J Korean Med Sci 2009; 24 (Suppl 2): S242-6
ISSN 1011-8934
DOI: 10.3346/jkms.2009.24.S2.S242

Consensus Making Process in the Korean Academy of Medical Sciences Guideline for Physical Impairment: Evaluation as a Social Process

The steering committee of the Korean Academy of Medical Sciences Guideline for Physical Impairment was fully aware of the social processes of disability evaluation from the beginning and thus, developed a series of strategies to examine and incorporate social property of the evaluation into the evaluation guide. Although those strategies could not be implemented to full extent because of lack of budget and time, we believe it worthwhile to share those in this paper as an example of general framework for developing disability evaluation. A series of strategies will be introduced and discussed that views the evaluation process as social per se, and propose a scheme that is designed to obtain growing legitimacy starting from core experts to expanded experts to general public. Also preliminary analyses on surveys of public attitude and experts' opinion with regard to the relative importance of each possible disability revealed the following three facts: 1) Public had difficulty weighing relative importance of many impairments. 2) Regarding some impairments including complex regional pain syndrome many doctors had varied opinions. 3) Public attitude did not always consistent with doctor's opinion. All these findings strongly suggest the need for developing strategies to draw consensus for legitimate and effective evaluation.

Key Words : *Disability Evaluation; Social Process; Consensus*

Yoosik Youm[1], Mi-Jin Lee[2], and Sun-Chul Hwang[3]

Department of Sociology[1], Yonsei University, Seoul; Graduate Program in Medical Law and Ethics[2], Yonsei University, Seoul; Department of Neurosurgery[3], College of Medicine, Soonchunhyang University, Bucheon, Korea

Received : 6 April 2009
Accepted : 30 April 2009

Address for correspondence
Yoosik Youm, M.D.
Department of Sociology, Yonsei University, 262 Seongsan-ro, Seodaemun-gu, Seoul 120-749, Korea
Tel : +82.2-2123-2431, Fax : +82.2-2123-2420
E-mail : yoosik@yonsei.ac.kr

*This work has been supported by the department of sociology, Yonsei University Institute of "Social Subsumption and Exclusion", a Brain Korea 21 program, Korea.

INTRODUCTION

Republic of Korea does not yet have a universal criterion for disability evaluation. In order to develop a single reasonable and consistent criterion that could be widely accepted across the nation, eleven committees for each medical specialty were recruited under the supervision of Korean Academy of Medical Sciences (KAMS): cardiopulmonary, digestive, endocrine, extremities, genitourinary, nervous system, pediatric development, psychiatric, skin & appearance, special sense, and spine. Each committee consisted of about ten specialists. Since both disability and developing disability evaluation are social processes in the sense that they are formed and changed by social norms and values, the steering committee developed a series of strategies from the very beginning that would help the evaluation criterion to be accepted as legitimate by majority of society. Unfortunately due to the lack of time and budget, however, the strategies could not be fully exercised. We plan to apply our strategies more fully in the next year with more money and time available. This paper introduces this series of strategies in order to share Korean experience.

Disability and disability evaluation as social processes

As widely acknowledged, in order for a certain impairment to be recognized as a disability, it must constrain a person in a way that substantially limits the activity, especially in relation to employment or education. In other words, a person with same impairment could be disabled or non-disabled, depending on time and space where he or she lives in. For example, substantial leg impairment could be fatal in a pre-modern society, but it is now much less critical in modern society where a variety of transportation tools are available. Also under the strong emphasis of the social value on facial appearance, some big permanent scars on face could lead to disability. Disability is the combined product of impairment and diverse social environments, including technology, social norms, and cultural values. Because of this social process of disability, we need to consider and even incorporate social norms and values of the current society when we evaluate disability.

Furthermore, social environment where impairment is embedded is not a single constant entity. Social values are constantly changing and also vary across societies or even groups

within one society. Thus, social surroundings of impairment must be recognized as an entity with fractured and diverse aspects rather than one simple universe. Considering this social characteristic of disability, the best strategy would be to incorporate diverse and even conflicting values and pressures from the beginning of the evaluation rather than treat them as a separate entity from the evaluation. The KAMS guide tried to adopt a series of strategies based on this viewpoint. Before we introduce those strategies, however, let me continue the debate on the social process of the evaluation procedure itself.

Disability evaluation in practice consists of a series of evaluations and re-evaluations. First, it should start from a small number of specialists. KAMS started with about ten people in each specialty. Let's call this group of medical specialist as core experts. But in many cases, this first step alone is not enough for successful evaluation procedure for two reasons. For any specialty field, there usually exist some vague grey areas that call for judgment call. Thus, the best approach might be to bring those unresolved issues to a bigger number of specialists: expanded experts. Also, expanded experts are necessary to obtain stronger legitimacy. Even if core experts could produce objective and scientific criteria to the full extent, as we discussed above, criteria themselves are social product, and thus, they need to be shared as many social members as possible to gain strong legitimacy and widely adopted by many experts later.

