



Data Management and Site-Visit Monitoring of the Multi-Center Registry in the Korean Neonatal Network

Chang Won Choi^{1,2} and Moon Sung Park³

¹Department of Pediatrics, Seoul National University College of Medicine, Seoul; ²Department of Pediatrics, Seoul National University Bundang Hospital, Seongnam; ³Department of Pediatrics, Ajou University School of Medicine, Suwon, Korea

Received: 25 April 2015

Accepted: 23 July 2015

Address for Correspondence:

Moon Sung Park, MD

Department of Pediatrics, Ajou University School of Medicine, 164 Worldcup-ro, Youngtong-gu, Suwon 16499, Korea
Tel: +82.31-219-5006, Fax: +82.31-219-4004
E-mail: drparkms@ajou.ac.kr

Funding: This work was supported by the Research Program funded by the Korea Centers for Disease Control and Prevention (2013-E63008-01).

The Korean Neonatal Network (KNN), a nationwide prospective registry of very-low-birth-weight (VLBW, < 1,500 g at birth) infants, was launched in April 2013. Data management (DM) and site-visit monitoring (SVM) were crucial in ensuring the quality of the data collected from 55 participating hospitals across the country on 116 clinical variables. We describe the processes and results of DM and SVM performed during the establishment stage of the registry. The DM procedure included automated proof checks, electronic data validation, query creation, query resolution, and revalidation of the corrected data. SVM included SVM team organization, identification of unregistered cases, source document verification, and post-visit report production. By March 31, 2015, 4,063 VLBW infants were registered and 1,693 queries were produced. Of these, 1,629 queries were resolved and 64 queries remain unresolved. By November 28, 2014, 52 participating hospitals were visited, with 136 site-visits completed since April 2013. Each participating hospital was visited biannually. DM and SVM were performed to ensure the quality of the data collected for the KNN registry. Our experience with DM and SVM can be applied for similar multi-center registries with large numbers of participating centers.

Keywords: Clinical Data Management; Site-Visit Monitoring; Prospective Registry; Multi-Center Study

INTRODUCTION

The Korean Neonatal Network (KNN), a national prospective registry of very-low-birth-weight (VLBW, < 1,500 g) infants in the Republic of Korea, was launched on April 15, 2013. As of March 31, 2015, clinical data for VLBW infants have been collected from 55 participating hospitals across the country. The KNN registry is funded by the Korea Centers for Disease Control and Prevention (Korea CDC) and is intended to identify clinical indicators for VLBW infant care through the investigation of clinical outcomes and their associated risk factors (1). The integrating principal investigator (PI) of the KNN registry is Dr. Won Soon Park at the Samsung Seoul Hospital (headquarters hospital) and the PIs involved are the staff neonatologists at participating hospitals across the country. The KNN registry is expected to serve as a research platform for the international comparison of neonatal care, development of standard neonatal care, and improvement of outcomes for VLBW infants in Korea (2).

The subjects for the KNN registry were VLBW infants born in KNN-participating hospitals or transferred to those hospitals within 28 days after birth. By March 31, 2015, there were 55 hospitals participating in the KNN registry. The KNN registry collects two sets of data: short-term data collected from the medical records of VLBW infants during their initial admission to the

neonatal intensive care unit (NICU) and long-term follow-up data collected at the outpatient clinic at 18-24 months of corrected age and 3 yr of chronological age. Since data collection for long-term follow-up data only began in earnest in 2015 at the majority of the participating hospitals, only the collection process for short-term registry will be covered in this article.

Data on 116 demographic and clinical variables were collected from the medical records of VLBW infants during their initial admission to the NICU for the short-term registry. Collected data are entered into an electronic case report form (e-CRF) at each participating hospital by the PI or clinical research coordinator (CRC) delegated by the PI at that hospital. Fifty-five hospitals across the country participated in the KNN registry. The area of the Republic of Korea is 99,720 km² and the longest distance among the participating hospitals is 411 km. Taking this into account, clinical data management (DM) and site-visit monitoring (SVM) are crucial to maintain quality of the collected data. In the present article, we describe the process of clinical DM and SVM performed during the establishment stage of the KNN registry.

