**VALIDATION OF THE ICE INDEX FOR ASSESSING SOCIAL CLASS IN THE ADULT POPULATION**

Antonio Cabrera de León (1, 2), María C Rodríguez Pérez (1), Santiago Domínguez Coello (1), Concepción Rodríguez Díaz (1), Cristobalina Rodríguez Álvarez (2), Armando Aguirre Jaime (1) and the CDC group.**

(1) Primary Health Care Unit. Nuestra Señora de la Candelaria University Hospital. Santa Cruz de Tenerife. Spain.
(2) Área de Preventive Medicine Preventiva. University of La Laguna. Spain.

**ABSTRACT**

Validation of the ICE Model to Assess Social Class in the Adult Population

**Background:** Social class has commonly been defined by the type of employment and it is assessed as a categorical variable. However, this approach has a number of drawbacks. The objective of this article is to develop and validate a readily standardizable quantitative indicator of social class and to show its ability to measure the impact of social class as a health determinant.

**Methods:** In 6729 individuals we measured income, crowding index, education, occupation and employment status. Two models were adjusted to study the neighborhood, dietary pattern and health problems.

**Results:** The model that included only income, crowding index and education (ICE) yielded an indicator that correlated with age (r = –0.28; p < 0.001) and consumption of potatoes (r = –0.17; p < 0.001) and salads (r = 0.10; p < 0.001). This indicator estimated that poor social classes were at significant risk for unemployment (OR=5,2), blue collar jobs (OR=40,9), residing in poor neighborhoods (OR = 30.2), low salad consumption (OR = 2.2) and high consumption of potatoes (OR = 4.5). They also had, especially in women, a higher risk of sedentarism (OR = 1.8), obesity (OR = 4.4), metabolic syndrome (OR = 3.4) and diabetes mellitus (OR = 2.0).

**Conclusions:** The ICE index was valid, not based on occupation or employment status, readily standardizable, and suitable for measuring social class and its impact of on health.

**Keywords:** Social class. Poverty.

* Funded by the Canary Island Research and Health Foundation and by the Health Research Fund (PI 070934)
** CDC group: Basilio Anía Lafuente (3), Francisco Hernández Díaz (1), Carlos Borges Álamo (1), Buenaventura Brito Díaz (1), Delia Almeida González (1), Ana González Hernández (1), Lourdes Carrillo Fernández (1), José Carlos del Castillo Rodríguez (1), Noelia Fernández Ramos (3), and José Juan Alemán Sánchez (1).
INTRODUCTION

The World Health Organisation considers poverty as the main cause of diseases. Socio-economic inequalities are the factors that best explain differences in health within a single society\(^1\textsuperscript{,}^2\textsuperscript{,}^3\). The richer and more educated social groups acquire healthy living habits earlier, hence the risk factors and disease diminishes in them too, while these remain high in the less favoured social classes\(^4\).

Irrespective of where it is measured, health status and mortality rates are far worse among poor social classes, although the magnitude of the differences between rich and poor depend on the distribution of wealth in each community and on the educational opportunities available in this community\(^5\). Traditionally, social classes have been defined by the kind of job that one does and this continues to be used by most authors to measure social class by categories of either occupation alone\(^6\textsuperscript{,}^7\) or in combination with other variables like education and income\(^8\textsuperscript{,}^9\). But this has many disadvantages, the first of which is that the classification and social importance of occupations changes over time and that it is difficult to standardise social classes to compare different moments in time, even within the same country. Moreover, occupation is classified in different ways in different countries, which is yet another obstacle for standardising it and for comparing different communities. In Spain, a classification has been proposed for measuring social class\(^9\), inspired by an earlier proposal made by Goldthorpe\(^10\). This measurement, based on professional occupation, is just as difficult to standardise, in fact the original author himself modifies it depending on the objectives of the study\(^11\). Another disadvantage is that it does not allow the classification of individuals declaring certain employment statuses (housework, students, unemployed, retired, disabled, etc.). A precarious employment status, particularly unemployment, implies a lack of income and material resources and, therefore, reflects material and social deprivation\(^12\).

Education, or rather a lack of education, is an indication of material and social deprivation\(^13\), the use of which in measuring social class is still of maximum importance but it is left out by classifications based solely on professional occupation. Education and disposable income are important components of social class, and the lack of both of these is associated with a greater frequency of health problems\(^13\). Another variable associated with the lower social classes is crowding, the presence of which is associated with poverty, a greater demand for health care\(^13\), residence in areas with high pollution levels\(^14\) or low levels of anthropometric development in childhood (failure to thrive)\(^15\). Some authors include this as a main component in defining social class\(^16\).

