Cohort profile: National Investigation of Birth Cohort in Korea study 2008 (NICKs-2008)

Ju Hee Kim, MD1,2, Jung Eun Lee2, So Min Shim2, Eun Kyo Ha, MD3, Dong Keon Yon, MD1, Ok Hyang Kim, MD1, Ji Hyeon Baek, MD1, Hyun Yong Koh, MD, PhD4, Kyu Young Chae, MD1, Seung Won Lee, MD, PhD2, Man Yong Han, MD1

1Department of Pediatrics, CHA Bundang Medical Center, CHA University School of Medicine, Seongnam, Korea; 2Department of Data Science, Sejong University College of Software Convergence, Seoul, Korea; 3Department of Pediatrics, Hallym University Kangnam Sacred Heart Hospital, Hallym University College of Medicine, Seoul, Korea; 4F.M. Kirby Neurobiology Center, Boston Children's Hospital, Harvard Medical School, Boston, MA, USA

Background: An adequate large-scale pediatric cohort based on nationwide administrative data is lacking in Korea.

Purpose: This study established the National Investigation of Birth Cohort in Korea study 2008 (NICKs-2008) based on data from a nationwide population-based health screening program and data on healthcare utilization for children.

Methods: The NICKs-2008 study consisted of the Korean National Health Insurance System (NHIS) and the National Health Screening Program for Infants and Children (NHSPIC) databases comprising children born in 2008 (n=469,248) and 2009 (n=448,459) in the Republic of Korea. The NHIS database contains data on age, sex, residential area, income, healthcare utilization (International Classification of Diseases-10 codes, procedure codes, and drug classification codes), and healthcare providers. The NHSPIC consists of 7 screening rounds. These screening sessions comprised physical examination, developmental screening (rounds 2–7), a general health questionnaire, and age-specific anticipatory guidance.

Results: During the 10-year follow-up, 2,718 children (0.3%) died, including more boys than girls (hazard ratio, 1.145; P<0.001). A total of 848,048 children participated in at least 1 of the 7 rounds of the NHSPIC, while 96,046 participated in all 7 screening programs. A total of 823 infants (0.1%) weighed less than 1,000 g, 3,177 (0.4%) weighed 1,000–1,499 g, 37,166 (4.4%) weighed 1,500–2,499 g, 773,081 (91.4%) weighed 2,500–4,000 g, and 32,016 (3.8%) weighed over 4,000 g. There were 23,404 premature babies (5.5%) in 2008 compared to 23,368 (5.6%) in 2009. The developmental screening test indicated appropriate development in 95%–98% of children, follow-up requirements for 1%–4% of children, and recommendations for further evaluation for 1% of children.

Conclusion: The NICKs-2008, which integrates data from the NHIS and NHSPIC databases, can be used to analyze disease onset prior to hospitalization based on information such as lifestyle, eating habits, and risk factors.

Key message
This national cohort study included all Korean children born in 2008 and 2009 observed over a period of more than 10 years. Our findings demonstrate that it is possible to analyze disease onset prior to hospitalization based on information such as lifestyle, eating habits, and risk factors by integrating National Health Insurance System data with national health screening data.

Introduction
Health care research using big data is regarded positively because of its easy access and ripple effect. Big data have been increasingly used to solve a variety of issues difficult to evaluate in basic or clinical research. There are various sources of data for medical research on children and adults in the Republic of Korea, including medical claims data of the Health Insurance Review and Assessment Service, the national health screening program for adults and the National Health Screening Program for Infants and Children (NHSPIC) of the National Health Insurance Service (NHIS),1 death data of the National Statistical Office,2 cancer patient registration data of the National Cancer Center,3 and the electronic medical records of all hospitals throughout the country.

The NHIS was first established in Korea in 1989, achieving universal coverage after only 12 years and accumulating nationwide healthcare data.1) The NHIS provides benefits for rehabilitation, births, deaths, and health promotion, and for prevention, diagnosis, and treatment of diseases and injuries. The NHIS also maintains records regarding healthcare utilization,
prescriptions, and national health screening programs. The NHIS records are ideal for public health and medical research because they contain data for the whole nation.4] Although several studies have been published utilizing the health screening and healthcare utilization database of the NHIS for adults,4,6 there are currently no comparable noteworthy studies for infants and children.