MATERIALS AND METHODS

Clinical data management

For proper management of clinical data collected for the KNN

Table 1. Clinical variables collected in the KNN registry

Category	Variable
Identification code	Hospital code, patient code, patient initial, subject initial, birth date, sex, date of entry agreement, age at entry
Maternal information	Age, parity, gravida, amniotic fluid amount, educational status for parents, country of origin for parents, marital status
Delivery information	Premature rupture of membrane, antenatal steroid, delivery mode
Basic information	Birth date, birth place, admission date, gestational age, apgar score, initial resuscitation at delivery room (oxygen/positive pressure ventilation/intubation/cardiac massage/epinephrine), birth weight/height/head circumference, at admission temperature/pH/base excess
Respiratory	Air leak, pulmonary hemorrhage, respiratory distress syndrome, surfactant (use/purpose/place/time/frequency) steroid (use/method/type of drug), at age of 28th day (use for oxygen/ventilator), at 36th week of corrected age (date/oxygen use/ventilator), duration of invasive/noninvasive ventilator/oxygen
Cardiovascular	Patent ductus arteriosus (PDA) (medication/type/timing), PDA ligation (date), pulmonary hemorrhage, hypotension
Central nervous system	Neonatal seizure, intraventricular hemorrhage, post-hemorrhagic hydrocephalus, cystic periventricular leukomalacia
Infectious	Congenital infection, bacterial/nonbacterial infection, meningitis, prophylactic antifungal agent use
Gastrointestinal	TPN duration, necrotizing enterocolitis (\geq stage 2/operation), idiopathic intestinal perforation (operation)
Retinopathy of prematurity	Retinopathy of prematurity (stage/operation/anti-VEGF treatment)
Anomaly	Serious congenital anomalies
Others	Full feeding day, transfusion of red blood cells
Discharge information	Date, hospitalization duration, type of discharge, assistant equipment or treatment, transfer (name of hospital/purpose), death (cause), at discharge (height/weight/head circumference)

This table was cited from the reference #1.

registry, a DM team was organized. The DM team consists of staff neonatologists at the major academic hospitals in the Seoul metropolitan area and clinical research associates (CRAs) at the headquarters hospital. All clinical data of the VLBW infants were entered into the e-CRF established using the iCReaT (Internet based Clinical Research and Trial management system). The iCReaT is a web-based electronic data capture system developed by the Korea CDC. One hundred and sixteen clinical variables are collected from the source documents (electronic or paper medical records) into the e-CRF at each participating hospital. The clinical variables collected are shown in Table 1. All personnel responsible for clinical data entry at each hospital completed a basic educational course officially provided by the Korea CDC for the usage of the iCReaT system. A manual of operation (MOP) for data entry was developed by the KNN short-term registry task force and distributed to all participating hospitals to ensure a clear and uniform definition of each variable by all data entry personnel. Data entry into the e-CRF at each hospital was restricted to the PI or a CRC delegated by the PI of that hospital. The PI must be the staff neonatologist and a CRC delegated by the PI could be another staff neonatologist or attending nurse at the participating hospitals. Rotating residents or fellows were not eligible as CRCs to maintain consistency of data entry.

At the data entry step, automated proof checks, which are the processes of using software for checking proofs for data correctness, were run by the iCReaT system, serving as edit checks. The DM team of the KNN created a checklist that was programmed into the iCReaT system during the e-CRF development stage. Most of automated proof checks programmed into the e-CRF were either for missing values or simple range checks. These automated proof checks prevent data omission and errors during data entry (3).