Second step would be, as we discussed above, to form a team of expanded experts. Depending on the magnitude of unresolved grey areas among core experts, the size of expanded experts should be determined. However, more critical issue here is how to recruit the members of expanded experts. The goal of forming core experts is not to obtain a representative sample of experts, but to recruit top experts to build up evaluation criterion. However, the reason we need expanded experts is different. We need a representative sample of experts as much as possible, because unresolved issues among core experts are not of technique but of social value or customs in many cases. Especially, considering legitimacy issue that could be critical once the evaluation criteria announced, the representativeness of the expanded experts is essential. Usually, there is no simple universal way of obtaining representative sample of medical experts and we will discuss this more later. The last step could be to probe and incorporate public opinions in various active ways. Since this last step usually is not considered as a part of evaluation process, we will discuss this later in detail.

MATERIALS AND METHODS

Consensus among core experts

The important issue regarding core experts is not the recruitment. Usually, it could be done by recommendation of top experts or appointment by an official national medical association. Once recruited, however, the consensus is the key to the successful evaluation procedure. Because of the nature of the evaluation, if possible, consensus instead of majority rule would be much more desirable.

Reaching a consensus poses two difficult challenges. First, it usually requires a moderator. However, it is hard to find a perfect moderator who can balance between every different or even conflicting views. Especially, expert committees like this also need an expert of the same field as a moderator, which makes him or her uneasy to act in a neutral way. Second, reaching a consensus needs more time and efforts than majority voting. Although consensus usually gives more satisfaction among participants, we sometimes have to rely on voting due to the lack of time.

If time permits, small number of experts like ten or so could reach a consensus. But, we do not usually have enough time for consensus regarding every issue. Here we suggest a way as an example to reach a consensus with the minimum activity of moderator in a relatively short time. It is called Multiperson Decision Making Problem (MDMP) model (1-5). This consists of the following six steps. 1) Opinion Expression: each expert expresses his or her own preference on the issue. Based on each opinion, the magnitude of consensus is measured. There exist several alternatives for this measure including variance that was used in Economics (5-8). Whatever measure we use, if it exceeds some pre-determined criterion, we can claim that we reach a consensus and all the following procedures are unnecessary. 2) Proximity measure: If the amount of consensus is not satisfactory, we measure proximity of each opinion from the whole opinion. This in principle measures how far each opinion is away from the whole committee. 3) Feedback: based on proximity measure, people with proximity score that are furthest from the collective score are recommended to change their opinion so that their opinions become closer to the collective one. 4) Changing the opinions: based on the feedback, some people change their opinion. 5) Now, new consensus score is calculated, and it is compared to the criterion. 6) If it exceeds the criterion, satisfactory level of consensus is reached and no more action is needed. If not, we will repeat from step 2) again until we will have satisfactory amount of consensus.

Consensus among expanded experts

As we discussed above, the critical issue here is to obtain a representative sample of experts in specific medical field to resolve legitimacy issue. The most widely used method was a snow-ball sampling based on core experts as seeds. We ask core experts to enumerate other experts in the same field. Once we contact those additional experts, we ask the same favor: to enumerate other experts. We repeat this snowballing until we hit the desired number of expanded experts. Although this

recruitment strategy is convenient and natural way to collect a small number of target people, it usually ends up with strongly biased sample. For example, if we start with male experts, the resultant sample would be male-biased.

Another popular strategy is to use a membership directory. Usually, medical experts must be registered and Korea is not an exception. If we can contact experts in a random way from the directory and do a survey with them, it would be a perfect representative sample. However, medical experts, especially in Korea, are busiest men due to tight schedule of caring patients. Thus, the response rate for survey is usually extremely low, and as result, the resultant final sample is also strongly biased.

One possible alternative which we plan to implement is to adopt a respondent driven sampling (RDS) method that adjusts biasedness introduced during snowball sampling (9-11). This method is handy in the sense that it is really similar to traditional snowball sampling except we collect some network information of respondents so that we can adjust biasedness of each respondent later. This method is widely adopted since its debut in the 1990s, especially among the studies that targeted for hidden small population such as jazz singers, drug users, men who have sex with men (MSM), etc.

