

UC Berkeley

Cal PREPARE

Title

Deriving Criteria Weights for Health Decision Making: A Brief Tutorial

Permalink

<https://escholarship.org/uc/item/4366g5pv>

Authors

Aragon, Tomas J
Dalnoki-Veress, Ferenc
Shiu, Karen

Publication Date

2012-06-08

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/3.0/>

Deriving Criteria Weights for Health Decision Making: A Brief Tutorial

Tomás J. Aragón, MD, DrPH,^{1,2} Ferenc Dalnoki-Veress, PhD³ Karen Shiu, MPH,²
Colleen Chawla, MPP,¹ Jenny Louie, MPP,¹ Greg Wagner, MPP,¹ Barbara A. Garcia, MPA¹

¹ *San Francisco Department of Public Health, San Francisco, CA*

² *University of California, Berkeley School of Public Health, Berkeley, CA*

³ *Monterey Institute of International Studies, Monterey, CA*

Abstract

When it matters most, to make decisions, set priorities, or allocate resources we must rank and select among available options (“alternatives”). To accomplish this we develop and weight criteria, and then evaluate the alternatives against these criteria. More important criteria should have higher weights. For important decisions we should use science-based, objective measurements and methods to derive criteria weights. In this tutorial we cover how to use the analytic hierarchy process (AHP) to derive criteria weights. More specifically, we illustrate how to conduct pairwise comparisons of criteria with respect to importance, likelihood, or preference. For numerical calculations, we use R—an open source programming language for statistical computing and graphics.

Keywords: Analytic hierarchy process, Decision sciences, Multi-criteria decision making, Priority setting, Resource allocation

1. Introduction

To make decisions, set priorities, and allocate resources we often have to rank and select among available options (“alternatives”). To accomplish this, we can develop criteria, weight them, and evaluate the alternatives against them. More important criteria should have higher weights and therefore should be derived using science-based, objective measurements and methods. Using the analytic hierarchy process (AHP) [1, 2, 3], we can derive such weights by conducting pairwise comparisons of criteria with respect to their importance, likelihood, or preference. With the AHP, we can integrate our interpretations of evidence (data, testimony, etc.) with our interpretations of qualitative factors (e.g., ethical values). Furthermore, we can explicitly assess and address value trade-offs.

While the AHP is the most common multi-criteria decision making (MCDM)¹ method in use worldwide, it is not commonly used in health and medical sciences aside from a few exceptions; however, it is starting to gain popularity [4, 5]. The purpose of this tutorial is to introduce the AHP to health professionals and begin explaining how to use pairwise comparisons for deriving criteria weights.

¹Also called multiple criteria decision analysis (MCDA)

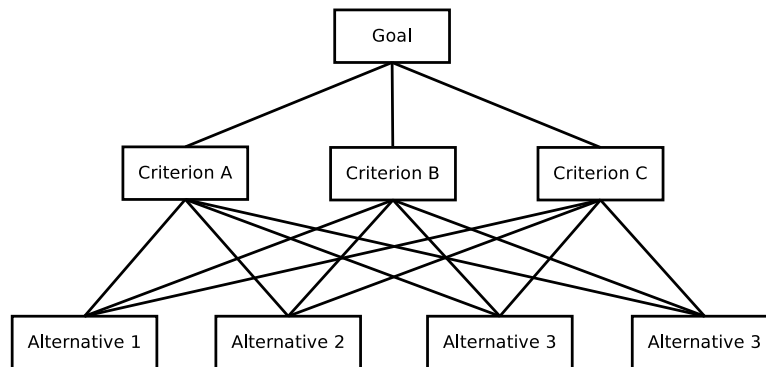


Figure 1: Generic analytic hierarchy process model

2. Methods

2.1. Overview

Many complex decisions require multiple considerations: conflicting requirements, value trade-offs, integration of qualitative data, limited evidence, competing stakeholder input, and time constraints. Much of our formal health methodologic training is in *analysis*: breaking down problems into components and studying their relationships. In contrast, complex decision making requires *synthesis* of components based on our interpretation of existing evidence or assumptions on how components are related. Synthesis requires a systems perspective and a different set of tools than analysis. For complex decision making we should employ a MCDM method [6]. The AHP is a popular MCDM method and involves the following steps:

1. Define the decision making goal
2. Select, organize, and weight criteria (this tutorial)
3. Apply criteria to alternatives and rank alternatives
4. Conduct sensitivity analysis