In short, social class can be measured in different ways through its multiple components, but it has been measured by categories since the beginning as this is implicit in the very concept of class. There have been two major schools of thought: Marx and Weber, which have provided a different view of the same concept, but both focus on work\(^17\). It is worth considering that, with the development of the middle classes, the boundaries between social classes become blurred, so it would be advisable to have quantitative measurements of them.

The objective of this paper is to develop a simple quantitative indicator of social class, whose components can be standardised to compare different communities, and to validate it in a large sample of the general adult population and see whether it is suitable for measuring the impact of social class on health.

SUBJECTS AND METHODS

A cross-sectional study of the subjects enrolled in the cohort known as the “Canary Island CDC Study” (CDC are the initials of Cardio-vascular, Diabetes and Cancer) was conducted. These had been selected from the general population between 2000 and 2005. The methodology used has been described elsewhere\(^18\). Sedentary life style was defined as not engaging in at least 30 minutes of moderate...
physical exercise a day\textsuperscript{19} and obesity as suffering a body mass index (BMI) $\geq 30$. The concept of abdominal obesity was also used, based on the abdominal waist/height ratio\textsuperscript{20}. Both the definition proposed by the ATP III\textsuperscript{21} and the one suggested by the IDF\textsuperscript{22} were used for metabolic syndrome. Diabetes mellitus was defined in accordance with ADA\textsuperscript{23} criteria.

The variables used to measure social class were the following:

A. Per capita family income (ratio between the average monthly income earned by all members of the family and the number of members of that family). This was divided into quartiles for use in this study.

B. Crowding index (ratio between the number of people living in the house and the number of available bedrooms). A value of 0 points was then allotted to this if the result was greater than 2. 1 point was given for ratios of 1 to 2 and 2 points if the ratio was less than 1.

C. Education, measured on an ordinal scale that reflects the years of education received: 1 point for illiterate people, 2 points for complete or incomplete primary studies, 3 points for complete or incomplete secondary studies and 4 points for complete or incomplete university studies.

D. Professional occupation, referring to the employment sector in which the participant works. This variable is included in the study because, despite the aforementioned disadvantages, it has been used historically to measure social class and the purpose was to analyse a class model including this variable, and another one that excluded it. The CDC questionnaire includes the National Classification of Occupations\textsuperscript{24}, which distributes subjects in large groups. To weight them, it gives 1 point for agriculture, 2 points for construction, 3 points to manual trades like carpenters or mechanics, 4 points for those working in transport, tourism, catering, household chores and others, and 5 points for the health, education and similar sectors.

E. Employment status at the time of enrolling in the study. For weighting purposes, 0 points are given to the unemployed, 2 to those in employment and 1 to all other situations (retired, students, housewives, employees unable to work for reasons incapacity).

These variables were used to create two mathematical models to measure social class. The first of these, inspired by previous proposals\textsuperscript{9,25}, included occupation. The second excluded the employment variables. That is:

Model 1 = Per capita family income, Occupation, Employment Status, Crowding Index, Education.

Model 2 = Per capita family income, Crowding Index, Education.

**Statistical analysis:** The continuous variables were summarised with their mean $\pm$ SE and the category variables by using the observed and relative frequency, or they are presented as proportions with their 95% confidence intervals (CI\textsubscript{95}). The correlation between continuous variables was estimated with the Pearson coefficient. Variance analysis was used to test the association of continuous variables with the category ones and Pearson’s chi squared test to study the association between category variables.

As it has been repeatedly shown that metabolic syndrome\textsuperscript{21} is associated with low social classes\textsuperscript{26-28} in developed societies, in order to derive these models, we started out by checking the association of each component variable with metabolic syndrome in 10% of the individuals studied (n =673), 23% of whom suffered the syndrome (n =155). As all the variables are associated with the syndrome, the two models were adjusted with multi-variant logistic regression, using metabolic syndrome as the dependent variable. The indicator produced by each model was obtained by totalling the points associated with each of its components, assigned according to the value of its rounded off regression coefficient. Then, aware of the fact that living in a rich neighbourhood (luxury estate) or a poor one (social housing estate) is an accepted indicator of social class\textsuperscript{26,30}, the participants that lived in two neighbourhoods that unmistakably meet
these criteria (n=369) were used to conduct an analysis of type II receiver operator curves (ROC), with residence as the status variable, comparing the C statistic or area under the curve generated by the two indicators of social class.