We conducted a new survey of 400 experts (doctors) on their opinion on physical impairment rate for 16 different impairments via faxes and obtained 195 completed cases (48.8% response rate) form September 2008 to January 2009. The Table 1 summarized the result. Although in general experts showed very consistent opinions, some impairments such as 'gastrointestinal fistulas' and 'complex regional pain syndrome' exposed substantial variance (or standard errors) among expert opinions. For example, response rate for 'complex regional pain syndrome' ranges from 0% to 70% and also inter-quartile range is 30%, which is relatively large. This finding strongly suggests the necessity of developing strategies for successful consensus making among experts since simple mean could be ineffective and even illegitimate when there exist wide variations of opinions.

Public opinion

We need to examine opinion of the public beyond expert group for two reasons. First, we need to weigh impairments across field. For example, we need to decide relative importance of eyes vs. legs. In this case, it is not recommendable to ask opinions of experts only. We need to examine public opinion when our decision requires more than one medical special field. Second, we also need to examine public opinion to gain stronger legitimacy. Even if the evaluation criteria and procedure are scientific and objective, the legitimacy of the criteria could be weak unless they are accepted by the public. Especially, considering sharp conflicts between interest groups with regard to disability evaluation, it could be critical to secure public legitimacy.

There could be two alternatives to probe public opinion. The first one is to conduct a survey of a representative sample of the public. The second alternative is to combine MDEMP model with a traditional survey. The first approach is most widely used. Also if we obtain a majority preference from this survey, it is the best strategy available. But what if we confirm the existence of sharp division, for example 52% vs. 48%? It would be hard to draw conclusions from this type of neck to neck competing and sometimes even conflicting opinions. This is actually what we found from a public opinion survey. This survey is conducted in February 2009 on a 840 representative sample of Korean adults through mobile phone text messaging. A survey research company named MBIZONE keeps the list of more than one million people who registered and agreed to be respondents for the surveys conducted by the company in the exchange of small amount of money. MBIZONE selected 840 representative adults out of registered one million people for this study. Table 2 summarized

Table 1. Summary of experts' opinion on physical impairment rates

Physical impairment	Cases	Mean rate	Standard error	Confidence interval (95%)
Persistent vegetative state	178	99.60	0.11	99.38-99.81
Complete quadriplegia (without self-respiration)	178	99.36	0.16	99.05-99.67
Complete quadriplegia (with self-respiration)	178	95.74	0.23	95.28-96.20
Complete loss of vision	178	84.07	0.39	83.31-84.83
Complete paraplegia	178	67.83	0.64	66.56-69.10
Amputation of the arms	178	75.31	0.72	73.89-76.73
Amputation of the legs	178	64.83	0.57	63.71-65.95
Complete loss of hearing	178	45.84	0.83	44.21-47.47
Complete loss of olfaction	178	7.25	0.50	6.27-8.24
Complete loss of taste	178	6.66	0.42	5.84-7.48
Micturation	178	35.85	0.55	34.78-36.93
Defecation	178	32.83	0.50	31.84-33.82
Sexual function (3rd to 4th decades)	178	20.49	0.51	19.48-21.50
Gastrointestinal fistulas (maximum)	170	52.85	1.06	50.75-54.95
External appearence	177	46.57	0.67	45.24-47.89
Complex regional pain syndrome (maximum)	165	44.93	1.38	42.20-47.67

Table 2. Summary of public opinion survey

	Cases	Mean*	Standard error	Confidence interval (95%)
Amputation of arms (vs. amputation of legs)	825.00	2.15	0.03	2.09-2.20
Amputation of arms (vs. complete loss of vision)	833.00	2.47	0.03	2.41-2.52
Amputation of one leg (vs. ankylosing spondylitis)	828.00	2.66	0.02	2.61-2.70
Loss of the thumb and index fingers in one hand (vs. dysfiguration from facial burn)	833.00	2.79	0.02	2.75-2.83
Complete loss of hearing (vs. amputation of one leg)	827.00	2.19	0.03	2.13-2.26
Right hemiplegia with useless arm and leg (vs. amputation of arms)	824.00	2.41	0.03	2.35-2.46
Amputation of legs (vs. complege paraplegia)	830.00	2.24	0.03	2.19-2.30
Dysfuguration from facial burn (vs. loss of sexual function)	823.00	1.33	0.02	1.29-1.38
Loss of the thumb and index fingers in one hand (vs. partial resectin of the stomach)	805.00	1.38	0.03	1.33-1.43
Complete paraplegia (vs. mental retardation)	827.00	1.86	0.03	1.80-1.92
Mental retardation (vs. complete loss of hearing)	821.00	1.88	0.03	1.81-1.94
Complete quadriplegia (vs. amputation of arms)	829.00	1.38	0.02	1.34-1.43
Amputation of one arm (vs. multi-segmental fusion of the lumbar spine)	816.00	1.94	0.03	1.88-2.00
Complete loss of hearing (vs. renal failure)	812.00	2.18	0.03	2.12-2.25
Complete quadriplegia (vs. persistent vegetative state)	828.00	2.12	0.03	2.07-2.18

*, 1 means the former one is more critical, 3 means the latter is more critical, and 2 means two are almost identical.

the result. The table summarized a series of questions to compare two impairments. For example, the first question reads, "Who is more disabled in Korea, 'person who lost both arms' or 'person who lost both legs'?" We coded 1 if a respondent chose the first one and coded 3 if he or she chose the latter one. We coded 2 if a respondent said, 'almost identical'.

