QUALITY CONTROL SYSTEM FOR MASS SCREENING IN JAPAN

H Naruse[1], A Tsuji[1], E Suzuki[1], N Watanabe[1], M Maeda[2], M Fukushi[3], N Ichihara[4]

[1]Quality Control Center for Mass-screening in Japan, Tokyo; [2]Department of Analytical Chemistry, School of Pharmaceutical Science, Showa University, Tokyo; [3]Department of Public Health, Sapporo City, Sapporo; [4]Institute of Public Health of Hokkaido, Sapporo, Japan

Abstract. Japan was the first country to establish a nationwide quality control system. When the Japanese Federal Government initiated Nationwide Neonatal Screening in 1977, the system officially included a Quality Control (QC) System that should cover all screening laboratories in Japan. This QC system is quite different from that for usual clinical chemistry. The aim of the National QC System for Neonatal Screening is evaluation of the accuracy of the tests and evaluation of the ability to detect suspicious samples with very mild abnormalities. For accomplishing the aim, the QC center established an inter-laboratory QC survey Screening laboratories having weak points can be identified through the inter-laboratory QC survey, and the Center must find a way to improve the ability of these screening laboratories. This requires a nationwide consensus regarding the cut-off levels of tested materials. Based on the cooperation of the Societies For Mass-screening, of Inborn Errors of Metabolism and of Pediatric Endocrinology, we set low cutoff levels for each compound to minimize the number of false negative cases. The system also included the evaluation of the quality of essential screening reagents and the special filter paper for blood collection (in partnership with the production companies). For this purpose, we developed some new methods for evaluating the standard-compounds for the various screening tests exactly, except in the case of TSH screening.

INTRODUCTION

The Japanese Society for Mass Screening (JSMS) plays an important role in the quality control (QC) system in Japan. When it was established in 1973, the Society had kind support from Robert Guthrie, Horst Bickel, Harvey Levy and other international pioneers in the field. At the end of 1977, it had a membership of 380 scientists in various fields with a current membership of around 550. Based on a proposal and requests from JSMS, the Maternal and Child Health (MCH) Division of the Japanese Federal Government established nationwide neonatal screening for phenylketonuria, maple syrup urine disease, homocystinuria and galactosemia in 1977. Screening tests for congenital hypothyroidism and congenital adrenal hyperplasia were added to the nationwide screening program in 1979 and 1988 respectively and the QC System soon covered screening for these disorders (Naruse, 1998).

The National QC Committee controls the QC System. Official members of the committee are: the Director of the MCH division of the Federal Government, specialists in various fields recommended by JSMS and core members of the QC Center for Mass-screening (QC-C). At present, four medical doctors who are specialists of inborn errors of metabolism (IEM), CH, CAH and neuroblastoma, one obstetrician who represents the Japanese Association for Maternal Welfare, the chairman of the QC committee of JSMS, and the representative of the Technical Staff Section of JSMS have been officially recommended by JSMS.

The duties of QC-C are described in Table 1. The QC for neonatal screening has special features. In usual clinical situations, the judgment of normal or abnormal (sick) newborns should be done by a clinician based on

Table 1. Duties of the Quality Control Center for Mass Screening in Japan.

1	Inter-laboratory quality assurance survey.
2	Quality control of calibrators for bacterial inhibition and other important reagents and filter paper for blood collection.
3	Distribution of "Control Materials" for internal quality control system of screening for inborn errors of metabolism.
4	Analysis of "cut-off levels" and "rate of second samples requested" in all screening laboratories in Japan.
5	Training, instruction, consultation for technical staffs in screening laboratories in Japan.
6	Distribution of "primary standards" for mass screening.

Table 2. Inter-laboratory Quality Assurance Program in Japan.

- Send "External Quality Control Materials" each month, which are mixtures of normal dried blood and simulated abnormal samples containing slightly elevated amounts of target compounds, to all screening laboratories in Japan.
- The total number of External Quality Control samples is fixed at ten but the number of abnormal samples is always changing.
- Technical staffs are asked to detect all samples that contain slightly elevated amounts of target compounds and report back promptly.
- When any abnormal samples are undetected, core staffs in the QC Center contact the responsible technical staff and discuss the reason for the mistake.

clinical findings (including various kinds of examinations and careful observations of the patient). However, in neonatal screening, clinical doctors usually cannot distinguish suspicious babies from normal ones. Only people in the screening laboratory will be able to pick up suspicious samples based on biochemical testing results. However, the judgment for distinguishing slightly abnormal samples from normal samples seems to be a very difficult task. Therefore, the QC system evaluates not only the accuracy of the screening tests, but also the adequacy and competency of judgment in detecting "not normal samples" or babies that should be followed up carefully by the clinical specialist (Naruse and Suzuki, 1998).