The indicator with the best results (model 2) was validated in this paper by checking its capacity to estimate the following relative risks in the remaining subjects of the sample (n= 6729 - 673 = 6056): (a) unemployment\(^12\), (b) poorly qualified professional occupation\(^9\), (c) residing in a poor neighbourhood\(^29,30\), (d) following typical dietary habits of poor social classes: high consumption of potatoes (grams/day)\(^31\) and pulses (grams/day)\(^32\), together with a low salad consumption (grams/day)\(^33\) and (e) presenting health problems that are presently associated with poverty (sedentary life style\(^4\), obesity\(^6,35\), metabolic syndrome\(^26,36\) and diabetes\(^37\)). The relative risks were estimated with the odds ratio (OR). To this end, logistic regression models adjusted for age and sex were used, with social class as the independent variable, expressed in quintiles and in which the dependent variable on each occasion was: (a) unemployed against working person [only the population aged 30 to 60 were taken for this variable, n= 5,578]; (b) poorly qualified occupation against highly qualified occupation [categories IVb and V of the SEE\(^9\) versus categories I, II and IIIa, taking only the population aged between 30 and 60 years old]; (c) potato consumption: the top quintile versus the lowest; (b) pulse consumption: top quintile versus the bottom; (c) salad consumption: top quintile versus the bottom; moreover, logistic models adjusted for age were also separately adjusted for men and women with the following dependent variables: (d) sedentary life style; (e) obesity; (f) abdominal obesity; (g) ATP III metabolic syndrome; (h) IDF metabolic syndrome IDF; (i) diabetes mellitus type 2.

The calculations were done with the Spanish SPSS\(^\circ\) version 15 statistical software package.

RESULTS

The population studied was made up of 3,816 women and 2,913 men. Table 1 shows the distribution of the variables used to develop the different social class models. It shows that there are significant differences between sexes with regard to monthly financial income, occupation and employment status.

The two social class models clearly discriminated between residing in rich and poor areas (figure 1). As model 2 showed the greater area under the ROC curve, it was chosen for validation. This model was called by the acronym ICE (Per capita family Income, in quintiles + 2 x Crowding Index + 3 x Education). Measured as a continuous quantitative variable, the ICE model presented a mean of 13.35±3.42 with no significant differences between sexes, with a range of values of between 4 and 21, and it showed inverse correlation with age (r= -0.28; p<0.001), with potato consumption (r= -0.17; p<0.001) and with the consumption of pulses (r= -0.03; p=0.01), and direct correlation with salad consumption (r = 0.10; p= 0.001). Its sensitivity to detect residence in a poor neighbourhood reached 97% for values of below 10 and 100% for those below 8.

In table 2, we can see that, in bi-variant analysis, the poor social classes (quintiles 1 and 2 of the ICE index) presented greater age, higher unemployment, worse professional occupation, they ate less salad, more potatoes and more pulses. Social classes generally lived in neighbourhoods that were in line with their income, even though there were small percentages of upper classes (10%) and lower classes (5%) living in the opposite neighbourhood from what one would expect.

The multi-variant analysis showed (table 3) that, even adjusting the possible effect of age and sex, the ICE index identified the poor social classes as those that face a high risk of unemployment, poorly qualified professional occupation and dietary habits with a low consumption of salads and a high consumption of potatoes. The consumption of pulses and the risk of living in a poor neighbourhood were also higher among these classes.
Adjusting for the effect of age and stratifying the analysis by sexes, the ICE model also identified poor classes as those that presented the greatest risk of suffering health problems, that is, those that presented a higher risk of sedentary life style, obesity, abdominal obesity, metabolic syndrome and diabetes mellitus (table 4). The gradient between rich and poor classes was steeper among women, who presented higher risks than men for each of these health problems.
Figure 1. Area under the curve obtained with 2 models to predict area of residence

<table>
<thead>
<tr>
<th>Social Class</th>
<th>Area under the curve (IC$_{95%}$)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1</td>
<td>0.787 (0.736, 0.838)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Model 2 (ICE)</td>
<td>0.812 (0.763, 0.861)</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

Model 1: Education, Income, Crowding Index, Professional Occupation, Employment Status.
Model 2 (ICE) = Income, Crowding Index, Education.
VALIDATION OF THE ICE INDEX FOR ASSESSING SOCIAL CLASS IN THE ADULT POPULATION

Table 3

Estimation of relative risk of unemployment, having a poorly qualified occupation, living in a poor neighbourhood and of having certain food consumption habits in each quintile of social class. Risks are estimated using logistic regression adjusted for age and sex, expressed as OR (CI95%).