The health care database for children began to be established when the NHSPIC were conducted to health insurance subscribers in November 2007 and expanded to all of Korea in January 2008.7] The NHSPIC conducts 7 surveys of children, from the age of 4 months to 72 months, and includes a general health questionnaire, developmental screening, assessment of oral health, a physical examination, an anthropometric examination, and anticipatory guidance. Growth, development, and death are important health concern of children around the world. In particular, the NHSPIC makes it easy to determine the nutritional status at infancy, since nutritional information of the screening program includes breastfeeding and complementary feeding, reported to be relevant with growth,8,9 development10,11 and clinical diseases12,13 at childhood. Also, it is useful for analyzing the impact of infantile disease outcomes,14 nutritional status,15 and safety education.16,17

This report describes the establishment of the National Investigation of birth Cohort in Korea study 2008 (NICKs-2008), a large-scale birth cohort study utilizing data from the NHSPIC, the most extensive database on health care in children in Korea, and from the NHIS, a database on the use of medical resources throughout the country. The purpose of this study was to provide basic data, including demographic data, anthropometric index, nutritional intake status, and developmental status for future research by integrating 2 large databases.

Methods

1. Cohort design and participants

The NICKs-2008 acquired data on the use of medical resources from the NHIS and data from the NHSPIC. The NICKs-2008 cohort consists of Korean children born in 2008 (n=469,248) and 2009 (n=448,459) who were followed up through 2017. The use of de-identified individual data for research purposes received ethical clearance under the current National Health Insurance Act. All participants of the cohort are still being tracked, provided they are still eligible for health insurance. Data from 2017 were collected and investigated in the present study. Personal information, including insurance copayment (which is based on income) and residential area were tracked in the eligibility database each year. The Public Information Sharing System, National Tax Service, and Ministry of Health and Welfare of Korea collected eligibility information, and these records are managed by the NHIS. Healthcare utilization information at all visits (inpatient, outpatient, and pharmacy visits) is monitored annually by the NHIS. Data for the cohort will be continuously updated by the NHIS. The protocol of this study was reviewed and approved by the Institutional Review Board of the Korea National Institute for Bioethics Policy (P01-201603-21-005).

2. National Health Screening Program for Infants and Children

The NHIS provides 7 NHSPIC screening surveys to all citizens from the age of 4 months to 71 months (first, age 4–6 months; second, age 9–12 months; third, age 18–24 months; fourth, age 30–36 months; fifth, 42–48 months; sixth, 54–60 months; seventh, age 66–71 months). The NHSPIC survey includes a general health questionnaire, developmental screening, oral health questionnaire, oral exam and guidance, anthropometric exam, physical examination, and anticipatory guidance (Table 1). The parents also provided basic information such as birth date, birth weight, and prematurity status. The general health questionnaire was completed for all 7 surveys and included subcategories of safety, accident prevention, nutrition, vision, and hearing. The other subcategories were oral health guidance (second screening), toilet training (third screening), emotion (fourth screening), hygiene (fifth screening), preschool education (sixth screening), and passive smoking and infection (seventh screening). The subcategory of nutrition collected data about breastfeeding and complementary feeding based on answers to the following questions: “What did you feed your child mainly?” (with possible answers of breast milk, formula, or both); “What kind of water did you use for formula milk?” (with possible answers of tap water, bottled water, purified water, or barley tea); “When did you start complementary feeding?” (with possible answers of before 4 months, 4–6 months, 7 months or later, or not yet started); and “How many times a day did your baby receive complementary feeding?” (with possible answers of 1, 2, 3, or 4 or more).