For this purpose, a special inter-laboratory QC survey was introduced (Table 2). Samples sent from QC-C for this purpose are called 'External QC Samples' (Ext. QC). As described in the next section, we have set "low cut-off levels" based on recommendations from clinical specialists, Ext. QC samples containing "slightly elevated amounts" of the target compounds have been routinely distributed. Based on these surveys, the accuracy of the screening tests in each laboratory and the adequacy of judgment regarding abnormal samples can be assessed. When Ext. QC contains high levels of the target compound, people may pick up abnormal ones rather easily. However, detecting slightly elevated samples is a rather difficult task.

During the initial months, many laboratories with untrained technicians, or screening laboratories using inadequate reagents did not detect some of the abnormal samples (Naruse, 1980). Now, however, almost all screening laboratories in Japan can detect these samples. When a laboratory shows repeated mistakes, the core staff of QC-C conducts thorough discussions with

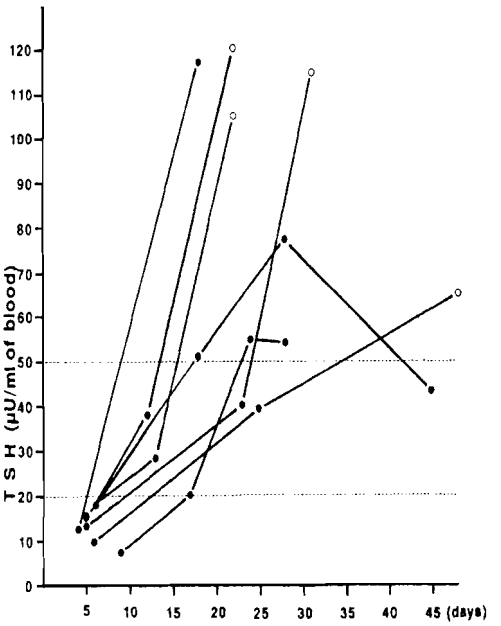

Fig 1. Time course of TSH levels in dried blood specimens in patients with congenital hypothyroidism with delay in marked TSH elevation detected at slightly elevated TSH levels in first specimen. • Indicates TSH levels in dried blood specimens. ○ Indicates TSH levels calculated from serum levels (:1.6).

technical staff at that laboratory, which may include on-site visits. Every year, JSMS organizes a training course for newly appointed technical staff of the screening laboratory in Japan.

NECESSITY FOR LOW CUT-OFF LEVELS IN NEONATAL SCREENING

The problem of cut-off levels has already been discussed in several international or regional meetings including 2nd and 3rd APRM. However, in the case of CH screening, we have not yet reached international consensus. Now, in Asia, many people are trying to develop nationwide or regional CH screening, so it is important to discuss this problem again. An article in J Med Screening Vol 7 (No 4), 2000, there was an article "Screening brief: Screening for congenital hypothyroidism" wherein the author recommended the cut-off level for TSH as 20 mU/l. If this unit means 20 mU/l (mU/ml) of whole blood, it is too high and we cannot accept the recommendation. However, if the author used "20 mU/l of serum", the value is almost the same as 12.5 mU/l of whole blood. Even, with the cut-off level of 12.5 mU/l whole blood, we will miss some mild cases of the real hypothyroid patients. When people

would like to discuss the cut-off levels of TSH, they should express the unit more carefully. We think that in the paper on neonatal screening, the unit per ml of whole blood will be better. However, some people are using "unit per serum" and sometimes we have confusion. The reason of the necessity for low cut-off levels should be described in the following section.

The main purpose of the QC for mass-screening is to reduce the cut-off level as low as possible in order to reduce the number of the false negatives. In the case of CH screening, radioimmunoassay (RIA) was initially used and, at that time, the cut-off level of TSH was 20-25 μU/ml (whole blood). Using this cut-off level, many mild CH or so-called "subclinical" cases may have been missed. In the case of mass-screening, it is important to find these subclinical cases. We introduced an enzyme immunoassay for TSH screening in 1979 and this improved the sensitivity of CH screening. In the early 1980s, commercial companies began to distribute sensitive enzyme linked immunosorbent assays (ELISA). JSMS and the Japanese Society of Pediatric Endocrinology recommended that the cut-off level for TSH should be lowered to 10 μU/ml of whole blood and since the mid-1980s, this cut-off level has been in use.

