

국가중점 개방데이터(건강검진정보) 사용자 매뉴얼 (ver 4.0)



국민건강보험공단
빅데이터운영실

목 차

제1절 개요

1. 「국가중점 개방데이터」란?	1
가. 개념	1
나. 개방 목적	1
2. 「국가중점 개방데이터」개발 추진 경과	1

제2절 국가중점 개방데이터(건강검진정보) 구성 방향성

1. 국가중점 개방데이터 선정기준	2
가. 안전한 데이터 중심의 개방(Safety) - 개인정보보호 중점	2
1) 그 자체로 개인을 식별할 수 있는 개인식별정보 제외	2
2) 개인정보 비식별화 처리기법 적용	2
나. 활용중심의 데이터 개방(Utilization) - 수요자 맞춤형 데이터 개방	2
다. 객관적이고 신뢰성 있는 데이터 개방(Reliability)	3
2. 개방 데이터셋의 생성 형태	3
가. 개방데이터셋 생성 방식 및 범위	3
나. 개방데이터셋 갱신 주기	3
다. 개방데이터셋 제공 형태	3

제3절 국가중점 개방데이터(건강검진정보) 상세

【건강검진정보】	4
1. 개념 및 구성	4
1) 개념	4
2) 구성	4
2. 개방항목 설명	4
3. 건강검진정보 이용 시 유의사항	8

제4절 국가중점 개방데이터(건강검진정보) 다운로드 방법

- 국민건강보험자료 공유서비스(NHIS) 웹페이지 활용 방법 소개 9

제5절 사용자 지원 사항

- 국가중점 개방데이터 활용가이드 및 사례 제공 11
- 국가중점 개방데이터 개방을 위한 웹사이트를 활용한 온라인 문의기능 지원 11
- 국가중점 개방데이터 다운로드 전에 데모자료 제공기능 지원 11

Q&A 묻고 답하기

제1절 개요



참조하세요.

1. 「국가중점 개방데이터」란?

가. 개념

국가중점 개방데이터는 정부3.0 정책 기조에 발맞추어 국민건강보험공단이 보유하고 있는 데이터 중 민간으로부터의 개방 요구가 높은 항목을 중심으로 개방되는 공공개방데이터이다. 국가중점 개방데이터는 공단이 국민건강보험가입자 역할을 수행함에 따라 축적된 국민건강보험가입자의 진료내역정보, 의약품처방정보, 건강검진정보이다. 공단에서는 안전한 데이터 개방을 위하여 개인정보와 민감성데이터는 제외 또는 마스킹 처리를 하였으며, 외부 전문가를 통하여 개방 데이터의 안전성과 관련된 자문을 실시하였다. 제공되는 데이터의 대상 연도는 2002년부터 2015년까지이며, 향후 대상 기간을 지속적으로 확대할 계획이다. 국가중점 개방데이터 개방을 통하여 관련 산업계의 활성화를 통하여 대국민 일자리 창출에 기여할 수 있기를 기대한다.

- ◎ 국민건강정보DB
국민건강보험공단에서 보유한 데이터를 빅데이터로 분석하여 개방데이터 발굴 및 생성을 목적으로 구축한 모집단DB 자격 및 보험료, 건강검진결과 및 문진 진료내역상세(병원/의원 이용 내역), 노인장기요양 등의 정보를 포함하고 있으며 총 테이블 수는 3,544개, 용량은 약 21,236GB 입

나. 개방 목적

첫째, 법률에 근거한 국민의 공공데이터 이용권을 보장하고, 국민건강보험공단이 수집·보유·관리하고 있는 주요 보건 의료 정보에 대한 국민의 알권리를 충족하고자 한다. **둘째**, 국가중점 개방데이터를 활용한 건강정보 관련 산업계의 새로운 형태의 사업 기회 창출을 돋고자 한다. **셋째**, 개인중심의 진료내역, 의약품처방정보를 일반국민에게 개방함으로서 의료기관의 고품질 의료 서비스 제공을 유도하고자 한다. 부가적으로는 건강정보 관련 산업의 발전으로 인한 보건 의료 산업 국가경쟁력 강화와 신사업 개발로 인한 일자리 창출 효과를 목적으로 한다.