The developmental screening test is the major screening item of the NHSPIC, and was administered at the second through seventh screenings (Table 1).19 From 2008 to 2013, the developmental screening tool was the Korean-Ages and Stages Questionnaire (K-ASQ) or the Denver II. However, because the Denver II must be administered by a trained examiner and takes more time, the K-ASQ was recommended.19 In 2014, the Korean Developmental Screening Test for Infants and children (K-DST), with results recorded by parents, was developed for preschool children (4–71 months). The K-ASQ has 5 domains (communication, gross motor skills, fine motor skills, problem-solving, and personal-social interrelation) and the K-DST has 6 domains (gross motor skills, fine motor skills, cognition, language, sociality, and self-care). In addition, the K-ASQ interprets the results at 3 different levels (appropriate, need for follow-up, and recommendation for further evaluation),19 and the K-DST interprets the results at 4 levels (advanced development, age-appropriate, need for follow-up, and recommendation for further evaluation).18 Parents of children with an evaluation of “need for follow-up” may be encouraged to promote development of their children in that specific area prior
to the next assessment, conduct a retest, or undergo further evaluation if interviews suggest a problem. An evaluation of “recommendation for further evaluation” means that developmental delay is not confirmed, but careful examination is required. When a child is classified as having a developmental delay based on further evaluation (“recommendation for further evaluation”), it was considered as a “need for consistent management.”

Anthropometric indices (weight, height, body mass index [BMI], and head circumference) were collected from physical measurements during each screening. Height was measured with the child in comfortable clothing while standing upright (for those over 2 years of age) or supine (for those less than 2 years of age) and recorded to the nearest 0.1 cm. If a child aged 2 years or older was measured while supine, 0.7 cm was subtracted from the measured value for calculation of BMI. An electronic scale was used to measure weight while children were dressed in simple clothes and was recorded to the nearest 100 g. The head circumference was measured using a rigid tape with a width of 0.7 cm. Measurements were performed by turning the head horizontally around the upper part of the left ear and the protruding part of the forehead while gently pressing the hair. The standardized scores (z scores) for height, weight, BMI, and head circumference in male and female children of different ages were obtained using the LMS (Lambda for the skew, Mu for the median, and Sigma for the generalized coefficient of variation) method with the 2017 Korean National Growth Charts for those 2 years or older, and using the World Health Organization (WHO) growth standards for infants and young children for those under 2 years of age.

### 3. National health insurance system database

The major variables of the NICKs-2008 cohort (Table 2) were from the NHIS database. The NHIS provides a public database of healthcare utilization, socio-demographic variables, and results from the NHSPIC. Healthcare utilization data are collected when an individual receives healthcare services and

| Table 1. Overview of data collected during the seven National Health Screening Program for Infants and Children screening sessions |
|---------------|-----|-----|-----|-----|-----|-----|-----|
| Screening session | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| | General health questionnaire | | | | | | |
| | Safety accident prevention | √ | √ | √ | √ | √ | √ |
| | Prevention of SIDS | √ | | | | | |
| | Nutrition | √ | √ | √ | √ | √ | √ |
| | Sleep | | | | | | |
| | Vision | √ | √ | √ | √ | √ | √ |
| | Hearing | √ | √ | √ | √ | √ | √ |
| | Oral health guidance | | | | | | |
| | Toilet training | | | | | | |
| | Emotion | | | | | | |
| | Hygiene | | | | | | |
| | Preschool education | | | | | | |
| | Passive smoking | | | | | | |
| | Infection | | | | | | |
| | Developmental screening | | | | | | |
| | Oral health questionnaire | | | | | | |
| | Oral examination and guidance | | | | | | |
| Anthropometric exam | | | | | | |
| | Height | √ | √ | √ | √ | √ | √ |
| | Weight | √ | √ | √ | √ | √ | √ |
| | Head circumference | √ | √ | √ | √ | √ | √ |
| | Vision acuity | | | | | | |
| Physical examination | | | | | | |
| | | | | | | |
| | | | | | | |
| | Anticipatory guidance | | | | | | |
| | Accidents and poisoning | √ | √ | √ | √ | √ | √ |
| | Nutrition | √ | √ | √ | √ | √ | √ |
| | Sleep | | | | | | |
| | Oral health | | | | | | |
| | Toilet training | | | | | | |
| | Hygiene | | | | | | |
| | Passive smoking | | | | | | |

SIDS, sudden infant death syndrome.

