ABSTRACT Although socioeconomic status is one of the essential factors in much health research, it is one of the most difficult constructs to measure. The aim of this study was to develop a relatively accurate, easy-to-calculate method to estimate socioeconomic status (SES) for health research in Iraq as well as similar profile countries in the developing world. After reviewing various methods of calculating SES, a new method is proposed based on the 3 main variables: education, occupation and wealth/income. Additional refinement was done based on experience and job status. Content validity was calculated through experts’ opinions and reliability through test/retest. Based on experts’ opinions, it was found that the calculated content validity ratio for the SES index was relatively high and the correlation coefficient reliability for the six variables (education, occupation, income, wealth, experience and job status) was very high. Thus, the suggested method is reasonably valid, reliable as well as easy to calculate.

Élaboration d’un indice socioéconomique pour la recherche en santé en Iraq
Introduction

Socioeconomic status (SES) has been shown by many studies to be related to almost all health conditions as well as their risk factors. These include but are not limited to: overall mortality, child mortality, maternal mortality, chronic diseases, communicable diseases, health behaviour and use of health services (1-4). Health researchers are usually required to analyse their finding according to SES. Outside the health field, studies have also shown that SES is related to many important aspects of life including language development, political participation, vulnerability and crime rates (5).

SES can be defined as “one’s access to financial, social, cultural, and human capital resources” (6). It has also been defined as “the relative position of an individual or family within a hierarchical social structure, based on their access to, or control over, wealth, prestige, and power” (7).
Although SES is one of the essential factors in the majority of health research, it is still one of the most difficult constructs to capture. The difficulty arises from the fact that SES is not a single variable like height or weight that can be easily measured. SES is the result of the interaction of many variables, which are themselves difficult to measure, such as income, wealth and occupation (6,8–11). This is especially problematic for health researchers who are not specialized in social sciences and just want a relatively accurate yet easy-to-calculate method to measure SES in their study (12).

The “big 3” variables of SES are education, occupation and income and/or wealth (6,13–15). When examining these 3 variables, we can see that they have 2 dimensions: the first is resourcefulness (education/culture) and the second is resources (income/wealth). Occupation can be viewed as an intermediate variable that can affect and be affected by both other variables (16): Education/culture ⇄ Occupation ⇄ Income/wealth.

In most cases, these 2 dimensions go together, therefore people with high education are more likely to have higher income and vice versa. But in some cases, they may depart from one another; for example, someone with only primary level education may have a successful private business and amass a great amount of wealth. On the other end of the spectrum, a newly graduated medical doctor may only have negative wealth (debts).

Although we have 2 distinct dimensions in the SES construct, we usually need to reduce it to one number to assign people into socioeconomic classes or levels. This number will be even more useful if it is on a numerical scale; then it would be more amenable to robust statistical analysis (6,8,9,11).

Almost all researchers agree that the 3 variables play a major role in defining a person’s SES but the way they use these variables differs to a great extent. Health researchers have used different methods, from assigning a linear number for each variable to the use of regression analysis, and more recently, principal component analysis to calculate the index (9,17).

Up to now, there has been no reasonably accurate and easy-to-calculate method to measure SES in Iraq that can help health researchers calculate SES for their studies. The aim of this study therefore was to develop such a method for Iraq and possibly for neighbouring and other developing countries.
Methods

We reviewed the currently used methods for calculating SES index from which we developed a new, relatively accurate and easy-to-calculate SES index for health research in Iraq and estimated its validity and reliability.

Development of SES equation

Based on the aforementioned concept of SES (6,13-16), we can create a simple equation that will calculate the SES as follows:

\[ \text{SES} = \text{Education} + \text{Occupation} + \text{Wealth/income} \]

The above variables need to be converted into numbers. The coding of education levels in this study was based on the UNESCO 2011 international standard classification of education (18). Education levels are divided into 8 levels. The lowest is illiterate which has the value of zero and the highest is doctoral degree (PhD) or equivalent and is given the value of 7 (Table 1).

For occupation, we assigned scores to the most common job titles available in Iraq and neighbouring countries based on the international standard classification of occupations (19). Jobs were classified into 6 categories and a score assigned to each (Table 2). The basic scores were adopted from other studies (20–22) but modified according to the relative importance and prestige of each occupation in Iraq. Modification of the scores was done according to the list of colleges/departments ranked by the required high school score for the academic year 2014-2015 (i.e. minimum grades required for admission to colleges/departments in the Kurdistan Region of Iraq). Every year tens of thousands of students together with their families decide on which occupation/subject to choose. It represents a community’s view of the relative value of each occupation/subject. In case of the wife and students, the occupation of the head of the household is used in calculations.