2. 「국가중점 개방데이터」 개발 추진 경과



2011년 전국민 5천 만 명의 9년간(2002~2010년)의 건강보험 정보를 수집하여 초기 국민건강정보DB를 구축한 것을 시작으로, 2013년 서울대학교 빅데이터 센터 등 16개 기관과 MOU를 체결하여 시범 연구사업을 추진하였다. 2014년에는 자료 공개 추진을 위하여 국민건강정보DB 공개 적정수수료 체계 개발 컨설팅, 「국민건강정보 자료제공 운영규정(규정 제130호)」을 제정하였고, 2015년 1월 표본코호트DB 이용신청 및 국민건강정보DB 연구자 지원을 위한 전용 홈페이지를 개발하였다. 2015년 12월 민간개방을 위해 진료내역정보, 의약품처방정보, 건강검진정보 데이터 세트으로 구성된 국가중점 개방데이터를 구축 제공하고, 2016년 12월(현재)에 추가 자료 2014~2015년도 자료를 업데이트하여 지속적으로 제공하였다.

제2절 국가중점 개방데이터 구성 방향성

1. 국가중점 개방데이터 선정기준

가. 안전한 데이터 중심의 개방(Safety) - 개인정보보호 중점

1) 그 자체로 개인을 식별할 수 있는 개인식별정보 제외

개인을 식별할 수 있는 개인식별정보(ex. 고유식별정보(주민번호, 국민건강보험자격번호 등), 쉽게 개인을 식별할 수 있는 정보(이름, 전화번호, 주소, 생년월일, 사진 등)) 등은 국가중점 개방데이터 항목에서 제외 하였다.

2) 개인정보 비식별화 처리기법 적용

타 정보와의 결합을 통해 개인 식별이 가능한 정보가 포함되어 있는지 여부를 사전 필터링 하여 재식별 가능성이 존재하는 데이터는 개방 대상에서 제외하거나, 개별 항목에 적합한 비식별화 처리기법을 적용하여 식별가능성을 제외하였다.

◎ 고유식별정보

법령에 따라 개인을 고유하게 구별하기 위해 부여된 식별정보로 ① 주민등록번호 ② 여권 번호 ③ 운전면허의 면허번호 ④ 외국인 등록 번호를 말함(개인정보보호법 시행령 제19조(고유식별 정보의 범위))

구 분	내 용
표본의 추출	<ul style="list-style-type: none">각 연도별 진료 및 건강검진 수진 환자 100만명 무작위 추출진료 환자, 약품처방 환자, 건강검진 수진자를 연도별로 각각 1백 만 명씩 무작위로 추출
자료별 조합제한	<ul style="list-style-type: none">동 자료는 단면자료로써 각 DB별 개인일련번호와 청구일련번호를 다르게 하여 제공(연계 불가능)
개인식별자 제거	<ul style="list-style-type: none">주민등록번호(13) → 개인일련번호(8자리)
범주화	<ul style="list-style-type: none">연령 그룹화연령 → 연령대(5세 단위)85세 이상은 '85세 이상'으로 범주화
데이터 마스킹	<ul style="list-style-type: none">민감상병 D, O, P, X, Y코드(5종, 114개)를 대분류화(예:D***)D(7개 코드) : 남성 생식기관의 양성 신생물O(81개 코드) : 임신, 출산 및 산후기, 자궁외 임신 등P(4개 코드) : 기타 선천성 감염 및 기생충질환X(1개 코드) : 알코올에 의한 불의의 중독 및 노출Y(2개 코드) : 특수목적 코드, 완력에 의한 성적 학대 등
최상위 지역코드 제공	<ul style="list-style-type: none">소규모 지역에 거주하는 표본의 인식을 고려하여 시도코드에 한하여 제공(17개 시도단위 제공)