Developmental screening tests were conducted using the Korean-Ages and Stages Questionnaire (K-ASQ) in 2008-2013 and using the K-ASQ or the Korean Developmental Screening Test for infants and children (K-DST) in 2014 and later.

| Table 2. Major variables in the National Health Screening Program for Infants and Children database |
|-------------|-----------------|
| Domain | Variable |
| Target health problems | Weight, height, BMI, head circumference |
| Complementary feeding | The time of introduction, type, and amount of complementary feeding, and its regularity |
| Development status | |
| Risk factors | |
| Diagnosis disease | Disease diagnosis (ICD-10 code); operation and procedure history, medication history (drug classification code, dose and duration of drug classification codes) |
| Birth history | Birth weight, prematurity |
| Lactation | Type of lactation, amount of lactation |
| Healthcare utilization | Date of visit, types of medical institution (clinic/hospital/tertiary hospitals/public health center), type of visit (inpatient/outpatient/ emergency/intensive care), length of stay, medical cost (insurer/patient) |
| Healthcare provider | Location, type of hospital, number of beds, medical equipment, human resources, specialties of physicians |
| Socio-economical and demographical factors | Age, sex, residential area, monthly insurance contributions (proxy for income) |

BMI, body mass index; ICD, International Classification of Diseases.
included on all inpatient and outpatient medical records. These data include International Classification of Diseases-10 (ICD-10) code, length of stay, treatment costs, services received, and prescription records (drug classification code, day prescribed). Economic status was classified by quintiles and was determined by the amount of insurance copayment (1, poorest quintile; 5, richest quintile). The residential status of the children was defined as “urban” if they resided in Busan, Daegu, Incheon, Gwangju, Daejeon, and Ulsan, and otherwise as “rural.” The healthcare provider database contains information about the types of healthcare institutions, healthcare human resources, and equipment that was used.

### Table 3. Demographic characteristics of participants in the National Investigation of Birth Cohort in Korea study 2008

| Characteristic          | Birth in 2008 (n=469,248) | Birth in 2009 (n=448,459) |
|-------------------------|----------------------------|----------------------------|
| Sex                     |                            |                            |
| Boys                    | 241,752 (51.5)             | 231,133 (51.5)             |
| Girls                   | 227,496 (48.5)             | 217,326 (48.5)             |
| Birth weight (kg)       |                            |                            |
| 500–999 g               | 3.19±0.46                  | 3.18±0.46                  |
| 1,000–1,499 g           | 367 (0.09)                 | 456 (0.11)                 |
| 1,500–2,499 g           | 1,644 (0.39)               | 1,533 (0.36)               |
| 2,500–4,000 g           | 18,671 (4.38)              | 18,495 (4.40)              |
| >4,000 g                | 388,902 (91.26)            | 384,179 (91.44)            |
| Prematurity             |                            |                            |
| Total                   | 23,404 (5.5)               | 23,368 (5.6)               |
| Death, total            | 1,424 (0.3)                | 1,294 (0.28)               |
| Birth-3 months          | 441 (0.09)                 | 432 (0.09)                 |
| 4–9 Months              | 286 (0.06)                 | 258 (0.05)                 |
| 10–21 Months            | 179 (0.04)                 | 162 (0.03)                 |
| 22–33 Months            | 95 (0.02)                  | 106 (0.02)                 |
| 34–45 Months            | 84 (0.02)                  | 74 (0.02)                  |
| 46–57 Months            | 63 (0.01)                  | 64 (0.01)                  |
| 58–69 Months            | 60 (0.01)                  | 38 (0.01)                  |
| >70 Months              | 216 (0.05)                 | 160 (0.03)                 |
| Birth place             |                            |                            |
| Seoul                   | 147,843 (31.51)            | 90,496 (19.29)             |
| Metropolis              | 96,887 (20.65)             | 106,079 (22.61)            |
| Urban area              | 167,357 (35.65)            | 187,027 (39.86)            |
| Rural area              | 52,107 (11.10)             | 59,965 (12.78)             |
| Economic status         |                            |                            |
| 1st quintile (lowest income) | 38,222 (8.4)           | 35,554 (8.3)               |
| 2nd quintile            | 69,678 (15.3)              | 61,812 (14.4)              |
| 3rd quintile (middle income) | 120,839 (26.6)         | 112,885 (26.3)             |
| 4th quintile            | 144,949 (31.9)             | 138,720 (32.3)             |
| 5th quintile (highest income) | 80,827 (17.8)          | 81,052 (18.9)              |

Values are presented as number (%) or mean±standard deviation.