After clinical data was entered on the e-CRF and confirmed by the PI of the participating hospital, data validation was performed by CRAs in the DM team. The DM process included data cleaning through electronic review using an external program for logic errors, query creation, query resolution, and corrected data revalidation (4). Because the e-CRF developed for the KNN registry included minimal text data, manual data review was not performed. The queries created fell into two categories: data entry errors and reconfirm errors. A reconfirm error occurred when a query was created but the data was confirmed as correct. The detailed process of DM was as follows: 1) the data were checked if they were within a preset range of values according to the distributed MOP for data entry; 2) the data were checked for logic errors using SAS software (SAS[®] 9.4, SAS Institute, Cary, NC, USA); 3) queries were created using the iCReaT system; 4) the corresponding PI was notified of the creation of queries via e-mail; 5) query resolution and data correction was confirmed; 6) additional queries were created for errors requiring further confirmation; and 7) the corrected data were revalidated and frozen if no further errors were found. A schematic outline of the KNN DM process is illustrated in Fig. 1. Authorization for data entry was given to the PI or a CRC delegated by the PI at the participating hospital, but data confirmation was authorized only by the PI. Data locking was under the authority of CRAs in the DM team at the headquarters hospital. The PI at each participating hospital only had access to the data collected from his or her hospital. Access to the whole data set from all participating hospitals was restricted to the CRAs in the DM team.

Site-visit monitoring

SVM included visit site selection, SVM team organization, confirmation of research-related documents, identification of unregistered cases, source document verification, and post-visit

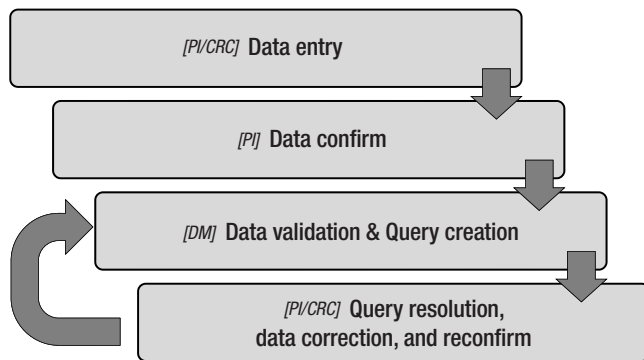


Fig. 1. A schematic outline of the Korean Neonatal Network data management process. PI, principal investigator; CRC, clinical research coordinator; DM, data management team.

report production. A total of 55 hospitals had participated in the KNN registry across the country by March 31, 2015. Participating hospitals with more than five registered VLBW infants were subject to SVM. For SVM, an SVM team was organized by dividing the country into the Seoul metropolitan area and six provinces (Fig. 2). In the first half of the registry period, the SVM team consisted of CRAs at headquarters hospital and PIs at participating hospitals located in the Seoul metropolitan area. However, for the second half of the registry period, PIs at participating hospitals in each province joined the SVM team and participated in SVM in their province. All participating hospitals were visited biannually. One or two PIs and one CRA visited one or two hospitals at a time. During SVM, institutional review board (IRB) approval status, informed consents, investigator files, and the e-CRF were checked. The most crucial parts of SVM were the identification of unregistered cases and source document verification. Unregistered cases were identified by comparing the number of VLBW infants admitted to the hospital with the number of created IDs in the e-CRF during a certain period of time. If unregistered cases were found, the reasons for omission were documented. Source documents were electronic medical records at most of the participating hospitals. Five to ten percent of the registered cases (at least five cases) were randomly selected and verified for any differences in 23 predetermined clinical variables between the medical records and e-CRF. After SVM, follow-up letters reporting the results of the site-visit and a corrective action request form were sent to the PI of the hospital by the CRA who conducted the SVM. The requested corrective action was to be completed and the corrective action request form signed and answered by the PI at the hospital within one month after the SVM. The CRA who performed the site-visit confirmed the receipt of the corrective action request form and provided feedback to the PI at the hospital visited.