We previously reported (Naruse *et al*, 1996) that about 25% of confirmed cases of CH in Japan had TSH levels between 10 and 15 μU/ml in the first samples. The percentage of the patients with results between 15 and 20 μU/ml is 4.4%. There is another important reason for a low cutoff. Fig 1 (Harada *et al*, 1995) shows changes of TSH in CH cases with a "so-called delayed TSH elevation". Some of these CH patients showed a very mild elevation of TSH in the beginning [around 10 μU/ml (whole blood) or slightly less], with a later rapid increase in TSH. Thus, care must be taken to detect mild elevations of TSH in initial screening samples from 4-6 day old newborns. In Japan, great care is taken in evaluating TSH results near 10 μU/ml. Relying on internal QC data, samples exhibiting TSH levels of 8-10 μU/ml result in a request for a second sample.

It has also been reported (J Med Screen, 2000) that the detection rate for CH using a cut-off of 20mU/l is around 90%. However, our studies (Naruse *et al*, 1996) show that in Japan, if the cut-off level is 20 mU/l, we might miss about 30% of CH patients. Some people believe that mild elevation of TSH may not cause the brain damage. However, Harada's data shows some babies who had mild elevation in the beginning might have remarkable elevation of TSH later. Thus, in the screening, we need to detect babies with even mild elevation of TSH to prevent brain damage at a later time. The National Academy of Clinical Biochemistry in the USA has noted that detecting milder CH patients (so-called "subclinical") is important. Similar opinions have been reported in various international meetings. In Japan, we have had the same opinion from specialists since the beginning of the CH screening era.

Some people in East Asia still depend on the old opinion that recommends high cut-off levels of 25-30 μU/ml TSH. On the other hand, many screening laboratories are now using cut-off levels of 10 μU/ml. When people wish to start the nationwide or regional screening, they need to try to set-up low cut-off levels to reduce the number of the false negative cases. In the case of CH screening, we will always have a certain number of false negatives. However, we can reduce the number of the false negative (next section). Fortunately, in Japan, we do not have the problem of early discharge. If people have the blood samples taken from 0-2 days old babies, CH screening may cause many problems. CH screening in such areas must develop another better method of screening. (The rapid progress of medicine seems to provide us some solution.)

Regarding cut-off levels for PKU, MSUD, galactosemia, homocystinuria and CAH in Japan, all are set up using low cut-off levels. JSMS had close communications with the Japanese Societies of IEM and of Endocrinology, and based on recommendations from specialists, guidelines for cut-off levels for these screening tests were issued. The detailed information about the cut-off levels are contained in "The Guideline for Neonatal Screening" issued by JSMS.

NATIONAL SURVEY OF FALSE NEGATIVE CASES FROM NEONATAL SCREENING

For the evaluation of the efficacy of nationwide neonatal screening and for the QC system, a national survey of false negative cases is important. K Aoki conducted a survey of false negative reports of PKU, MSUD and homocystinuria in cooperation with IEM specialists and the Japanese Cooperation Projects of Special Formula (Aoki *et al*, 1999). This report detected 4 false negative reports of homocystinuria, 1 report of PKU, and 4 reports of MSUD. The reason for the missed PKU was not known. For homocystinuria, the methionine cut-off level was 1.5 – 2 mg/dl but not all cases have elevated methionine. Some screening laboratories measure homocysteine using high performance liquid chromatography (HPLC). However, there is not yet any scientific data comparing the efficacy of screening by both methods. In case of MSUD, the three cases were the

intermittent type. The details of the remaining MSUD case is known.

H Inomata has been conducting a survey of false negative cases of CH in cooperation with endocrinologists. Since nationwide screening began until the end of fiscal year 1999, 35 cases of primary CH were diagnosed outside of screening while 6194 cases of primary CH patients were detected and treated. Thus, the percentage of false negative cases to detected cases is 0.56%. Twenty-seven cases had lower TSH levels than the cut-off level in the first sample and three cases were missed for unknown reasons, probably due to clerical error. There are no clear reports as to why the other cases were missed.