- Birth weight data were obtained from parental responses in the National Health Screening Program for Infants and Children (NHSPIC) database. Missing data: 43,000 (2008) and 28,331 (2009).
- Prematurity status was from parental responses to the general questionnaire of the NHSPIC and defined as birth prior to 36-week gestation. Missing data: 43,644 (2008) and 29,406 (2009).
- Metropolis areas were defined as 6 metropolitan cities (Busan, Incheon, Gwangju, Daegu, Daejeon, and Ulsan), urban areas as cities, and rural areas as noncities. Missing data: 5,054 (2008) and 4,892 (2009).
- Economic status was classified into quintiles based on insurance copayment. Missing data: 14,733 (2008) and 18,436 (2009).

### Results

#### 1. Demographic characteristics

Basic demographic characteristics were shown in Table 3. A total of 51.5% of the children born in each year were boys. There were only minor differences in birth weight (3.19 kg vs. 3.18 kg) and prematurity (5.5% vs. 5.6%) for the 2 years. A total of 823 infants (0.1%) who weighed less than 1,000 g were born in 2008 and 2009 together. At birth, 3,177 infants (0.38%) weighed 1,000 to 1,500 g and 37,166 infants (4.39%) weighed 1,500 to 2,500 g. Most children were born in urban areas (2008, 35.6%; 2009, 39.9%) and most were also born in Seoul (2008, 31.5%; 2009, 19.3%). Mortality rates were 0.3% in 2008 and 0.28% in 2009. Most deaths (55.13%) occurred within the first 12 months, including 30.97% within the first 3 months. In addition, the risk of death was greater in males than females (hazard ratio [HR], 1.145; 95% confidence interval [CI], 1.061–1.235; P=0.0005) (Fig. 1). Notably, relative to the richest economic quintile, the HR for death was significantly greater in the poorest quintile (HR, 1.233; 95% CI, 1.051–1.447; P=0.01) and the second poorest quintile (HR, 1.402; 95% CI, 1.229–1.599; P<0.001) (Fig. 2).

#### 2. National Health Screening Program for Infants and Children

A total of 848,048 children (92.43%) participated in one or more of the 7 rounds of NHSPIC (Table 4). The highest participation rate was in the 4th session for children born in 2008 (n=256,085, 54.57%) and in the 5th session for children born in 2009 (n=280,319, 62.51%). In addition, 40,360 children born in 2008 or more of the 7 rounds of NHSPIC (Table 4). The highest participation rate was in the 4th session for children born in 2008 (n=256,085, 54.57%) and in the 5th session for children born in 2009 (n=280,319, 62.51%). In addition, 40,360 children born in 2008 and 55,686 children born in 2009 participated in all 7 screening programs.

![Fig. 1. Cumulative numbers of deaths of males and females.](image-url)
We examined the z scores for height, weight, BMI, and head circumference at each of the screening sessions (Table 5). Children with height z scores of 1.63 or more were significantly more common in the first screening (22.79%) and the second screening (10.31%), followed by a slight decrease from the third screening (7.15%) to the fifth screening (5.0%), and slight increases at the sixth screening (6.3%) and seventh screening (8.43%). The z scores for weight and head circumference had trends similar to those for height.

3. NHSPIC - The type of feeding and the time of the introduction of complementary feeding

The general health questionnaire of the NHSPIC collects data on breastfeeding and complementary feeding. A total of 44.2% of children received breastfeeding, 36.1% received formula, 19.3% received breastfeeding and formula, and 0.4% received a special formula (Fig. 3). Analysis of the timing of the introduction of complementary feeding (Fig. 4) indicated that most of children (72.1%) began complementary feeding within 4 to 6 months, 7.8% started before 4 months, 19.8% started after 6 months, and 0.3% started after 12 months.

4. NHSPIC - Developmental screening test

Our analysis of the results from the developmental screening test (Table 6) indicated the highest percentage of children who needed follow-up was at the first screening (4.1%), and the need for follow-up in the other screenings ranged from 1.2% (sixth screening) to 3.3% (seventh screening). The percentage of children with appropriate development ranged from 95.1% (first screening) to 97.8% (sixth screening). About 1% of the children received recommendations for further evaluation.
during each screening, and 0.1 to 0.4% of children were recommended for consistent management.

Discussion

The NICKs-2008 is a study of a nationwide birth cohort for infants and children based on 2 large national administrative databases. These basic data serve as a foundation for future analyses of this cohort. The present study describes the demographic characteristic information, nutritional patterns and the results of developmental screening of children in this birth cohort.