Ethics statement

In all KNN-participating hospitals, a grant for the KNN registry

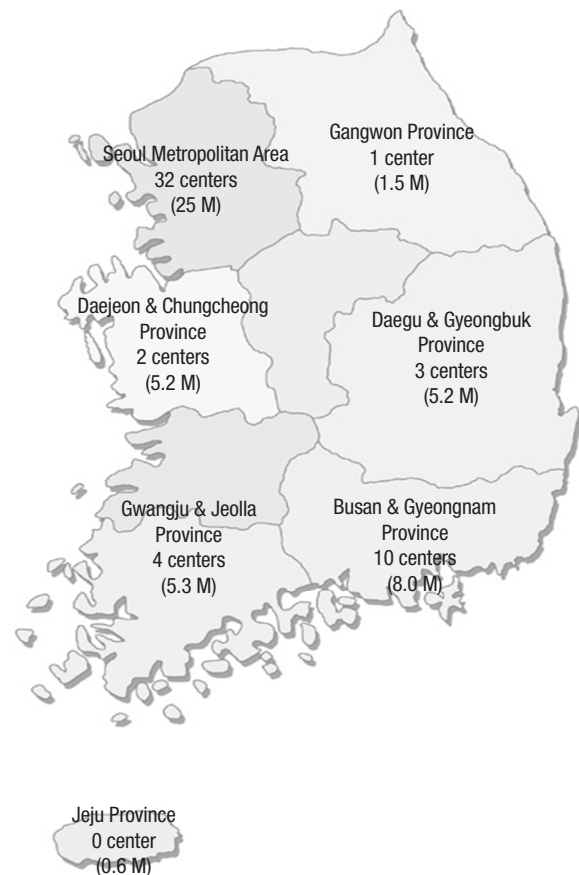


Fig. 2. The regional distribution of the participating hospitals. Fifty-two hospitals were visited since April 2013. The 61.5% of the hospitals were located in the Seoul metropolitan area. The numbers in the parentheses indicate the population of each province. M, millions.

was obtained from the IRB at each participating hospital and consent from the parents of VLBW infants was obtained (MED-OBS-13-014).

RESULTS

Clinical data management

By March 31, 2015, 4,063 VLBW infants were registered in the KNN registry, with 1,693 queries created since April 2013 (Fig. 3). As of March 31, 2015, 1,629 queries have been resolved and 64 queries remain unresolved. Of the 1,629 created queries, 1,515 queries indicated a data entry error and 114 queries, a reconfirm error. A second query was created for 67 cases (Fig. 4). The ratio of the number of queries created to the number of registered VLBW infants was quite different among the participating hospitals, ranging from 6.3% to 211.1% (median, 32.1%). Queries were created in all variables in the e-CRF, but concentrated on several variables. The three most common variables for query creation were the date of 36 weeks postmenstrual age, bronchopulmonary dysplasia, and total duration of invasive mechanical ventilation. Half of the queries were created for these three

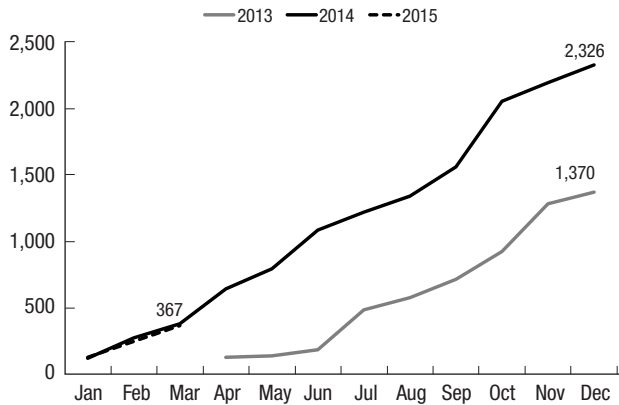


Fig. 3. The cumulative numbers of very-low-birth infants registered in each year since April 2013.

variables, from the 116 variables in the e-CRF. The average time between query creation and query resolution also differed greatly among the participating hospitals, ranging from 3 days to 180 days (median, 18 days).

Site-visit monitoring

By November 28, 2014, 52 participating hospitals had been visited, with 136 SVMs completed since April 2013. Each participating hospital was visited biannually. In 2013, 37 participating hospitals were visited once between September 5 and December 18. In the first half of 2014, 47 participating hospitals were visited from April 29 to August 4. In the second half of 2014, 52 participating hospitals were visited from October 8 to November 28. The 52 participating hospitals were located throughout the country but were concentrated in the Seoul metropolitan area (32/52) (Fig. 2). By November 28, a total of 42 personnel, including three CRAs at the headquarters hospital and 39 PIs at the participating hospitals, contributed to 136 SVMs in a SVM team since April 2013 (Fig. 5). One PI participated in SVM 1-6 times per year (median, twice per year). Each SVM lasted 2-4 hr. The IRB approval status, informed consents, investigator files, and e-CRF were checked by the visiting CRA, requiring about one hour. Identification of unregistered cases and source document verification were performed by the visiting PI, requiring about 2-3 hr, depending on the number of registered VLBW infants. The reasons for unregistered cases fell into three categories: death of the subject, absence of parental consent, or delayed registration. The problem of deregistration due to death of the subject was resolved by modifications to the study protocol. An updated study protocol enabled the inclusion of dead VLBW infants without parental consent. This updated study protocol was granted by the IRB at each participating hospital. Causes of discrepancy between the source document and the e-CRF fell into four categories: simple entry error, misunderstanding of the inclusion and exclusion criteria, misunderstanding of the definition of the variables, and differing interpretations of