A generic AHP model is displayed in Figure 1. In pairs, each criterion is compared to each another:

- Criterion A vs. Criterion B
- Criterion A vs. Criterion C
- Criterion B vs. Criterion C

For each pairwise comparison, we ask if one criterion is more important (or effective, likely, preferred, etc.) than the other? If yes, by how much more? To determine the relative “how much more,” we use the fundamental scale—a qualitative, ordinal scale with ratio properties (see Table 1 on the next page). The judgment of “how much more” is based on the qualitative description (moderate, strong, very strong, and extreme) and *not* on the quantitative intensity of the associated values (1, 2, . . . , 9); it is our interpretation of the relative importance of one criterion compared to another. If possible, the interpretation should be guided by the review of evidence (data, testimony, etc.).

Table 1: The fundamental scale for pairwise comparisons (Source: Saaty, 2008, [2])

Intensity (i)	Definition	Explanation
1	Equal importance ^a	Two activities contribute equally to the objective
3	Moderate importance	Experience and judgment slightly favor one activity over another
5	Strong importance	Experience and judgment strongly favor one activity over another
7	Very strong (or demonstrated) importance	An activity is favored very strongly over another; its dominance demonstrated in practice
9	Extreme importance	Evidence favoring one activity over another is of the highest possible order of affirmation
2, 4, 6, 8	For compromise between above values	
$1/i$	Reciprocals of above intensities	

^a Or likelihood, preference, or other factor

If two criteria are *equal*, then these “two activities contribute equally to the objective.” If a criterion is *moderate* over another, then “experience and judgment slightly favor [this] activity over another.” If a criterion is *strong* over another, then “experience and judgment strongly favor [this] activity over another.” If a criteria is *very strong* over another, then “[this] activity is favored very strongly over another; its dominance demonstrated in practice.” If a criterion is *extreme* over another, then “evidence favoring [this] activity over another is of the highest possible order of affirmation.”

The intensity score (i) is a ratio with valid reciprocal values ($1/i$). For example, if Criterion A, compared to Criterion B, is scored with intensity value i , then Criterion B, compared to Criterion A, has the reciprocal intensity value $1/i$. For each pairwise comparison, only one valuation is required. For n criteria, there will be $n(n-1)/2$ comparisons. To improve validity and practicality, n should not be more than 7 (± 2). If the number of criteria seem “too many,” then try clustering those that are equal—these pairwise comparisons will receive an intensity value of 1.

2.2. Derivation of criteria priority weights

Figure 2 on the following page shows the measurement tool for the pairwise comparisons of Criteria A, B, and C. Criterion A is *strong* in importance compared to Criterion B; Criterion C is *moderate* in importance compared to Criterion A; and Criterion C is *very strong* in importance compared to Criterion B.

Next, we create a comparison matrix using our results. We will do this incrementally starting with the diagonal values which are always 1s.

$$\begin{array}{c}
 \text{Criterion A} \\
 \text{Criterion B} \\
 \text{Criterion C}
 \end{array}
 \begin{array}{ccc}
 \text{Criterion A} & \text{Criterion B} & \text{Criterion C} \\
 \left(\begin{array}{ccc}
 \mathbf{1} & & \\
 & \mathbf{1} & \\
 & & \mathbf{1}
 \end{array} \right)
 \end{array}$$

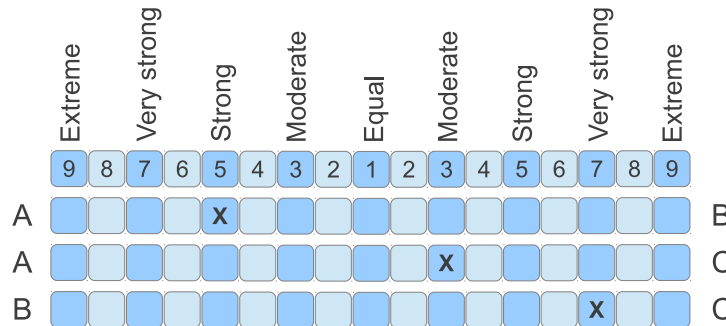


Figure 2: Fundamental scale measurement tool for the pairwise comparisons of Criteria A, B, and C.