Our demographic results showed that the number of children born annually in Korea has decreased significantly from 636,780 in 2000 to 649,248 and 484,482 in 2008 and 2009, respectively.23 Recent improvements in neonatal care have led to an increase in the percentage of low birth weight infants (<2,500 g), from 3.79% born in 2000 to 4.86% and 4.87% born in 2008 and 2009, respectively.23 Consistent with results showing that child mortality rates are affected by the socioeconomic and educational statuses of parents,24 our study showed that the risk of death was higher in the poorest than in the richest economic quintile.

Growth is one of the most important health concerns at childhood. In particular, obesity is a common chronic inflammatory disease linked with metabolic diseases. Moreover, the increasing prevalence of obesity in children and the relationship between childhood and adult obesity indicate the importance of determining the prevalence of and risk factors for obesity in children.25 We found that the prevalence of obesity, defined as a BMI $z$ score $>1.63$, was 7.85% at 6 years of age (NHSPIC seventh session) and increased from 3 to 6 years of age. The NICKs-2008 provides a useful database to analyze factors associated with the risk of childhood obesity. In contrast, the proportion of children with $z$ scores $>1.63$ in height for age, weight for age, and head circumference for age decreased sharply between the second and third screenings. The finding about more elevated $z$ scores in the first and second screenings is possibly due to the discrepancies between the $z$ scores obtained from the WHO growth standards for infants and the actual growth of Korean children under 2 years of age.22 We believe that future studies are needed to assess utility of this growth chart for assessments of the growth of Korean infants in this age range.

Nutrition has an important effect on growth and development in children.27 Breastfeeding has many advantages, including reductions of respiratory infections, gastrointestinal infections, allergic disease, diabetes, obesity, and neurodevelopment disor-

---

**Table 6. Outcomes of the 7 developmental screening sessions of the National Health Screening Program for Infants and Children**

| Session No. (age) | Participating children (n) | Appropriate development, n (%) | Need for follow-up, n (%) | Recommendation for further evaluation, n (%) | Need for consistent management, n (%) |
|-------------------|---------------------------|--------------------------------|--------------------------|---------------------------------------------|-------------------------------------|
| 2 (9–12 months)   | 428,798                   | 407,804 (95.1)                 | 15,520 (4.1)             | 3,474 (0.8)                                 |                                     |
| 3 (18–24 months)  | 501,881                   | 482,346 (96.1)                 | 15,230 (3.0)             | 4,303 (0.9)                                 | 2 (0.0)                             |
| 4 (30–36 months)  | 534,478                   | 513,870 (96.1)                 | 13,706 (2.6)             | 6,293 (1.2)                                 | 609 (0.1)                           |
| 5 (42–48 months)  | 521,053                   | 505,563 (97.0)                 | 9,238 (1.8)              | 4,639 (0.9)                                 | 1,613 (0.3)                         |
| 6 (54–60 months)  | 506,427                   | 495,245 (97.8)                 | 6,140 (1.2)              | 3,310 (0.7)                                 | 1,775 (0.4)                         |
| 7 (66–71 months)  | 436,538                   | 443,365 (95.7)                 | 15,082 (3.3)             | 3,643 (0.8)                                 | 1,448 (0.3)                         |

The developmental screening sessions 2, 3, 4, and 5 used the Korean Age and Stages Questionnaire (K-ASQ) in all participating children. Missing data: 836 (session 2), 209 (session 3), 178 (session 4), 138 (session 5), 31 (session 6), and 1 (session 7).

During session 6, 502,975 children were screened using the K-ASQ and 3,452 children using the Korean Developmental Screening Test for Infants and Children (K-DST). During session 7, 277,479 children were screened using the K-ASQ and 186,059 children using the K-DST. For K-DST, results of “advanced development” and “age appropriate” were defined as “appropriate.”
ders. In addition, the early or late introduction of complementary feeding has been reported to be associated with growth in children, including rates of overweight and obesity. The NHSPIC can provide easy access to national data about the type of feeding and the timing of the introduction of complementary feeding in infancy. By integrating these data with data from the NHSPIC and NHIS claim data, it may be possible to analyze the relationships of breastfeeding and the timing of the introduction of complementary feeding with various subsequent health problems.