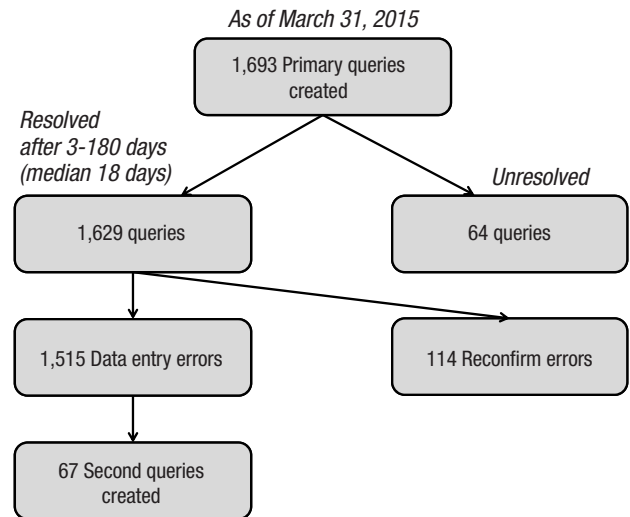


Fig. 4. The result of query management of the Korean Neonatal Network as of March 31, 2015. A reconfirm error occurs when a query was created but the data was confirmed as correct.

clinical findings. Problems identified during source document verification were archived for incorporation into an updated version of the MOP for data entry. The last part of the site-visit was used to interview the PI at each hospital about difficulties in data registration at the hospital. Problems identified during SVM and feedback from the visited hospitals were discussed and incorporated into an updated checklist for subsequent SVMs by the data subcommittee members at periodic off-line meetings.

DISCUSSION

The KNN registry project is the first monumental nationwide registry for preterm infants overseen by the Korean Society of Neonatology. As of March 31, 2015, 55 hospitals, including almost all major NICUs in Korea, have participated. DM and SVM were crucial in ensuring data quality in the collection of 116 clinical variables from the 55 participating hospitals. Efforts to ensure data quality began at the e-CRF development stage. A short-term registry task force was formed to select the short-term clinical variables to be collected and prepare a detailed MOP for each variable. Thanks to the detailed MOP for data entry, confusion in data entry at the participating hospitals was markedly reduced. However, discrepancies in the definitions of clinical variables continued to be a problem throughout the registry project. Thus, a complementary on-line bulletin board to answer data entry questions was established and operated by the official KNN registry homepage (<http://knn.or.kr/index.jsp>). Furthermore, entry errors from misinterpretation of the definition of each variable were found, discussed, and corrected during the site-visits. To reduce errors in the data entry phase, automated proof checks were programmed into the e-CRF. However, these automated proof checks played a limited role (miss-



Fig. 5. Site-visit monitoring performed in the fifty two participating hospitals across the country. By November 28, 2014, 52 participating hospitals had been visited. A total of 42 personnel, including three clinical research associates at the headquarters hospital and 39 principal investigators at the participating hospitals contributed to site-visit monitoring since April 2013.

ing values and simple range checks) in the reduction of data entry errors. Therefore, electronic review for the detection of more complex logic errors was performed using an external program (SAS®) by the CRAs in the DM team at the headquarters hospital after data entry and confirmation by the PIs at the participating hospitals. Queries were created from the data entry errors and sent to the corresponding hospital to be resolved by the PI at his or her hospital. Revalidation of the corrected data was performed by the CRAs in the DM team and second queries were sent to the corresponding hospital when necessary. Through this cycling DM process, data quality could be ensured (5).

