15.1, 54.4)	88.7 (10.5, 167.0)	17.0 (-1.2, 35.3)	5.9 (3.3, 8.6)	5.9 (0.4, 11.4)	23.8 (-10.0, 57.6)	7.4 (-27.2, 42.1)	14.4 (6.8, 22.0)
Mean Log Deviation	16.1 (14.4, 17.7)	17.0 (15.3, 18.6)	8.2 (7.0, 9.5)	3.1 (2.3, 3.9)	1.9 (0.8, 2.9)	2.6 (0.5, 4.7)	1.3 (-1.7, 4.2)	3.7 (3.3, 4.1)

* Absolute indicator for inequality

Higher estimates means wider inequality.

Table S4. Sex- and age-specific estimates (95% confidence interval) of inequality indicators for heavy smoking rates among daily smokers by education levels.

Inequality indicators	age group, years							Total
	25-34	35-44	45-54	55-64	65-74	75-84	85-94	
Men								
Rate Difference*	14.7 (11.7, 17.7)	21.7 (15.7, 27.7)	17.6 (10.5, 24.7)	5.0 (-0.8, 10.9)	5.8 (2.2, 9.4)	7.2 (1.3, 13.2)	NA	9.4 (6.3, 12.6)
Between-Group Variance*	16.9 (10.6, 23.2)	17.4 (11.2, 23.6)	3.1 (0.5, 5.7)	2.1 (-0.4, 4.6)	4.0 (-0.7, 8.8)	3.3 (-2.5, 9.1)	NA	2.8 (1.7, 4.0)
Rate Ratio	2.1 (1.8, 2.4)	2.4 (1.7, 3.5)	2.0 (1.4, 3.0)	1.2 (1.0, 1.4)	1.3 (1.1, 1.5)	1.9 (1.2, 2.9)	NA	1.5 (1.3, 1.8)
Index of Disparity	3.5 (1.7, 5.2)	7.4 (1.2, 13.7)	9.1 (1.8, 16.4)	1.3 (0.4, 2.1)	2.3 (0.6, 3.9)	5.0 (-0.1, 10.0)	NA	4.0 (1.6, 6.4)
Mean Log Deviation	2.6 (1.7, 3.4)	1.2 (0.8, 1.6)	0.2 (0.0, 0.4)	0.1 (0.0, 0.2)	0.4 (0.0, 0.8)	1.5 (-0.7, 3.6)	NA	0.2 (0.1, 0.3)
Women								
Rate Difference*	13.0 (9.0, 17.0)	13.6 (9.4, 17.7)	3.5 (-1.0, 8.1)	6.5 (2.5, 10.5)	NA	NA	NA	6.1 (3.9, 8.2)
Between-Group Variance*	11.9 (4.3, 19.5)	12.5 (5.3, 19.7)	0.7 (-1.5, 3.0)	3.1 (-1.1, 7.3)	NA	NA	NA	3.4 (1.7, 5.2)
Rate Ratio	4.1 (2.3, 7.0)	2.9 (2.0, 4.1)	1.3 (0.9, 1.9)	1.9 (1.2, 3.0)	NA	NA	NA	1.8 (1.4, 2.2)
Index of Disparity	12.9 (0.8, 24.9)	7.2 (1.6, 12.9)	1.9 (-1.4, 5.3)	7.0 (-0.7, 14.6)	NA	NA	NA	3.2 (0.3, 6.2)
Mean Log Deviation	6.3 (2.8, 9.9)	4.1 (2.0, 6.2)	0.3 (-0.3, 0.9)	1.3 (-0.4, 3.0)	NA	NA	NA	1.5 (0.7, 2.2)

* Absolute indicator for inequality

Higher estimates means wider inequality.

Abbreviation: NA, not applicable because evaluation using three or more educational categories was not available.