To complete the matrix, we read across the rows. Since “Criterion A is *strong* in importance compared to Criterion B,” we add the entry 5 into the matrix at row 1 (A), column 2 (B):

$$\begin{array}{c}
 \text{Criterion A} \quad \text{Criterion B} \quad \text{Criterion C} \\
 \text{Criterion A} \begin{pmatrix} 1 & & \\ & 5 & \\ & & 1 \end{pmatrix} \\
 \text{Criterion B} \\
 \text{Criterion C}
 \end{array}$$

Since “Criterion C is *moderate* in importance compared to Criterion A,” we add the entry 3 into the matrix at row 3 (C), column 1 (A):

$$\begin{array}{c}
 \text{Criterion A} \quad \text{Criterion B} \quad \text{Criterion C} \\
 \text{Criterion A} \begin{pmatrix} 1 & & \\ & 5 & \\ 3 & & 1 \end{pmatrix} \\
 \text{Criterion B} \\
 \text{Criterion C}
 \end{array}$$

Since “Criterion C is *very strong* in importance compared to Criterion B,” we add the entry 7 into the matrix at row 3 (C), column 2 (B):

$$\begin{array}{c}
 \text{Criterion A} \quad \text{Criterion B} \quad \text{Criterion C} \\
 \text{Criterion A} \begin{pmatrix} 1 & & \\ & 5 & \\ 3 & & 1 \end{pmatrix} \\
 \text{Criterion B} \\
 \text{Criterion C}
 \end{array}$$

Since the reverse comparisons have intensity values of $1/i$, we can complete the comparison matrix by filling it in with the reciprocal values:

$$\begin{array}{c}
 \text{Criterion A} \quad \text{Criterion B} \quad \text{Criterion C} \\
 \text{Criterion A} \begin{pmatrix} 1 & & \\ & 5 & 1/3 \\ 1/5 & & 1/7 \end{pmatrix} \\
 \text{Criterion B} \\
 \text{Criterion C}
 \end{array}$$

To derive the criteria priority weights we solve for the normalized right eigenvector of the

comparison matrix (see Appendix for detail).

Table 2: Criteria priority weights

Criterion (j)	Priority Weight (p_j)
A	0.2790
B	0.0719
C	0.6491
Total	1.0000
(In)consistency ratio	0.0624

2.3. Consistency of judgments

While multiple pairwise comparisons improve accuracy, decisions makers’ judgments still cannot be measured with absolute certainty and therefore can be inconsistent with their valuations. For example, if a decision maker prefers A to B, and then B to C, we can expect A to be preferred to C. However, inconsistency arises when the decision maker prefers C to A. Inconsistency is measured by the consistency ratio (CR) and it is generally acceptable if $CR < 0.10$. When CR becomes relatively large (> 0.10), then its reasons should be explored. Inconsistencies can result from unintentional errors, lack of concentration during the comparison process, or even misunderstandings. An advantage of the AHP is that it allows us to identify, explore, and correct these inconsistencies.

3. Application

Consider a local health department (LHD) that is committed to becoming a high performance, learning organization through robust strategic effectiveness, performance management, and quality and equity improvements. They have adopted the AHP to improve their decision making, priority setting, and resource allocation processes. The planning unit has developed a priority setting tool (Figure 3) to assist them in prioritizing health programs so that they can align with the agency’s strategic directions.

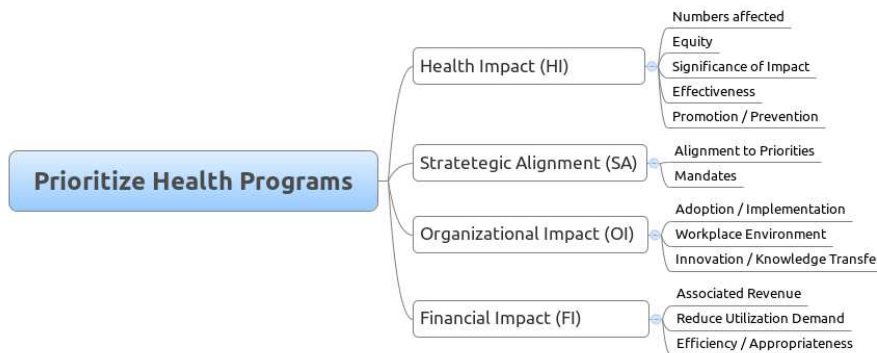


Figure 3: AHP for Prioritizing Health Programs (Adapted from [7])

