Summarizing Health Inequalities in a Balanced Scorecard

Methodological Considerations

Nathalie Auger, MD, MSc, FRCPC¹

Marie-France Raynault, MD, MSc, FRCPC²

ABSTRACT

The association between social determinants and health inequalities is well recognized. What are now needed are tools to assist in disseminating such information. This article describes how the Balanced Scorecard may be used for summarizing data on health inequalities. The process begins by selecting appropriate social groups and indicators, and is followed by the measurement of differences across person, place, or time. The next step is to decide whether to focus on absolute versus relative inequality. The last step is to determine the scoring method, including whether to address issues of depth of inequality.

MeSH terms: Population statistics; inequalities; epidemiology; health status indicators; information dissemination

RÉSUMÉ

L'association entre les déterminants sociaux et les inégalités de santé est bien connue. Il faut maintenant des outils efficaces pour transmettre ces informations. Cet article présente un de ces outils (le « tableau de pointage équilibré » ou Balanced Scorecard) qui peut être utilisé pour synthétiser les données sur les inégalités sociales. Dans un premier temps, on sélectionne les indicateurs et les groupes sociaux appropriés. Les différences entre les personnes, les lieux et dans le temps sont ensuite mesurées, et l'accent peut être mis sur les inégalités relatives ou absolues. On porte enfin un jugement sur ces différences et les niveaux d'inégalités.

The association between social determinants and health inequalities is well recognized.¹ While much work has been done on the measurement of social inequalities in health,^{2,3} relatively less has been done on methods for disseminating such information. Because of the complexity involved in measuring social inequality,⁴ translating this information into a simple and understandable form can be challenging.

This article addresses how to use the Balanced Scorecard for the purpose of summarizing social inequality data. Briefly, the Balanced Scorecard is an information translation tool originating from the marketing sector.⁵ Examples of applications of the Balanced Scorecard in public health include its use in summarizing surveillance data by health departments,^{6,7} and for policy purposes.⁸ Limitations of the Balanced Scorecard, however, reinforce the need to properly use this tool.⁵

There are six steps in the development of a Balanced Scorecard. The first step is to select appropriate social groups, the second to select indicators, the third to choose the person/place/time dimension, the fourth to select the type of inequality, the fifth to define the scoring system, and the sixth to address depth of inequality. These steps will be discussed in turn.

Step 1: Selection of social groups – the columns of the table

In order to describe inequality in a population, it is crucial to look at data across different social groups. Certain social groups are commonly used because of known inequalities. Failure to stratify by social groups risks masking underlying inequality.⁹ For instance, a community with low overall mortality may have unacceptably high mortality in some areas. Although this problem may seem obvious, it is not unusual to find Balanced Scorecards with indicators not displayed by subgroup.⁸

The choice of social group depends not only on the objective of the Balanced Scorecard, but also on availability of social data. While the ideal situation is one in which the data itself is used to create social groups, databases frequently lack information on social characteristics. In such cases, a geographic proxy may be used, such as Wilkins' neighbourhood income quintile.¹⁰ Wilkins used census data to classify neighbourhoods into

1. Institut national de santé publique du Québec, Department of Social and Preventive Medicine, Université de Montréal, Montréal, QC

2. Direction de santé publique de Montréal, Observatoire montréalais des inégalités sociales et de la santé, Chair of the Department of Social and Preventive Medicine, Université de Montréal

Correspondence: Dr. Nathalie Auger, 190, boulevard Crémazie Est, Montréal (Québec) H2P 1E2, Tel: 514-864-1600, Fax: 514-864-1616, E-mail: nathalie.auger@inspq.qc.ca

Acknowledgements: The authors are indebted to Denis Roy, Eric Litvak, Cristin Muecke, and the reviewers for their contributions to earlier versions of this manuscript.

quintiles based on the proportion of individuals living below the poverty threshold. Similarly, census data can be used to group census tracts according to social indicators such as education, employment, ethnicity rates, etc. Another option is to classify areas with a composite index, such as the Pampalon index which classifies Quebec enumeration areas according to material and social deprivation.¹¹ Since vital statistics registries, disease registries, administrative/ insurance data, and hospitalization data frequently contain residential information such as the postal code, these databases can be linked to the census data, allowing for the calculation of subgroup-specific rates.¹²

Step 2: Selection of indicators – the rows of the table

After choosing the social groups, the next step is to select the indicators. In general, there are two types of indicators: 1) determinants of health, and 2) health status indicators.⁹ Ideally, the literature should be reviewed to identify those indicators that are important to the problem being studied. When such information is not available, innovative techniques can be used to identify indicators. For example, a modified nominal group technique has been used to identify indicators for a Balanced Scorecard on the performance of public health units.¹³ Preferably, indicators should have sufficient variability across social groups and over time. Furthermore, stability of indicators should be considered since stratifying social groups could lead to small numbers.