There are several important differences between the NICKs-2008 database and other databases used to date for pediatric research. Table 7 summarizes the data resources that have been widely used for pediatric research in Korea. The Korean National Health and Nutrition Examination Survey database includes the results of blood tests, urinalysis and anthropometric measurements, but these studies were of cross-sectional design and included relatively few children and adolescents. The Korean Youth Risk Behavior survey included a relatively large sized sample, but it was also of cross-sectional design and all survey results were obtained online. The Panel study on Korean children was a prospective, longitudinal, follow-up cohort study. However, the sample size was relatively small and it included limited information about disease and healthcare utilization, as it utilized interviews and questionnaires to obtain results. The National Health Insurance Service-national sample cohort had a wider database than the NICKs-2008, including blood test results and medications. However, since it is a representative 2% sample cohort, the sample size for each age group was relatively small due to the age-specific characteristics of children.

In contrast, the NICKs-2008 had several advantages, as it included all Korean children born in 2008 and 2009. The low attrition rate over a period of more than 10 years was likely due to the rigorous procedures used to collect nationwide data. A second advantage was that 92% of children completed the questionnaire survey and underwent physical examination, anthropometric examination, and developmental screening tests. Thus, these data can be reliably used to study the relationship of personal exposure to multiple factors with physical growth and developmental delay. A third advantage was that the data from the NHIS contained extensive information on healthcare utilization. Integration with the national health screening data makes it possible to analyze disease onset prior to hospitalization based on information such as lifestyle, eating habits, and exposure to risk factors.

However, the NICKs-2008 cohort also has certain limitations. Only limited information is available regarding health-related behaviors, because those data were from reports by parents at the nationwide health screenings. In addition, because recorded information is used for the Korean fee-for-service payment and reimbursement system, the ICD diagnosis code might not accurately reflect a patient’s true health status. Also, in contrast to adults, there is a paucity of blood data for children and adolescents.

The NHSPIC was initially implemented in 2007, but the participation rate remained low over the next 6 years, during which the participation rate per session was only about 50%. Also, the NHSPIC does not include assessments of allergic disorders, which are common in children and of concern to their parents. The importance and validity of the NHSPIC should be communicated to the public and the shortcomings of the NHSPIC should be corrected to increase participation rates.

In conclusion, by integrating NHIS data with national health screening data, the NICKs-2008 could enable analysis of disease onset prior to hospitalization based on information such as lifestyle, eating habits, and exposure to risk factors. These findings may encourage the establishment of additional administrative cohorts for infants and children.

**Footnotes**

Conflicts of interest: No potential conflict of interest relevant to this article was reported.

Acknowledgments: This work was supported by a National Research Foundation of Korea (NRF) grant funded by the Korean

---

**Table 7. Big data resources available for pediatric research use in Korea**

| Title | Design | Organizer | Period | Subject | Data |
|-------|--------|-----------|--------|---------|------|
| Korean National Health and Nutrition Examination Survey | Rolling sample survey Cross-sectional study | Korea Centers for Disease Control and Prevention | Every 3 yr | All ages (about 4,500-5,000 children aged <18 years per session) | Anthropometric data, blood pressure, pulse, grasping power, dental checkup, blood test, urinalysis, questionnaire |
| Korean Youth Risk Behavior Survey | Sample survey Cross-sectional study | Korea Centers for Disease Control and Prevention | Annual | About 60,000 adolescents per a session (first year of junior high school to third year of high school) | Online questionnaire based |
| Panel Study on Korean Children | Prospective longitudinal follow-up cohort study | Korea Institute of Child Care and Education | Annual follow-up | About 2,000 children born in March to May in 2008 | Interview, questionnaire |
| National Health Insurance Service-National Sample Cohort | Retrospective sampling survey | National Health Insurance Service | Annual update | A representative 2% sample | Participant’s insurance eligibility, medical treatment, medical care institutions, general health examinations |
How to cite this article: Kim JH, Lee JE, Shim SM, Ha EK, Yon DK, Kim OH, et al. Cohort profile: National Investigation of Birth Cohort in Korea study 2008 (NICKs-2008). Clin Exp Pediatr 2021;64:480-8. https://doi.org/10.3345/cep.2020.01284