K Tachibana has been following the false negative cases of CAH since the beginning of nationwide screening (Tachibana *et al*, 2001). According to this report, there were six false negative cases. In four of these, the first samples were tested using a reagent that had elevated cross-reactivity to other compounds. If cross-reactivity is high, ether extraction should be used, however in three of four cases, this was not done. Three other cases had a 17-OHP value below the cut-off and a late rise in 17-OHP may account for these results.

RECENT PROGRESS RELATING TO THE GENERAL QUALITY OF SCREENING LABORATORIES IN JAPAN

1) Certificate for "Qualified Technologists in Neonatal Screening" (Naruse, 1998)

Several years ago, JSMS established a system to issue a special certificate for "Qualified Technologists for Mass-screening" based on an agreement with the Federal Government. The technical staffs were recognized as "Qualified Technologists" upon: 1) having participated in mass-screening for a certain year, 2) having given a certain number of reports or scientific papers in an academic meeting, and 3) having participated in a certain number of training courses organized by JSMS, universities or special foundations. It is hoped that in the future, core staffs of all screening laboratories will have this title and that this system will encourage young technologists and improve the quality of screening laboratories in the future.

2) The standard software for data-management in the screening tests

A missed CH case prompted the national QC committee to recommend standard software for all data management and reports to the family. The MCH Division of the federal government provided a grant for the development of this software and a core group in the technical staff section of JSMS spent several years developing the "Standard Software For Data Management in Neonatal Screening" (Fukushi *et al*, 1999) to be distributed by JSMS. The group also established a standard model for "Internal Quality Control System of Neonatal Screening Tests" (Ichihara *et al*, 1988). This software will also provide data management for internal QC. Use of this software is intended to avoid careless mistakes and improve the internal QC system of the screening laboratories in Japan.

3) Development of accurate assay methods for the standards in the screening tests

During the first few years, we established exact assay methods for amino acids and galactose in dried blood using HPLC and enzymatic methods. However, some people wanted to have some back-up data on the accuracy of HPLC methods. Therefore, we tried to establish a more accurate method for evaluating the standards for Phe. With support from the National Center For Neurology and Psychiatry, carbon-13 labeled phenylalanine was obtained and an exact assay method using GC/MS was established for dried blood samples during 1985-87. Using this method, random dried blood standards were analyzed and the HPLC method was validated.

Screening for CAH began in Japan using an EIA assay for 17-OHP in 1980. Nationwide screening for CAH is completely dependent upon the EIA method (Tsuji *et al*, 1983). Two companies now make three different types of ELISA kits for 17-OHP. The cutoff level is 8-10 ng/ml (whole blood). In initiating the Japanese Inter-laboratory QA Survey for CAH, it was realized that adult QC samples could not be used because newborn blood contains high amounts of a fetal steroid that can lead to cross reaction, even in good kits with high specificity. By mixing 17-α-hydroxy-pregnenolone-3-sulfate (17OHP-3-Sulf) with adult blood, control materials simulating newborn blood can be prepared. A. Kanbegawa kindly synthesized sufficient amounts of 17OHP-3-Sulf for the QA survey. In preparing external QC materials, 100 ng/ml of the compound is added to adult blood (Hct 55%) and mixed with appropriate amounts of 17-OHP to prepare known concentration materials. Since 1990, when QA surveys for CAH were begun, no serious difficulties were observed through 1999. However, beginning in 2000, B Company's ELISA kits gave higher 17-OHP results than A company's kit. M Maeda developed a column-switching LC/MS method using deuterated 17-OHP (Komiya *et al*, 2001). Primary

standards prepared in this way should eliminate differences such as those above and are being added to the QC-C system.

A nationwide screening system for neuroblastoma analyzing VMA and HVA in urine also exists in Japan. This screening has met with opposition both within Japan and outside. Recently, however, a national committee analyzing the utility of this screening gave the first report that suggested that neuroblastoma screening reduced infant mortality. Recently, the Federal Government decided to initiate an inter-laboratory QA survey for neuroblastoma screening. This screening is performed by HPLC and several different HPLC systems are used. The inter-laboratory QA survey discovered some problems.

The best way to solve this problem will be the standardization of these HPLC methods. The QC-C is tasked with improving the present screening situation and for this purpose, a good primary standard is needed for both VMA and HVA. Presently, M Maeda is trying to establish the LC/MS method using carbon-13 labeled VMA and deuterated HVA. Establishing a good procedure for preparing primary standards for these two analytes will continue to be an important topic.