<table>
<thead>
<tr>
<th>Social Class (ICE)</th>
<th>Quintile 1</th>
<th>Quintile 2</th>
<th>Quintile 3</th>
<th>Quintile 4</th>
<th>Quintile 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unemployed status</td>
<td>5.4 (3.4, 8.6)</td>
<td>4.5 (2.9, 7.1)</td>
<td>3.0 (1.8, 5.0)</td>
<td>2.1 (1.3, 3.5)</td>
<td>1</td>
</tr>
<tr>
<td>Poorly qualified occupation</td>
<td>40.9 (28.7, 58.4)</td>
<td>18.0 (13.7, 23.7)</td>
<td>7.4 (5.8, 9.5)</td>
<td>3.4 (2.8, 4.3)</td>
<td>1</td>
</tr>
<tr>
<td>Residence in poor neighbourhood</td>
<td>23.1 (8.3, 64.5)</td>
<td>39.2 (11.9, 76.9)</td>
<td>16.1 (6.6, 39.5)</td>
<td>3.3 (1.8, 6.4)</td>
<td>1</td>
</tr>
<tr>
<td>Low salad consumption1</td>
<td>2.2 (1.7, 3.0)</td>
<td>2.2 (1.7, 2.9)</td>
<td>1.8 (1.4, 2.4)</td>
<td>1.1 (0.8, 1.4)</td>
<td>1</td>
</tr>
<tr>
<td>High potato consumption1</td>
<td>4.5 (3.4, 6.1)</td>
<td>3.4 (2.6, 4.6)</td>
<td>3.4 (2.5, 4.6)</td>
<td>1.6 (1.2, 2.1)</td>
<td>1</td>
</tr>
<tr>
<td>High pulse consumption1</td>
<td>1.6 (1.2, 1.3)</td>
<td>1.7 (1.2, 2.3)</td>
<td>1.8 (1.3, 2.5)</td>
<td>1.3 (0.9, 1.7)</td>
<td>1</td>
</tr>
</tbody>
</table>

1Lower quintile for salad consumption and higher for potato and pulse consumption.

Table 4

Estimation of the relative risk of suffering from sedentary lifestyle, obesity, metabolic syndrome and diabetes mellitus in the lowest quintile in comparison with the maximum social class quintile according to ICE. Risks were estimated as OR (CI95%) and adjusted for age.

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>Men</th>
<th>OR (CI95%)</th>
<th>p</th>
<th>Women</th>
<th>OR (CI95%)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sedentarism</td>
<td></td>
<td>1.6 (1.2, 2.2)</td>
<td>&lt;0.001</td>
<td></td>
<td>1.8 (1.4, 2.4)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Obesity</td>
<td></td>
<td>1.8 (1.2, 2.4)</td>
<td>&lt;0.001</td>
<td></td>
<td>4.4 (3.2, 6.1)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Abdominal Obesity*</td>
<td></td>
<td>2.2 (1.7, 2.9)</td>
<td>&lt;0.001</td>
<td></td>
<td>5.4 (4.1, 7.2)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>ATP Metabolic Syndrome</td>
<td></td>
<td>1.3 (1.0, 1.8)</td>
<td>0.100</td>
<td></td>
<td>3.4 (2.4, 4.8)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>IDF Metabolic Syndrome</td>
<td></td>
<td>1.4 (1.1, 1.9)</td>
<td>0.010</td>
<td></td>
<td>2.5 (1.9, 3.4)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Diabetes Mellitus</td>
<td></td>
<td>1.8 (1.1, 2.8)</td>
<td>0.016</td>
<td></td>
<td>2.0 (1.2, 3.1)</td>
<td>0.004</td>
</tr>
</tbody>
</table>

*Waits / height ratio > 0.55. ATP = Adult Treatment Panel III. IDF = International Diabetes Federation.

DISCUSSION

We have proved the validity of the ICE index for measuring social class. It produces a simple indicator that is easy to obtain from interviews. One traditional problem of models of the structure of society that divide them into social classes is their lack of empirical validation, plus the fact that the few examples of validation cover models based on professional occupation. Having tested the validity of ICE, one must highlight the fact that, apart from being simple to obtain, it can also be applied to any person, without presenting difficulties for estimating the results for individuals that declare certain employment statuses, as was mentioned before. Moreover, as it does not include professional occupation, this indicator has the advantage of avoiding the aforementioned problems concerning occupation: it can easily be standardised and used to compare different communities. Of the components of the ICE model, per capita family income compiles more accurate information on an individual’s disposable income than the information provided solely by personal income; education is a classic component of social class as it is not only a person’s wealth, but also his know-how and his/her culture too, that defines this; finally, including the crowding index in the measurement of social class improves the information about an individual’s lifestyle.