The environment of the 55 hospitals participating in the KNN registry varied enormously, not only in the number of NICU beds and attending doctors or nurses, but also in the severities of the conditions of the NICU patients (6). Therefore, a careful understanding of the different environments at each participating hospital was crucial to enroll and maintain as many participating hospitals as possible. In this respect, SVM is an important part of the KNN registry. Thanks to the spontaneous con-

tributions of 39 PIs at participating hospitals, 136 SVMs were performed since the establishment of the KNN registry. The number of participating hospitals increased from 37 to 55 during the project. Of note is the lack of hospitals dropping out of the KNN registry since the beginning of the project. SVM might have encouraged PIs in the participating hospitals to continue registration of VLBW infants. Data quality was also improved through SVM. The identification of unregistered cases at each participating hospital reduced selection bias, which contributes to erroneous conclusions. Source document verification was a crucial part of SVM, along with the identification of unregistered cases. Several errors that were not recognized through the DM process were identified and corrected on-site. Furthermore, confusion over the inclusion or exclusion criteria, definition of the variables, and differing interpretations of the clinical findings were identified and resolved during the site-visit. Although the importance of SVM cannot be emphasized enough, it is not feasible to perform biannual site-visits for 52 participating hospitals due to limited financial and human resources. If not for the voluntary participation of 39 PIs, SVM of the 52 hospitals

would have been impossible. The KNN registry is a nationwide project, and region-based SVM would be more desirable in the future. In the second half of the registration period, PIs in each province participated in SVM in their province, together with the SVM team from the Seoul metropolitan area. PIs will gradually take over local SVM in their provinces, making the project regionalized and sustainable. Based on the passion and voluntary devotion that the Korean neonatologists have shown in the KNN registry, a more advanced nationwide multi-center research network which targets quality improvement and clinical trials will be able to be established.

Maintenance of data quality is crucial in this multi-center registry project. In the KNN registry, data quality was ensured by clinical DM and SVM. By combining these two approaches, the weakness of each method could be complemented. The experience of the KNN registry using clinical DM and SVM can be applied for similar multi-center registry projects with a large of number of participating centers.

ACKNOWLEDGMENTS

We thank Raejin Lee, R.N. and Hee-Young Lee, R.N. at the Samsung Seoul Hospital for their excellent contribution to the DM and SVM and 39 PIs who voluntarily participated in the SVM for their devotion to the KNN registry project.

DISCLOSURE

All authors have no potential conflicts of interest.

AUTHOR CONTRIBUTION

Conception and coordination of the study: Park MS, Choi CW. Design of ethical issues: Park MS, Choi CW. Acquisition of data: Choi CW. Data review: Park MS, Choi CW. Statistical analysis: Choi CW. Manuscript preparation: Choi CW. Manuscript approval: Park MS, Choi CW.

ORCID

Moon Sung Park <http://orcid.org/0000-0002-9101-3581>

Chang Won Choi <http://orcid.org/0000-0003-1911-0253>

REFERENCES

1. Chang YS, Ahn SY, Park WS, and Committee on Program and Planning and Advisory Committee of Korean Neonatal Network. *The establishment of the Korean Neonatal Network (KNN)*. *Neonatal Med* 2013; 20: 169-78.
2. Thakkar M, O'Shea M. *The role of neonatal networks*. *Semin Fetal Neonatal Med* 2006; 11: 105-10.
3. Prokscha S. *Edit checks*. In: Prokscha S. editor. *Practical guide to clinical data management*. 3rd ed. Boca Raton: CRC Press Taylor & Francis Group, 2012, p37-42.
4. Prokscha S. *Cleaning data*. In: Prokscha S. editor. *Practical guide to clinical data management*. 3rd ed. Boca Raton: CRC Press Taylor & Francis Group, 2012, p73-84.
5. Seol HC. *Guideline for electronic data processing and management in clinical trial*. Cheongju: Ministry of Food and Drug Safety, 2012.
6. The Executive Committee of Korean Neonatal Network. *2013 Korean Neonatal Network Annual Report*. Cheongwon: Korean Centers for Disease Control and Prevention, 2014.