bariatric surgery and on the identification of additional target areas. Our capacity building programme has trained over 250 individuals within and without Amil, including policy developers within regional and national health ministries.

Conclusion An IHTA programme within a BHCO may offer a capacity building model useful for national development.

| P270 | DEVELOPMENT OF QUALITY INDICATORS OF LOW RISK LABOR PROVIDED PRIMARILY BY MIDWIVES: BASED ON CLINICAL PRACTICE GUIDELINES WITH A MODIFIED DELPHI METHOD |

[1,2]K Ueda, [1]S Ohtera, [1,3]K Misato, [1]T Nakayama. [1]Department of Health Informatics, Kyoto University School of Public Health, Kyoto, Japan; [2]Morinomiya University of Medical Sciences School of Nursing, Osaka, Japan; [3]Kansai University of Health Sciences School of Nursing, Osaka, Japan

Background In Japan some labour units have started services for low risk labour provided primarily by midwives with the assistance of obstetricians. However, real practices varied with facilities and quality of care has not been sufficiently assured.

Objectives We aimed to develop quality indicators of low-risk labour provided primarily by midwives in Japan.

Methods We systematically searched databases to select related indicators and clinical practice guidelines. Using a modified Delphi method, a multidisciplinary panel including two women who experienced labour and were not health professionals was assembled and a meeting was held. After the panel members rated the candidates individually, consensus was attained for each indicator.

Results An electronic search retrieved 104 guidelines (2043 key recommendations) from which 32 guidelines (166 key recommendations) were selected. Some recommendations were excluded due to difficulties in measuring and others due to evidence against following certain guidelines. Additionally, 31 existing quality indicators were identified. The panel discussed 25 candidate indicators: 18 indicators were adopted, six indicators were not adopted, one indicator was modified. Four indicators were added later following discussion. Consequently, 23 quality indicators of care were chosen.

Discussion A pilot test of these indicators at multiple facilities should contribute to future development of guideline-based quality indicators.

Implications for Guideline Developers/Users Guideline developers should be more conscious in making recommendations that are potentially used as quality indicators.

| P271 | COMPUTERIZED GUIDELINES ANALYSIS: 5-YEARS EXPERIENCE AND USER FEEDBACK |

[1]G Georg, [2]J Siddle, [2]M Truran, [2]M Cavazza. [1]Haute Autorité de Santé, Saint-Denis, France; [2]Teesside University, Middlesbrough, UK

Background In 2007, HAS (Haute Autorité de Santé, France) started experimenting the use of computerised guidelines analysis with the G-DEE software as part of the guidelines production process. G-DEE uses shallow natural language processing techniques to recognise the expression of recommendations in a guidelines' text, and structures the document accordingly using on-screen text highlighting and colouring. Project managers were offered, on an entirely voluntary basis, the possibility of using this tool during the final phases of guidelines production to i) analyse overall guidelines structure and ii) review the formulation of individual recommendations. To date, this experiment has concerned 38% of guidelines that have been officially released.

Objectives Improving the performance and usability of the deployed G-DEE system.

Methods We have collected feedback from guidelines' project managers who have identified clear benefits in terms of readability of individual recommendations, overall text balance, fine-tuning and testing of alternative wording. Main limitations identified were processing time, which required additional planning and the absence of a tuning facility that would allow different levels of grammar coverage.

Results Notable enhancements include faster document processing (x10) and incremental document analysis allowing direct text input in the system.

Discussion These will make its use possible throughout the production process, e.g. directly as part of consensus meetings once the first draft is available. In addition, the possibility to define additional grammar files for text processing supports the processing of new document types (shorter guidelines, summaries) which may adopt different writing styles.

| P273 | EFFECTS OF KOREAN-AGREE SCORING GUIDE ON IMPROVING THE RELIABILITY OF THE SCORES |

[1]M Oh, [1,2,5]H Jo,[3,4,5]Y Lee. [1]Department of Preventive Medicine, Kangwon National University Hospital, Chuncheon, Korea; [2]Department of Health Management and Policy, Kangwon National University, Chuncheon, Korea; [3]Department of Laboratory Medicine and Genetics, Soonchunhyang University Hospital, Bucheon, Korea; [4]Department of Laboratory Medicine and Genetics, Soonchunhyang University, Bucheon, Korea; [5]The Executive Committee for Clinical Practice Guideline, The Korean Academy, Seoul, Korea

Background The Korean translated AGREE II instrument has been distrubuted to Korean medical societies in 2011. But inter-rater disagreement issues still existed. The Korean AGREE II scoring guide was therefore developed to reduce inter-rater differences.

Objectives To examine effects of the Korean AGREE II scoring guide on reduction of the inter-rater differences.

Methods Appraisers were randomly assigned to group A or B. The Korean AGREE II instument was provided to all participants, but the Korea AGREE II scoring guide was offered to only group A. Total of 14 appraisers participated and each gudeline was assessed by 8 appraisers. To evaluate the reliability of the Korean AGREE II scoring guide, correlation of scores among appraisers and domain-specific intra-class correlation (ICC) were calculated.

Results Most scores of two groups were comparable. And scoring guide user group showed higher reliability at all guidelines. It showed higher correlation among appraisers and higher ICC values at almost all domains (from $0.272{\sim}0.792$ to $0.667{\sim}0.888$ in CPG A, from $0.273{\sim}0.682$ to $0.424{\sim}0.815$ in CPG B).

Discussion Scoring guide reduces the inter-rater disagreement and improves the overall reliability of the Korean-AGREE II instrument.