나. 활용중심의 데이터 개방(Utilization) - 수요자 맞춤형 데이터 개방

국가중점 개방데이터는 개방되는 데이터 항목 및 제공 방법 등에 관련된 수요조사를 실시하였으며, 해당 수요조사 결과를 반영하여 개방항목 및 제공형태 등을 구성하였다. 단 수요조사를 통하여 요구되었던 항목의 일부는 민감성 데이터로 분류되어 개방이 제한되었으며, 공단은 수요자 요구에 대한 지속적인 모니터링을 통하여 실효성 있는 데이터가 개방될 수

있도록 노력할 계획이다.

다. 객관적이고 신뢰성 있는 데이터 개방(Reliability)

국가중점 개방데이터가 다양한 분야에서 믿고 활용될 수 있도록 개방데이터 발굴 시점부터 데이터 값의 신뢰성이 낮거나 검증이 어려운 항목은 개방 항목 대상에서 제외하였다. 예를 들어 건강검진정보 문진항목 중 수검자가 직접 기입하는 항목 등은 데이터 값의 오류를 정제하는 과정을 거쳐 객관성과 신뢰성을 담보하기 위해 노력하였다.

- ◎ 구체적으로 어떤 건강검진 결과를 제공하는지는 「제3절 국민건강정보 데이터 상세 - 4 건강 검진정보- 가. 개념 및 구성- 1) 개념 참조」

2. 개방 데이터셋의 생성 형태

가. 개방데이터셋 생성 방식 및 범위

국가중점 개방데이터는 2002년부터 2015년 기간에 해당하는 국민건강보험가입자 100만 명의 진료내역, 의약품처방내역, 건강검진결과에 대한 정보이다. 100만 명의 선별은 무작위 선별 기법을 적용하였으며, 본 매뉴얼은 건강검진정보에 관한 내용으로 구성하였다.

건강검진정보는 해당 년도에 건강검진을 수진한 국민건강보험가입자 100만 명을 무작위로 선별하고, 항목 선정 과정을 거쳐 선정된 가입자의 기본정보와 검진결과정보를 추출하여 국민건강정보 데이터셋을 구성하였다.

나. 개방데이터셋 갱신 주기

국가중점 개방데이터의 모든 개방데이터셋은 매년 12월에 2년 전 데이터를 추가하는 방식을 취하여 1년 단위로 데이터를 갱신 한다. 예를 들면 2014년도 데이터는 2016년 12월 추가되는 형태이다.

다. 개방데이터셋 제공 형태

데이터셋의 자료 구조(포맷)는 CSV(Comma Separated Value)이며, 데이터 활용 및 다운로드의 편의성을 위하여 연도별로 분할된 형태로 데이터를 제공한다.

제3절 국가중점 개방데이터 상세

【건강검진정보】

1. 개념 및 구성

1) 개념

건강검진정보란 2002년~2015년(14개년)동안 국민건강보험의 직장가입자와 40세 이상의 피부양자, 세대주인 지역가입자와 40세 이상의 지역가입자의 「일반검강검진」 결과와 이들 일반건강검진 대상자 중에 만40세와 만66세에 도달한 이들이 받게 되는「생애전환기건강진단」의 결과이다.

구분	검진대상		검진주기
	1차	2차	
일반건강검진	<ul style="list-style-type: none"> 지역세대주, 만40세 이상 세대원 직장가입자, 만40세 이상 직장피부양자 	1차 검진 결과 질환의심자(고위험군)	2년 1회 (비사무직 매년)
생애전환기 건강진단	<ul style="list-style-type: none"> 만 40세에 도달한 일반건강검진 대상자 만 66세에 도달한 일반건강검진 대상자 	1차 수검자 전체	해당 연령

* 직장가입자 및 지역세대주는 연령제한 없이 건강검진 대상자가 됨. 단 본 건강검진정보에는 20세 미만의 검진결과는 제외하고 구축하였음.