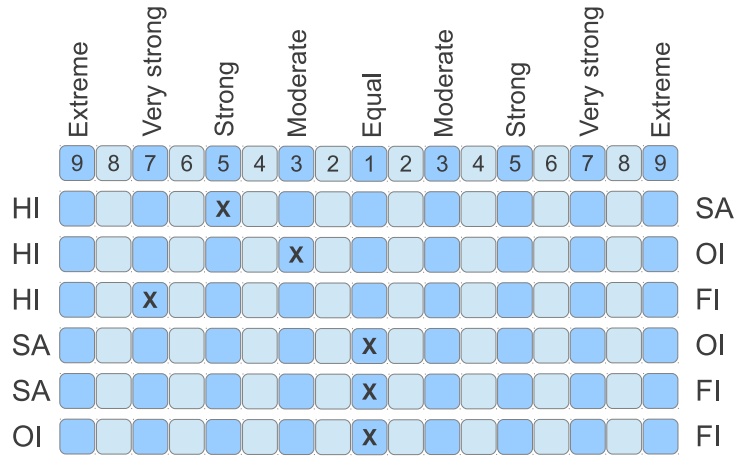


Figure 4: Dr. Juan Nieve’s pairwise comparisons of top-level health prioritization criteria using the fundamental scale

Now, the executive leadership team is charged with assessing the top level criteria and deriving criteria weights. The top level criteria are the following:

- Health Impact (HI)
- Strategic Alignment (SA)
- Organizational Impact (OI)
- Financial Impact (FI)

The executive team completed a session where they defined and approved the criteria. Each executive was then provided with a criteria scoring tool which they were instructed to use their experience, expert judgment, and understanding of existing evidence to score the criteria using pairwise comparisons.

Dr. Juan Nieve is a public health officer and he believes that Health Impact is the most important criterion. He has no strong feelings about the other criteria and considers them equal; his scores are displayed in Figure 4. From his intensity scores we can construct his comparison matrix:

$$\begin{matrix}
 & \text{HA} & \text{SA} & \text{OI} & \text{FI} \\
 \text{Health Impact (HI)} & \left(\begin{matrix} 1 & 5 & 3 & 7 \end{matrix} \right) \\
 \text{Strategic Alignment (SA)} & \left(\begin{matrix} 1/5 & 1 & 1 & 1 \end{matrix} \right) \\
 \text{Organizational Impact (OI)} & \left(\begin{matrix} 1/3 & 1 & 1 & 1 \end{matrix} \right) \\
 \text{Financial Impact (FI)} & \left(\begin{matrix} 1/7 & 1 & 1 & 1 \end{matrix} \right)
 \end{matrix}$$

In contrast, Mr. Donald Trumpini is a finance officer and he stays up-to-date on the anticipated fiscal impacts of health care reform. He thinks programs that generate revenue should be weighted moderately higher in order to ensure a financially sustainable health system. Here is

his comparison matrix (scoring tool sheet not shown):

	HA	SA	OI	FI
Health Impact (HI)	(1	1	1/3
Strategic Alignment (SA)		1	1	1/3
Organizational Impact (OI)		1	1	1/5
Financial Impact (FI)		3	3	5
)			

The derived criteria weights for each comparison matrix are displayed in Table 3. Notice that Dr. Nieve and Mr. Trumpini value each criteria differently; this is appropriate and expected. The AHP allows us to measure these valuations, making the priority-setting process transparent.

With group decision making we must aggregate individual-level data to get overall priority weights. We have two approaches:

1. aggregate individual priority (AIP) weights that were derived from individual judgment matrices; or
2. aggregate individual judgment (AIJ) matrices first into one matrix, and then derive the overall priority weights.

Both approaches are covered next.

3.1. Aggregating individual priority weights

Aggregating individual priority (AIP) weights is useful when we want to honor, recognize, or study individual valuations or between-person variability. This may also be useful if we want to identify differences that should be discussed, clarified, or resolved. Unfortunately, individuals may not respond honestly or participate fully if they do not want their valuations to be scrutinized.

The criteria AIP weights (p_i) were calculated using the geometric mean (Equation 1),

$$p_j = \sqrt[n]{\prod_{i=1}^n p_{ij}}, \quad (1)$$

where n is the number of decision makers. Then the criteria priority weights were normalized

Table 3: Criteria priority weights, including normalized geometric mean

Criterion (j)	Nieve	Trumpini	Aggregated
	Weights	Weights	Weights
	p_{1j}	p_{2j}	p'_j
Health Impact	0.6125	0.1581	0.3676
Strategic Alignment	0.1253	0.1581	0.1663
Organizational Impact	0.1454	0.1401	0.1686
Financial Impact	0.1167	0.5437	0.2976
Total	1.0000	1.0000	1.0000
(In)consistency ratio	0.0259	0.0123	0.0178