There is not a “gold standard” for social class to validate a theoretical construct that attempts to measure it, but the employment and occupation variables are the most widely accepted for this purpose. Well, ICE shows a 5 times greater risk of unemployment among the poorer classes, despite the fact that the information was gathered in years that, historically, the unemployment rate in the Canary Island Region was at its lowest. And as for professional occupation, ICE detects a higher than 40% risk of the poor classes holding a poorly qualified job. Moreover, there is a wide set of well identified variables that are widely accepted as indicators of social class. In this paper, we have opted to use neighbourhood, dietary habits and some typical health problems of social class to validate the ICE model. In particular, neighbourhood has repeatedly been shown to be a marker of class both in Spain and abroad. In our case, the results clearly show the validity of ICE for finding an association between each social class and the kind of neighbourhood where they live. Detecting the OR with such high values is infrequent in any investigation and identifies the
strong association between each social class and its predictable environment. With the information available, it is not possible to deduce the reasons that lead 10% of the rich class to live in a poor neighbourhood. The complex networks of family, social, economic relations, etc., are determining factors that may explain this variability. Moreover, the sky high house prices in Spain enable us to speculate that there are layers of population that live in lower class neighbourhoods, despite the fact that their education and income would place them in the upper class.

With regard to dietary habits, we have opted to measure these with food items whose consumption we know that varies with social class. In rich countries, potato consumption is clearly higher these days among the poorer classes\textsuperscript{31}. In this population, adults from the Canary Islands, this has also been repeatedly confirmed with models of social class based on professional occupation\textsuperscript{43, 44}. The ICE model shows that a diet with a high potato consumption is three times as frequent among the poorer classes than among the rich (28% among the poor versus 8% among the rich), which, when adjusted for age and sex, gives an OR of over 4 between both classes. And it also shows a clear linear trend of diminishing consumption as the social class of the population climbs.

Pulses too, are a group of food items that have been identified as being preferably consumed among the poorer classes, both in the Canary Islands\textsuperscript{45, 46} and among other Spanish populations\textsuperscript{32}. In this group, ICE shows that a high consumption of pulses among the poorer classes is approximately twice as frequent as among the rich. The estimation of relative risk of consumption does not show such high values in this case as for potatoes, but it does show a 60% increase in these risks among the poorer classes. The fact that the consumption trend is not linearly diminishing when class rises may be attributed to the survival of the Mediterranean Diet and dietary advice to maintain a frequent pulse intake, which seem to be followed by a large proportion of the upper classes.

Finally, concerning food, the salad group has repeatedly been identified as showing more frequent high consumption among the richer classes both in Spain\textsuperscript{45, 46} and abroad\textsuperscript{31}. In concordance, ICE identifies a low consumption of salads among the poorer classes twice as frequently as among the rich, and the OR adjusted for age and sex present values higher than 2 between these classes. The linear trend of low salad consumption is clearly decreasing as social class increases.

We have analysed health problems that are highly prevalent in modern societies, for which their greater frequency among poorer classes had been previously shown: sedentary life style\textsuperscript{34}, obesity\textsuperscript{6, 35}, metabolic syndrome\textsuperscript{26, 36} and diabetes mellitus\textsuperscript{37}. Stratifying by sexes, ICE shows not only that it is the poorer classes that face a greater risk, but also that the gradient between poor and rich classes has greater health repercussions on women. The higher risk among women has been described in Spanish populations concerning perceived health\textsuperscript{45}. In other populations it has been shown that social structure (socio-economic differences, family support, etc.) is a more important factor for women’s health, while behavioural factors (alcohol, physical activity, etc.) are more important for men’s health\textsuperscript{46}.

As a limitation to our study, one could argue that the selection of ICE based on areas under the curve using a status variable (residence), to then validate it for the same variable for which it is selected, is a circular argument. However, we consider that we have validated ICE with the results obtained for unemployment, professional occupation, dietary habits, the health problems analysed and the distribution of these by sexes. The results obtained concerning the residence are shown because the differences between classes are very important, and only confirm the validation of ICE in comparison with other class markers.

In conclusion, ICE is an indicator of social class that is easy to develop and its validity has been proven in the general adult population of the Canary Islands. As it does not include professional occupation among its components, it can be easily standardised and it offers a simple means of measuring social class quantitatively in studies that require an analysis of the impact of social class on health.

**BIBLIOGRAPHY**