2) 구성

총 34개의 변수로 가입자 일련번호와 ① 수진자 기본정보 : 성, 연령, 거주지 시도코드와 같은 기본정보 ② 건강검진결과 및 문진정보 : 신체, 몸무게, 허리둘레 등 신체사이즈 정보와 혈압, 혈당, 콜레스테롤, 요단백, 간마지피티와 같은 병리검사결과 시력과 청력, 구강검사와 같은 진단검사결과 그 외 음주와 흡연 여부에 대한 문진결과로 구성되어있다.

2. 개방항목 설명

NO	제공항목			속성정보		비고
	표준 항목명	영문명	설명	표현형식/단위	예시	
1	기준년도	HCHK_YEAR	<ul style="list-style-type: none"> 해당 정보의 기준년도를 제공함 	YYYY	2009	
2	가입자 일련번호	IDV_ID	<ul style="list-style-type: none"> 해당가입자에 부여한 일련번호 - 1 ~ 1,000,000 	N	1	
3	성별코드	SEX	<ul style="list-style-type: none"> 해당 정보 대상자의 성별을 제공함 - 성별 : 1(남자), 2(여자) 	N	1	●

◎ 건강검진이란?

- 건강상태 확인과 질병의 예방 및 조기발견을 목적으로 건강검진기관을 통하여 진찰 및 상단, 이학적 검사, 진단검사, 병리검사, 영상의학 검사 등 의학적인 검진을 시행하는 것
(건강검진기본법 제3조(정의) 제1호)

◎ 국가건강검진이란?

- 국가와 지방자치단체가 시행 하는 건강 검진
 1. 「모자보건법」에 따른 영유아에 대한 건강검진
 2. 「영유아보육법」에 따른 영유아에 대한 건강검진
 3. 「학교보건법」에 따른 초중고등학교 학생의 건강검진
 4. 「청소년복지지원법」에 따른 청소년 건강진단
 5. 「국민건강보험법」에 따른 건강검진
 6. 「산업안전보건법」에 따른 일반건강진단
 7. 「의료급여법」에 따른 건강검진
 8. 「암관리법」에 따른 암검진
 9. 「노인복지법」에 따른 건강진단
 10. 그 밖에 보건복지부령으로 정하는 건강검진
(건강검진기본법 제3조(정의) 제3호 각 호)