Although education and occupation account for most of the SES score, the equation was made more sensitive by adding wealth and income. The 2 variables to reflect wealth and income were house and car ownership respectively. However, car and house ownership do not have equal weights and their weights are also much lower than those assigned for education and occupation. Thus house ownership was given a weight of 0.5 and car ownership a weight of 0.1. Hence our equation became:
SES = Education + Occupation + House ownership * 0.5 + Car ownership * 0.1

To further refine our equation, we added another variable for age (which can represent experience and can affect income/wealth). For ease of data collection and calculation, we deducted 20 from the current age and divide by 100 to produce a reasonable weight. The next version of the equation was therefore:

SES = Education + Occupation + House ownership * 0.5 + Car ownership * 0.1 + (age-20)/100

To accommodate special cases (retired/unemployed people/deceased head of household), we assigned a weight for previous job by deducting 1 unit. Our final equation was therefore:

SES = Education + Occupation + House ownership * 0.5 + Car ownership * 0.1 + (age-20)/100 - Retired/unemployed/deceased

**Validation**

To validate the proposed SES method, we tried a number of scenarios to demonstrate the limits of each variable, i.e. different education levels, different occupations, house and car ownership, age and employment status.

**Validity**

To estimate the content validity of this new method, a modified Lawshe content validity ratio (CVR) (23) was used.

\[
\text{CVR} = \frac{\text{ne} - \text{N2}}{\text{N2}}
\]

where ne denotes number of expert panellists indicating an “essential” component and N2 denotes total number of experts. The new method was given to 12 experts in the field of public health.
health during a workshop that was held to estimate its content validity. They were asked whether any of the 6 components of the equation (education, occupation, income, wealth, experience and job status whether retired, unemployed or deceased) were: not needed, needed but not essential or essential. Additionally, their opinion was sought about the weights assigned to each component: whether the weights were underestimated, overestimated or appropriate. If underestimated or overestimated, they were asked to provide their estimate. Their feedback was analysed using the above formula and their recommendations were taken into consideration.

**Reliability**

To estimate reliability, a specially designed questionnaire (Figure 1) was administered twice to the same sample of 54 individuals, selected from employees of the College of Medicine and Nursing at Halwer Medical University in Erbil, and considered to be representative of different SES levels with different values for the 6 variables. Correlation coefficients were calculated for the 6 variable of SES. Data were collected by face-to-face interview.

**Ethical considerations**

The study was approved by the research ethics committee of the College of Medicine, Hawler Medical University.

Informed oral consent was obtained from each participant in the reliability analysis.

**Results**

Our final equation to calculate SES was:

\[
\text{SES} = \text{Education} + \text{Occupation} + \text{House ownership} \times 0.5 + \text{Car ownership} \times 0.1 + \frac{(\text{age}-20)}{100} - \text{Retired/unemployed/deceased}
\]

The minimum score would be 0 for an unemployed, illiterate, young manual labourer with no house and car, and the maximum 14.05 for a medical doctor of 65 years (which is the usual age of retirement in Iraq) who has a car and a house. The calculated SES score can either be divided into equal parts (3: high, middle and low socioeconomic levels, or 5: very low, low, middle, high, very high SES), or the distribution of scores in the sample can be checked and a decision made accordingly. The distribution will likely be skewed to the left, i.e. the majority will
be in the lower SES level.

**Validity**

Table 3 shows the different scenarios using the SES equation with different socioeconomic indices. For example, the SES of case number 4 in Table 3 is calculated as follows: Education: 5 (college); Occupation: 6 (medical doctor); House: 0 (No); Car: 1 (Yes); Age: 25 years. Therefore:

\[
\text{SES} = 5 + 6 + (0 \times 0.05) + (1 \times 0.1) + (25-20)/100 - 0 = 11.15.
\]

**Content validity**

Based on the opinion of the experts, the CVR for the SES index was 0.72.

**Reliability**

The correlation coefficients for the 6 components of SES were 0.99 for education, 0.95 for occupation, 0.99 for income, 0.95 for wealth, 0.98 for experience and 0.99 for job status, with an average of 0.98.