At present, we have an unsolved problem regarding TSH analytical standards. We need a good 'gold standard' for TSH assays and also an accurate method for measuring TSH very exactly. We have tested various kinds of TSH material. At this moment, we have no reliable standard material which does not have a lot-to-lot difference. A. Hayakawa has established an excellent method for eliminating impurities of the r-h-growth hormone resulting in a very pure product. This method was investigated as a means of purifying TSH but was found to be difficult to use with TSH. Removing impurities from TSH continues to be a important purpose of the research in Japan and the world.

CONCLUSION

The QC system for mass-screening in Japan has contributed positively and significantly towards improving the general level of mass screening in Japan. The system is improving through the cooperation of the different scientists involved. However, the QC system still has some weak points. Presently, there is no truly blinded QC being used in our surveys. It would be preferable to send a sample through the system with a fictitious name in order to truly test the system. However, in Japan, there are many objections to performing a truly blinded Inter-laboratory QA survey. For example, there must be full agreement of the doctors who are taking care of the newborn and of the public health nurses in the health care centers, which are present in all towns and villages. The public must be convinced of the necessity for such a special QA survey. There are probably other weak points in the QA system but through the activities of JSMS, problems will slowly and steadily be overcome so that the level of mass-screening in Japan will improve.

REFERENCES

Aoki K, Inomata H, *et al.* False negative cases of neonatal screening for MSUD and Homocystinuria (in Japanese). Report from National Project for "Establishment of Effective System Preventing False Negative cases of Neonatal Screening." (Project leader: Y. Kuroda) In: The report of research project of MHW of Fiscal Year 1999; 1: 293-4.

Fukushi M, Arai O, *et al.* Establishment and distribution of " Standard Software for Data Management and for Internal Quality Control in Neonatal Screening" (in Japanese). Report from National Project for Establishment of Effective System Preventing False Negative Cases of Neonatal Screening (Project leader Y Kuroda). In: The report or research project of MHW of Fiscal Year 1999;1: 307-8.

Harada S, Ichihara N, *et al.* Later manifestation of congenital hypothyroidism predicted by slightly elevated thyrotropin levels in neonatal screening. *Screening* 1995;3:181-92.

Ichihara N, Terai I, *et al.* An effective internal quality control system of the screening for inborn errors of metabolism – questionnaire of usage of internal quality control material. *J Jpn Soc Mass-screening* 1998;8 (suppl 1):171-2.

Komiya T, Maeda M, *et al.* Determination of 17α-hydroxyprogesterone (17-OHP) in standard dried blood disc using column switching LC/MS (in Japanese). *J Jpn Soc Mass-screening* 2001;11:95.

Naruse H. System of neonatal screening for inborn errors of metabolism in Japan. In: Bickel H, Guthrie R, Hammersen G, eds. Neonatal screening for inborn errors of metabolism. Berlin: Springer-Verlag, 1980: 299-305.

Naruse H. Neonatal screening in Japan - past and present. *J Jpn Soc Mass-screening* 1998;8:17-28.

Naruse H, Suzuki E, *et al.* Problem of primary standard in CH screening (2). Lot to lot difference of recombinant h-TSH (in Japanese). Report from National Project for "Research on Effective Mass-screening System". (Project leader T Aoki) In: Project Report of Fiscal Year 1997: 141-3.

Naruse H, Suzuki E. Quality control in neonatal screening in Japan. In: Proceedings of the 6th International

Symposium on Quality Control – Osaka. Amsterdam: *Excerpta Medica, Curr Pract Ser* 1998; 44: 23-38.

Naruse H, Watanabe N, Harada S, *et al.* Necessity of low cut off levels of TSH in neonatal hypothyroid screening. In: Levy HL, Hermos RJ, Grady GF, eds. Proceedings of the third meeting of the International Society For Neonatal Screening, Boston, USA Oct. 20-23,1996. Watertown:IKON-MAP 1996: 228-31.

Tachibana K, Inomata H, *et al.* National survey for false negative cases through CAH screening and evaluation of the cut-off level of 17-OHP in the CAH Screening (in Japanese). *J Jpn Soc Mass-screening* 2001;11:47-52.

Tsuji A, Maeda M, *et al.* Fluorescence enzyme immunoassay of 17α-hydroxy- progesterone and its application to mass screening for congenital adrenal hyperplasia. In: Naruse H, Irie M (eds.). Neonatal Screening Amsterdam: *Excerpta Medica Int Congress Ser* 1983; 606: 324-9.