NO	제공항목				속성정보		비고																																								
	표준 항목명	영문명	설명		표현형식/단위	예시																																									
4	연령대 코드(5세 단위)	AGE_GROUP	<ul style="list-style-type: none"> 기준년도에 수진자의 나이를 5세 단위로 그룹화(범주화)하여 구분한 코드 - 5세 단위 그룹화, 85세 이상은 85+로 그룹화 - 2002~2013년 까지 <table border="1"> <thead> <tr> <th>그룹</th><th>연령대</th><th>그룹</th><th>연령대</th></tr> </thead> <tbody> <tr><td>1</td><td>20~24세</td><td>8</td><td>55~59세</td></tr> <tr><td>2</td><td>25~29세</td><td>9</td><td>60~64세</td></tr> <tr><td>3</td><td>30~34세</td><td>10</td><td>65~69세</td></tr> <tr><td>4</td><td>35~39세</td><td>11</td><td>70~74세</td></tr> <tr><td>5</td><td>40~44세</td><td>12</td><td>75~79세</td></tr> <tr><td>6</td><td>45~49세</td><td>13</td><td>80~84세</td></tr> <tr><td>7</td><td>50~54세</td><td>14</td><td>85세+</td></tr> </tbody> </table>					그룹	연령대	그룹	연령대	1	20~24세	8	55~59세	2	25~29세	9	60~64세	3	30~34세	10	65~69세	4	35~39세	11	70~74세	5	40~44세	12	75~79세	6	45~49세	13	80~84세	7	50~54세	14	85세+	11 ●							
그룹	연령대	그룹	연령대																																												
1	20~24세	8	55~59세																																												
2	25~29세	9	60~64세																																												
3	30~34세	10	65~69세																																												
4	35~39세	11	70~74세																																												
5	40~44세	12	75~79세																																												
6	45~49세	13	80~84세																																												
7	50~54세	14	85세+																																												
<ul style="list-style-type: none"> 2014년 이후 <table border="1"> <thead> <tr> <th>그룹</th><th>연령대</th><th>그룹</th><th>연령대</th></tr> </thead> <tbody> <tr><td>1</td><td>0~4세</td><td>10</td><td>45~49세</td></tr> <tr><td>2</td><td>5~9세</td><td>11</td><td>50~54세</td></tr> <tr><td>3</td><td>10~14세</td><td>12</td><td>55~59세</td></tr> <tr><td>4</td><td>15~19세</td><td>13</td><td>60~64세</td></tr> <tr><td>5</td><td>20~24세</td><td>14</td><td>65~69세</td></tr> <tr><td>6</td><td>25~29세</td><td>15</td><td>70~74세</td></tr> <tr><td>7</td><td>30~34세</td><td>16</td><td>75~79세</td></tr> <tr><td>8</td><td>35~39세</td><td>17</td><td>80~84세</td></tr> <tr><td>9</td><td>40~44세</td><td>18</td><td>85세+</td></tr> </tbody> </table>					그룹	연령대	그룹	연령대	1	0~4세	10	45~49세	2	5~9세	11	50~54세	3	10~14세	12	55~59세	4	15~19세	13	60~64세	5	20~24세	14	65~69세	6	25~29세	15	70~74세	7	30~34세	16	75~79세	8	35~39세	17	80~84세	9	40~44세	18	85세+			
그룹	연령대	그룹	연령대																																												
1	0~4세	10	45~49세																																												
2	5~9세	11	50~54세																																												
3	10~14세	12	55~59세																																												
4	15~19세	13	60~64세																																												
5	20~24세	14	65~69세																																												
6	25~29세	15	70~74세																																												
7	30~34세	16	75~79세																																												
8	35~39세	17	80~84세																																												
9	40~44세	18	85세+																																												
5	시도코드	SIDO	<ul style="list-style-type: none"> 해당 수진자 거주지의 시도코드 - 2012년부터 세종특별자치시가 신규로 편입됨에 따라, 2011년까지의 데이터에는 해당 항목이 존재하지 않음 <table border="1"> <thead> <tr> <th>코드명</th><th>시도명</th><th>코드명</th><th>시도명</th></tr> </thead> <tbody> <tr><td>11</td><td>서울특별시</td><td>42</td><td>강원도</td></tr> <tr><td>26</td><td>부산광역시</td><td>43</td><td>충청북도</td></tr> <tr><td>27</td><td>대구광역시</td><td>44</td><td>충청남도</td></tr> <tr><td>28</td><td>인천광역시</td><td>45</td><td>전라북도</td></tr> <tr><td>29</td><td>광주광역시</td><td>46</td><td>전라남도</td></tr> <tr><td>30</td><td>대전광역시</td><td>47</td><td>경상북도</td></tr> <tr><td>31</td><td>울산광역시</td><td>48</td><td>경상남도</td></tr> <tr><td>36</td><td>세종특별자치시</td><td>49</td><td>제주특별자치도</td></tr> <tr><td>41</td><td>경기도</td><td></td><td></td></tr> </tbody> </table>	코드명	시도명	코드명	시도명	11	서울특별시	42	강원도	26	부산광역시	43	충청북도	27	대구광역시	44	충청남도	28	인천광역시	45	전라북도	29	광주광역시	46	전라남도	30	대전광역시	47	경상북도	31	울산광역시	48	경상남도	36	세종특별자치시	49	제주특별자치도	41	경기도			N	26 ●		
코드명	시도명	코드명	시도명																																												
11	서울특별시	42	강원도																																												
26	부산광역시	43	충청북도																																												
27	대구광역시	44	충청남도																																												
28	인천광역시	45	전라북도																																												
29	광주광역시	46	전라남도																																												
30	대전광역시	47	경상북도																																												
31	울산광역시	48	경상남도																																												
36	세종특별자치시	49	제주특별자치도																																												
41	경기도																																														
6	신장(5CM 단위)	HEIGHT	<ul style="list-style-type: none"> 검진자의 키(5CM 단위) <ul style="list-style-type: none"> ❖ 예) 100~104CM → 100CM 																																												
7	체중(5KG 단위)	WEIGHT	<ul style="list-style-type: none"> 검진자의 몸무게(5KG 단위) <ul style="list-style-type: none"> ❖ 예) 25~29KG → 25KG 	N/Kg	45																																										
8	허리둘레	WAIST	<ul style="list-style-type: none"> 검진자의 허리둘레 <ul style="list-style-type: none"> ❖ 허리둘레 항목은 2008년부터 건강검진 문진항목으로 추가되었기 때문에 기준년도가 2002년부터 2007년까지인 경우 해당 항목 값이 결측 처리되어 제공됨 																																												
9	시력(좌)	SIGHT_L_EFT	<ul style="list-style-type: none"> 수검자의 좌측 눈의 시력 <ul style="list-style-type: none"> - 0.1~2.5 사이의 값으로 표기하며 0.1 이하의 시력은 0.1, 실명은 9.9로 표기 	N	0.5																																										
10	시력(우)	SIGHT_R_Ight	<ul style="list-style-type: none"> 수검자의 우측 눈의 시력 <ul style="list-style-type: none"> - 0.1~2.5 사이의 값으로 표기하며 0.1 이하의 시력은 0.1, 실명은 9.9로 표기 																																												
11	청력(좌)	HEAR_L	<ul style="list-style-type: none"> 수검자의 좌측 귀의 청력 	N	1																																										