**Discussion**

Health researchers have used a variety of approaches to measure SES.

1. Direct question about SES: this is one of the least accurate methods to measure SES because of social desirability bias. Usually the questionnaire includes a questions like “What is your SES?” with 3, or less frequently 5, options to select from, such as high, moderate and low. The problem with this method is that people at either extreme of SES tend to report their SES as moderate because of the sensitivity of the issue (14,24).

2. Using a single variable as a proxy for SES: this has the advantage of simplicity but ignores the fact that SES is a multidimensional construct. There are several commonly used single proxy variables.

1. Education: this is easy to measure but represents only 1 dimension of the SES construct. In addition, people with the same education may have very different SES (14).

2. Occupation: this is probably the best candidate but the problem is usually with how to capture and code it. One difficulty in capturing occupation is that many people, especially those
who work in the public sector, supplement their income with a second job.

3. Income: this is a relatively good indicator of SES but the problem is its sensitivity. People do not like to disclose their income, especially when they are above the median, or have in-kind or multiple sources of income. Wealth is even more problematic to capture than income. An alternative but less direct method is to ask if income is sufficient. The options usually listed include: not enough, marginally enough, enough and more than enough. The problem with this approach is that it is very subjective and is a measure of a person’s contentedness with material wealth rather than its value.

4. Residence: where people live is strongly related to their SES. When people achieve a higher SES, they usually choose another area to live in. The problem with using residence as a proxy for SES is twofold. First, each neighbourhood of the city needs to be ranked according to a scale, usually the value of one square metre of land. Second, using residence is not externally valid, i.e. it is not generalizable and cannot be used for other areas even in the same country. Therefore we need to have specific values for each neighbourhood for all the cities (10,12).

5. Crowding index: although this could serve as a rough indicator of socioeconomic index, it does not always reflect SES. For example, many old people live alone in a big house but they do not have a high SES. Another example, in rural areas having more children is regarded an asset as they represent free unpaid workers and this may not necessarily translate into more room (25).

6. In addition to the variables mentioned above, there are many others that can be used in specific cases to refine SES, for example number of languages spoken and the proficiency level, and computer skills.

3. Creating an index from a combination of variables, more commonly a combination of education, occupation and income/wealth, to represent different aspects of the SES construct. Researchers have used a variety of methods for assigning weights to each variable. The accuracy of these methods depends on how the weight is assigned and how the data are captured.

1. Using linear weight: here an arbitrary weight is given for each variable. The problem with such a method is the relative weight given to each component (26).

2. Regression analysis: to be able to assign weights based on regression, a nationally representative, comprehensive and accurate data set is needed and this may not be available. Another problem with using regression is confounding between the variables leading to over-representation of one dimension of SES or another (17).

3. Principal component analysis: this is a relatively new and promising method of assigning weight to different components of the SES construct but like regression, a nationally representative, accurate and detailed data set is needed (9,17,27).

One of the most frequently used methods for calculating SES in the Middle East is the method
developed by El-Gilany and colleagues (26). Using their method, the researcher has to ask more than 36 questions to calculate SES, which could be a major distraction from the main topic of the study. In addition, they give 3 times the weight to education compared to occupation, with unemployed/retired having a value of zero. Although they assigned weight for the occupation of the wife, they assigned a value of zero for housewife. Thus a doctor whose wife is a primary school teacher will have much lower SES than one whose wife is a doctor, and a civil engineer who has been unemployed for the past 3 months will have a much lower SES compared with his employed civil engineer colleagues. They also assigned equal weights to a television and computer among many other household possessions. However, a computer is a much better indicator for SES than a television or even a car, not because of its monetary value but because it is a proxy for more education (12). Another issue with their method is that they include data about health in their SES calculation, which may not be necessary.

We used a combination of variables to develop our index and assigned relative weights to each component variable.

**Strengths of the proposed index**

1. **Small number of variables:** we used only 6 variables to calculate SES and most of them are already collected for other purposes, e.g. age, education and occupation. This has the advantage of a shorter interview time. According to Ganzeboom and colleagues (20), education and occupation account for more than 90% of the SES weight. That is why we only needed 2 more variables to represent income/wealth, the third big contributor to SES.

2. **Easy to measure variables:** the variables used in our method are easy to measure.