Ethical concerns should also be kept in mind; indicators can potentially stigmatize groups that may already be socially disadvantaged. For example, reporting on highly stigmatized infectious diseases, such as tuberculosis, could inadvertently worsen the reputation of such areas.

Step 3: Choice of person/place/time

The simplest Balanced Scorecard is one that addresses only one dimension, such as the comparison of social or geographic groups at one point in time.^{6,7} In such a table, the columns correspond to the different social/geographic groups. This comparison allows policy-makers to identify needy areas.

Figure 1. Illustration of a setting characterized by decreasing relative inequality and increasing absolute inequality

This figure describes the trend in inequality for two social groups at two time points. At the first time point, the absolute difference between groups A and B is 6 (7 minus 1). At the second time point, the absolute difference increases to 8 (10 minus 2). Meanwhile, the relative difference between groups A and B at the first time point is 7 (7/1). At the second time point, this ratio decreases to 5 (10/2). Thus, one measure of inequality changes favourably over time, while another changes unfavourably over time, to the point of actually reversing the interpretation of the trend in inequality. Note further that the setting is one in which the health status of both social groups improves over time (diagonal arrows).

On the other hand, a Balanced Scorecard may address changes over time in a social or geographic group. For example, one Balanced Scorecard developed by the UK has a column for “trend since baseline” (long-term trend) and another for “direction of latest data” (short-term trend).⁸ Showing trends over time allows policy-makers to see whether their policies are on track.

An ideal Balanced Scorecard addresses more than one dimension, for example by showing changes both across groups and over time. For the UK Balanced Scorecard, this would mean adding a third column to show how the social group compares to others at a given point in time.

Step 4: Choice of type of inequality

In the literature, a distinction is made between absolute and relative disparities.^{9,14} Absolute inequality generally refers to the additive difference between two indicators. For example, two areas with disease rates of 7% and 10% have a 3% absolute difference. Relative inequality, on the other hand, refers to a ratio between two values. In the above example, relative inequality could be expressed as 10 divided by 7, or a ratio of 1.4. While the calculation and interpretation of absolute inequality is simpler, relative measures provide different information, as illustrated in the following

example. Consider two new areas with disease rates of 1% and 4%. Here, the absolute difference remains unchanged at 3, but the relative difference increases from 1.4 to 4. If measurements over time are also brought in, it is even possible for absolute inequality to decrease, without a coincident decrease, or possibly even an increase in relative inequality (see Figure 1 for an illustration).

It is also possible to have very large relative differences coexisting with small absolute differences.¹⁴ In addition, the perception of inequality is affected by the direction of the indicator, that is, whether the indicator is described in terms of favourable or unfavourable events.¹⁴ For example, *unemployment* rates of 5 and 10% have a larger relative difference than the complementary *employment* rates of 90 and 95%; meanwhile, absolute inequality is the same for both.

These examples show how absolute and relative inequality are different yet complementary concepts. Although not all forms of inequality need be addressed in a Balanced Scorecard, these distinctions must be kept in mind since they may change the interpretation of the data. An excellent review of the differences between absolute and relative inequality, including statistical issues in their measurement, is available online.¹⁴

Step 5: Defining the scoring method

The fifth step is the process whereby social groups are assigned a score. This step involves selecting the nature, scores, and criteria for assigning scores. The nature of scores refers to the use of either symbols or numeric values. Symbols are advantageous because they summarize trends in the data. They are particularly helpful when trends are not obvious in the data. Consequently, symbols make it easier to interpret a large number of indicators. Consider a Balanced Scorecard which has two types of indicators, some where an increase is favourable (e.g., life expectancy) and others where an increase is unfavourable (e.g., infant mortality). In both cases, the interpretation of the trend is different. When a Balanced Scorecard contains many such indicators, there is a greater chance for the data to be misinterpreted when numbers rather than symbols are used. The main disadvantage of symbols is that readers cannot see the values or range of the indicators. The remainder of this discussion focusses on Balanced Scorecards composed of symbols.

The next consideration is the choice of reference. For comparisons at one point in time, the reference could be either the average of all groups⁷ or only the group that is most well off. Another possibility is to compare to groups in other areas.⁶ Similarly, in assessing changes over time, the reference period could be either a specific point in time⁸ or the average trend over time.