We obtained two important implications from Table 2. First, we can find some impairments had the mean value around 2. The closest one is 1.94 for the comparison between 'amputation of one arm' and 'multi-segmental fusion of the lumbar spine'. Public survey produced little information about the relative importance of this type of impairments. Second, public opinion does not always consistent with experts' view. For example, Table 1 clearly showed that Korean experts believed amputation of arms was more critical than amputation of legs (95% confidence intervals are '73.89-76.73' vs. '63.71-65.95'). Additional table (not shown) revealed that only 4.5% of experts believed that amputation of legs was more critical while 70% believed that amputation of arms was more critical. However the order is reversed in public opinion. In Table 2, the mean value of 'amputation of arms' vs. 'amputation of legs' was 2.15, which means that public believed that amputation of legs was more critical. According to additional analysis (not shown), the proportion of people who believed that amputation of legs was more critical was 44% while 30% of the respondent believed that amputation of arms was more serious. Both findings robustly implicated the need of developing efficient and legitimate procedure of consensus formation.

The second alternative to investigate public opinion is to integrate MDMP model with public survey. Because MDMP model includes feedback and changing opinions, it is not possible to apply this to huge number of people in person. Instead, it is possible to apply this to about one thousand people via mobile phone. Once we select a representative sample from MBIZONE, we can give feedbacks and obtain new changing opinions through its system. Considering the fact that mobile phone survey only takes about one day for about 1,000 people, this is a plausible strategy in the future.

DISCUSSION

The KAMS realized that disability evaluation itself is a social product, therefore, tried to incorporate social property of evaluation into evaluation process from the beginning. Due to the lack of budget and time, we could not implement all the strategies we planned. We just finished the first draft of the evaluation guide from core experts. We plan to apply strategy 1, 2, or 3 as time permits, since we now enter into the second year period of our project. Evaluation guide must change across societies and times but we believe that our set of strategies could be an example of a general framework where opinions of experts and public are systematically examined and thus legitimacy could strongly be held.

REFERENCES

1. Bordogna G, Fedrizzi M, Pasi G. *A linguistic modeling of consensus in group decision making based on OWA operators. IEEE Trans Syst Man Cybern A Syst Hum 1997; 27: 126-32.*
2. Bryson N. *Group decision-making and the analytic hierarchy process: Exploring the consensus-relevant information content. Comput Oper Res 1996; 23: 27-35.*
3. Carlsson C, Ehrenberg D, Eklund P, Fedrizzi M, Gustafsson P, Lindholm P, Merkuryeva G, Riissanen T, Ventre AG. *Consensus in distributed soft environments. Eur J Oper Res 1992; 61: 165-85.*
4. Chiclana F, Herrera F, Herrera-Viedma E. *Integrating three represen-*

tation models in fuzzy multipurpose decision making based on fuzzy preference relations. *Fuzzy Sets Syst 1998; 97: 33-48*.

5. Herrera-Viedma E, Herrera F, Chiclana F. *A consensus model for multiperson decision making with different preference structures*. *IEEE Trans Syst Man Cybern 2002; 32: 394-402*.
6. Alston RM, Kearl JR, Vaughan MB. *Is there a consensus among economists in the 1990s? American Economic Review, Papers and Proceedings 1992; 82: 203-9*.
7. Fuller D, Geide-Stevenson D. *Consensus among economists: revisited*. *J Econ Educ 2003; 369-87*.
8. Kearl JR, Pope CL, Whiting GC, Wimmer LT. *A confusion of economists*. *American Economic Review, Papers and Proceedings 1979; 69: 28-37*.
9. Heckathorn DD. *Respondent driven sampling: a new approach to the study of hidden populations*. *Social Problems 1997; 44: 174-99*.
10. Heckathorn DD. *Respondent driven sampling II: deriving statistically valid population estimates from chain-referral samples of hidden populations*. *Social Problems 2002; 49: 11-34*.
11. Heckathorn DD, Semaan S, Robert S, Broadhead RS, Hughes J. *Extensions of respondent-driven sampling: a new approach to the study of injection drug users aged 18-25*. *AIDS Behav 2002; 6: 55-67*.