3. **No sensitive questions:** the questions used for data collection are not sensitive and people are not asked about income or wealth.

4. **Content validity:** based on expert opinion, the CRS for the SES index was 0.72 which is relatively high; it is much higher than the critical value suggested by Lawshe (23) for 12 experts, which is 0.56. This indicates that our method of calculating the proposed index sufficiently covered all of the dimensions of the SES.

5. **External validity:** the variables used are not specific to one area and are generalizable so can be used in different governorates of Iraq, and probably other countries in the Middle East and other developing countries with a similar socioeconomic profile. In addition, it can be used to compare SES between different societies because the measure is absolute and not relative.

6. **Time sensitivity:** some variable can have a very strong relation with SES but their effect changes with time. For example, a few years ago, only people with high SES had a mobile phone but today almost everyone has one. Our proposed index avoids such time-sensitive variables.

7. **Reliability:** with an average correlation coefficient of 0.98, our index is fairly reliable because of the nature of the variables selected and the method used to measure them. The high reliability score could be attributed to the fact that data collection involved only 6 variables that are easy to collect, non-sensitive and unambiguous. The Cronbach alpha could not be used to estimate reliability in our method because it assumes unidimension, which does not apply to
our case as SES is a multidimensional construct (28–30).

8. List of common occupations: providing this list ensures that data about occupation are collected in an accurate and detailed way, avoiding vague terms such as “kasb” (meaning bread-winner or earner). This is a common response when the field of occupation is left blank to be filled by the respondent (12).

9. Reasonable weights for each component: education and occupation have most of the weight when calculating SES because the third component of the big 3, i.e. wealth/income, is closely related to them, and these 2 components have almost equal weights (8). We included house and car ownership as proxies for wealth and income respectively. House is probably the most single expensive item a person can own. While a car is typically much less expensive than a house, what is important about car is the cost to run it and its daily maintenance. A person may be able to buy a car but they may not be able to afford the daily expenditure to run it. Nevertheless, home possessions do not make a big difference and cannot move someone from one occupation category to another. For example, a medical doctor who does not have a house will not have a much lower SES than a colleague who does. In addition, we took years of experience into consideration, calculated from age, which can affect wealth/income. However we deducted 20 from the current age of the individual as this is the average age at which people start a job.

10. Including special cases: many of the already used methods of calculating SES usually assign a value of zero for retired, unemployed, student and housewife (26). Accordingly, an engineer who has been unemployed for the past few months may end up with a lower SES than a manual labourer. This pitfall is avoided by deducting one unit from the SES level of the respondent's previous job in case of unemployment or being retired. This is based on the fact that a retired/unemployed person does not lose all the prestige of the previous job but the weight of the previous occupation will be reduced by 1 job category level. On the other hand, for the wife and student, the occupation of the head of the household is used for the calculation. Historically, as well as at present, the position of women in society is often defined by the head of household’s socioeconomic characteristics rather than their own (31).

Limitations of the proposed index

Like any method used for calculating SES, this method also has its own weaknesses. The most obvious is the fact that we are reducing a complex multidimensional construct to one number. Another issue is change over time; the value of education and a degree can change over time and depends on the literacy rate. The same is true for occupation; although relatively stable, some occupation titles have been introduced in the past few years, such as those related to computers, while other occupations may die out. That is why SES needs to be updated from time to time (16,22,32,33).

The proposed equation is thought to be reasonably accurate for most health studies but if the researcher is interested in even more detail, then the education and occupation of the wife should also be taken into consideration. The same equation can be used for calculating the SES of the wife and then the results added together.
Conclusion

Currently, most health research in Iraq uses unreliable and/or invalid methods for calculating SES. Our study aimed to fill this gap by proposing a new method which is valid, reliable and easy to calculate.

Acknowledgements

The authors would like to acknowledge the valuables remarks of Professor Jamal Ameen on a previous version of this manuscript and the faculty members of the Department of Community Medicine, College of Medicine at Hawler Medical University for their valuable feedback.

Funding: None.

Competing interests: None declared.

References


8. Caro DH, Cortés D. Measuring family socioeconomic status: An illustration using data from PIRLS 2006. IERI Monograph Series Issues and Methodologies in Large-Scale


12. Omer W. Health seeking behavior and consumer satisfaction with health care services in Erbil city [PhD thesis]. Erbil: Hawler Medical University; 2012.