Selecting the actual scores is fairly straightforward. For the comparison of groups at one time point, one option is to use the following four scores: 'data favourable compared to other data', 'data unfavourable compared to other data', 'data show no significant difference', and 'insufficient data available'.⁷ For trends over time, one option is: 'data moving in right direction', 'data moving in wrong direction', 'data show broadly constant trend or no significant movement', and 'insufficient data available to determine a trend'.⁸ The symbols themselves must also be chosen. For the aforementioned four scores, one could use, respectively, a check mark, x mark, equivalent sign, and solid triangle.⁸ Other symbols have been assessed in the literature.¹⁵

Once the scores are selected, the criteria for assigning scores must be defined. This step involves deciding what differences or changes are significant for policy. This can be understood through an example, the scoring system of high school report cards. Report cards frequently use letter symbols to grade students. No differentiation is made between students who score 85% or 95%; both receive A grades. The policy implication is that both students are judged to be equally successful and are equally rewarded, even though one student is clearly superior. Similarly, a Balanced Scorecard must be based on differences important for policy. These criteria require the consensus of all partners and should, to the extent possible, be based on evidence. Different criteria may have to be set for each indicator. Statistical significance testing and confidence intervals may be useful adjuncts, but should be used with caution since statistically significant differences may not always be significant for policy.⁵ Similarly, changes important for policy are not always statistically significant.⁵ In fact, the goal of a policy-maker is to implement policies before differences become statistically significant. The main point is that the criteria for assigning scores should be carefully thought out and set *a priori*.

Step 6: Deciding whether to include measures of depth of inequality

The depth or extent of the difference between two measurement points is another important consideration. Depth of inequality is a useful planning adjunct for decision-makers who have to decide how to allocate limited resources; the social groups that are the worst off may be more deserving than other disadvantaged groups. Depth can be assessed for both absolute and relative inequality. Once scores have been assigned, depth can be incorporated by varying the shading of the symbols. One drawback is that introducing depth adds complexity to a Balanced Scorecard. It may be easier to address depth in an appendix. An appendix would also be useful for readers wishing more detailed information.

In summary, the Balanced Scorecard is useful as a means to summarize health inequalities in a population. This article

discusses the methodological issues involved in its development. It should also be acknowledged that there is a need for input from and evaluation with stakeholders. No matter how well conceived, the Balanced Scorecard cannot achieve its objectives without being acceptable and useful to the stakeholders.

REFERENCES

1. Marmot M, Wilkinson RG. *Social Determinants of Health*. Oxford: Oxford University Press, 1999.
2. Regidor E. Measures of health inequalities: Part 1. *J Epidemiol Community Health* 2004;58:858-61.
3. Regidor E. Measures of health inequalities: Part 2. *J Epidemiol Community Health* 2004;58:900-3.
4. Berkman LF, Kawachi I (Eds.). *Social Epidemiology*. New York: Oxford University Press, 2000.
5. Auger N, Roy D. The Balanced Scorecard: A tool for health policy decision-making. *Can J Public Health*. 2004;95(3):233-34.
6. Direction de la Santé Publique de Montréal. *Social Inequalities in Health: Annual Report of the Health of the Population*. Montréal (Canada): 1998.
7. Pageau M, Choinière R, Ferland M, Sauvageau Y. *Le portrait de santé - Le Québec et ses régions*. Québec Public Health Institute, 2001.
8. Department for Work and Pensions. *Opportunity for all: A Summary of the Fourth Annual Report 2002*. London: 2002 Sept. Available online at: http://www.dwp.gov.uk/publications/dwp/2002/oppal-fourth/opp_summ.pdf (Accessed August 15, 2006).
9. Braveman PA. Monitoring equity in health and healthcare: A conceptual framework. *J Health Popul Nutr* 2003;21(3):181-92.
10. Wilkins R, Berthelot J-M, Ng E. Trends in mortality by neighbourhood income in urban Canada from 1971 to 1996. *Health Rep* 2002;13 Suppl:S1-S27.
11. Pampalon R, Raymond G. A deprivation index for health and welfare planning. *Chron Dis Can* 2000;21(3):104-13.
12. Wilkins R. Use of postal codes and addresses in the analysis of health data. *Health Rep* 1993;5:157-77.
13. Robinson VA, Hunter D, Shortt SED. Accountability in public health units: Using a modified nominal group technique to develop a balanced scorecard for performance measurement. *Can J Public Health* 2003;94(5):391-96.
14. Keppel KG, Percy JN, Klein RJ. Measuring progress in Healthy People 2010. *Healthy People 2010 Stat Notes*. 2004;25:1-16.
15. Schunemann HJ, Best D, Vist G, Oxman AD. Letters, numbers, symbols and words: How to communicate grades of evidence and recommendations. *CMAJ* 2003;169(7):677-80. Erratum in: *CMAJ* 2004;170(7):1082.

Received: June 22, 2004

Revisions requested: October 1, 2004 and February 1, 2006

Revised mss: November 2, 2005 and February 22, 2006

Accepted: May 25, 2006