(Equation 2):

$$p'_j = p_j / \sum_j p_j, \quad (2)$$

Although an arithmetic mean can be calculated, the geometric mean is more appropriate because these weights have ratio properties, meaning that ratio comparisons are valid. For completeness, the arithmetic mean formula is provided (Equation 3):

$$p_j = \frac{\sum_{i=1}^n p_{ij}}{n}, \quad (3)$$

3.2. Aggregating individual judgment weights

An alternative approach is aggregating individual judgment (AIJ) weights. In the AIJ method we aggregate the comparison matrices first and then derive the criteria priority weights. Because comparison matrices contain ratio measures, we must take the geometric mean and not the arithmetic mean:

$$J = \sqrt[n]{\prod_{i=1}^n J_i}, \quad (4)$$

where J is the aggregated comparison matrix of geometric means calculated from the individual (i) comparison matrices using Equation 4. Here is the aggregated comparison matrix J :

	HA	SA	OI	FI
Health Impact (HI)	1.00	2.24	1.73	1.53
Strategic Alignment (SA)	0.45	1.00	1.00	0.58
Organizational Impact (OI)	0.58	1.00	1.00	0.45
Financial Impact (FI)	0.65	1.73	2.24	1.00

Like before, we can derive the criteria priority weights from the aggregated comparison matrix J above (see Table 4, column 3). Table 4 also displays the results comparing the AIJ and AIP methods. Geometric mean results are almost identical meaning either method is appropriate. The arithmetic mean is reasonable with the caveat that ratio comparisons would not be appropriate. The AIJ method is selected over the AIP method when we want to have a single comparison matrix to represent the collective judgment of a group of decision makers, and we are not interested in evaluating individual priority weights.

Table 4: Comparison of aggregated priority weights using the AIP and AIJ methods

Criterion	AIP		AIJ
	Arithmetic Mean	Geometric Mean	Geometric Mean
Health Impact	0.3853	0.3676	0.3678
Strategic Alignment	0.1417	0.1663	0.1661
Organizational Impact	0.1428	0.1686	0.1683
Financial Impact	0.3302	0.2976	0.2978
Total	1.0000	1.0000	1.0000

4. Discussion

In this brief tutorial we showed how to use the analytic hierarchy process (AHP) for deriving criteria priority weights. Using AHP pairwise comparisons we measure whether one criteria has dominance over another, and if yes, what is the relative intensity of this dominance. The relative intensity is based on an ordinal scale with ratio properties called the fundamental scale. The relative intensity represents our comparative interpretation with respect to importance, likelihood, preference, impact, or other factor of interest. AHP allows us to combine our interpretations of evidence (from data, testimony, etc.) with qualitative attributes such as preference or other “intangibles.” This fact alone makes AHP incredibly powerful and practical. At worst, AHP improves our decision making.

Although not shown in this tutorial, the criteria priority weights are further used to assess alternatives. In spite of its simplicity, there is usually resistance to applying more rigor to decision making. For important decisions that matter we must overcome our fears, biases, and methodologic limitations. While there is no foolproof approach, many of us realize that we are not proficient at employing rigorous, evidence-based methods for decision-making. While we may be perceived to make “good decisions” some or most of the time, we have no clear way of *demonstrating* that our team made the “best decision” every time—even when it matters.

The guts of the AHP is the selection and weighting of criteria, and the application of these weighted criteria to the alternatives we are considering. The best decisions are group decisions using the most knowledgeable and impacted stakeholders to develop the criteria and to score the alternatives. The criteria can be based on interpretation of quantitative (e.g., rate ratios) or qualitative (alignment to organizational strategy) data. The ability to measure and incorporate qualitative attributes (“intangibles”) is very powerful! Key stakeholders’ strong preferences can also be incorporated explicitly in this method. Finally, decisions can be explained, rationalized, and reviewed to assess which factors had the biggest influence on the final decision or ranking (also called sensitivity analysis). While the AHP does require some matrix algebra, this can easily be handled using a freely available software (see Appendix).

To conclude, we would argue that MCDMs like the AHP are just “systematic common sense” applied to important decisions, and that sense of control and confidence will increase—not decrease—with use. Applying MCDMs can transform an organization, even if it is only applied to simple (but important) decisions. MCDMs will create a systematic approach to decision making, priority setting, and resource allocation. At worst, the development of clear goals, selection of criteria, and evaluation of alternatives will help improve the decision making process.

Appendix A. Deriving criteria weights using R

R is an open source, multi-platform program for statistical computing and graphics. It can be downloaded from www.r-project.org. The R Project site contains numerous free tutorials for learning R basics. The following assumes minimal R proficiency.

Now, how do we use R to derive the criteria priority weights from Dr. Juan Nieve's comparison matrix?

	HA	SA	OI	FI
Health Impact (HI)	1	5	3	7
Strategic Alignment (SA)	1/5	1	1	1
Organizational Impact (OI)	1/3	1	1	1
Financial Impact (FI)	1/7	1	1	1

Very simple, here is R code entered at the R command prompt that accomplishes this:

```
> x = c(1, 5, 3, 7, 1/5, 1, 1, 1,
+ 1/3, 1, 1, 1, 1/7, 1, 1, 1)
> xm = matrix(x, nr = 4, nc = 4, byrow = TRUE)
> xm
      [,1] [,2] [,3] [,4]
[1,] 1.000000 5 3 7
[2,] 0.200000 1 1 1
[3,] 0.333333 1 1 1
[4,] 0.142857 1 1 1
> eigen.xm = eigen(xm)
> prop.table(as.numeric(eigen.xm$vectors[,1]))
[1] 0.6125463 0.1253280 0.1453999 0.1167258
```

Above we used the `eigen` function to derive the right eigenvector; we used the `as.numeric` function to simplify complex number notation; and we used the `prop.table` function to normalize the eigenvector.

Continuing previous calculation, here is how we calculate the consistency ratio (CR):

```
> nn = 4 #number of criteria
> rand.ci = 0.89 #random CI from Table A.5
> lambda.max = eigen.xm$values[1]
> consist.index = (lambda.max - nn)/(nn - 1)
> consist.ratio = consist.index/rand.ci
> consist.ratio
[1] 0.02585234+0i
```

The consistency ratio is just the consistency index divided by the random consistency index (Table A.5 on the next page). The consistency index is given by

$$CI = \frac{\lambda_{max} - n}{n - 1}, \quad (A.1)$$

where n are number of criteria, and λ_{max} is the maximum eigenvalue which was provided by the `eigen` function in R. Details of consistency ratio calculations are provided elsewhere [2]

Table A.5: Random consistency index (*RI*) for *n* criteria

<i>n</i>	<i>RI</i>	<i>n</i>	<i>RI</i>
1	0.00	9	1.45
2	0.00	10	1.49
3	0.52	11	1.51
4	0.89	12	1.54
5	1.11	13	1.56
6	1.25	14	1.57
7	1.35	15	1.58
8	1.40	–	

References

- [1] Saaty TL. Decision Making for Leaders: The Analytic Hierarchy Process for Decisions in a Complex World, New Edition 2001 (Analytic Hierarchy Process Series, Vol. 2). 3rd ed. RWS Publications; 2008. Available from: <http://amazon.com/o/ASIN/096203178X/>.
- [2] Saaty TL. Fundamentals of Decision Making and Priority Theory With the Analytic Hierarchy Process (Analytic Hierarchy Process Series, Vol. 6). 2nd ed. RWS Publications; 2006. Available from: <http://amazon.com/o/ASIN/0962031763/>.
- [3] Forman EH, Selly MA. Decision by Objectives. 1st ed. World Scientific Pub Co Inc; 2002. Available from: <http://amazon.com/o/ASIN/9810241429/>.
- [4] Liberatore MJ, Nydick RL. The analytic hierarchy process in medical and health care decision making: A literature review. European Journal of Operational Research. 2008;189:194. Available from: <http://dx.doi.org/10.1016/j.ejor.2007.05.001>.
- [5] Dolan JG. Multi-criteria clinical decision support: A primer on the use of multiple criteria decision making methods to promote evidence-based, patient-centered healthcare. Patient. 2010;3(4):229–248. Available from: <http://dx.doi.org/10.2165/11539470-000000000-00000>.
- [6] Baltussen R, Niessen L. Priority setting of health interventions: the need for multi-criteria decision analysis. Cost Eff Resour Alloc. 2006;4:14. Available from: <http://dx.doi.org/10.1186/1478-7547-4-14>.
- [7] Mitton C, Dionne F, Damji R, Campbell D, Bryan S. Difficult decisions in times of constraint: Criteria based resource allocation in the Vancouver Coastal Health Authority. BMC Health Serv Res. 2011;11:169. Available from: <http://dx.doi.org/10.1186/1472-6963-11-169>.