◎ 요단백 판정 기준?

- 성인인 경우 하루 500mg 이상, 소아는 1시간 동안 체포면적 1제곱면적 당 4mg이상의 단백이 배설될 경우 명백한 단백뇨, 이보다 적은 경우(하루 30~300mg)의 단백이 배설되는 경우 미세단백뇨
- 시험지검사법(dipstick method)으로 시험지에 소변을 적신 후 60초 이내에 초록색으로 변색하는 정도로 판정하여 음성(-), 약양성(±), 30mg/dL은 +1, 100mg/dL은 경우 +2, 300mg/dL은 +3, 1000mg/dL은 +4로 판정

◎ 시력 측정 방식?

- 물체의 형태나 그 존재를 구분하는 눈의 기능을 형태각이라 하고 그 정도를 나타내는 것이 시력,
- 2점을 2점으로서 식별할 수 있는 최초 2점간 거리를 측정하고 이것을 시각으로 표시하고 그 역수에 비례하는 양으로 표시

◎ 청력이란?

- 소리의 강도에 대한 청각의 감수성 뜻이며 데시벨(dB)로 표시됨
- 정상인은 0dB, 난청자일수록 그수가 커짐

NO	제공항목			속성정보		비고
	표준 항목명	영문명	설명	표현형식/단위	예시	
		EFT	- 1(정상), 2(비정상)			
12	청력(우)	HEAR_RI_GHT	<ul style="list-style-type: none"> 수검자의 우측 귀의 청력 - 1(정상), 2(비정상) 	N	1	
13	수축기 혈압	BP_HIGH	<ul style="list-style-type: none"> 검진자의 최고 혈압으로 심장이 수축해서 강한 힘으로 혈액을 동맥에 보낼 때의 혈관 내압 	N/m mHg	140	
14	이완기 혈압	BP_LWST	<ul style="list-style-type: none"> 검진자의 최저 혈압으로 심장의 완기시의 혈압 	N/m mHg	81	
15	식전혈당 (공복혈당)	BLDS	<ul style="list-style-type: none"> 검진자 식사 전 혈당(혈액 100mL당 함유 되어 있는 포도당의 농도) 수치 	N/mg /dL	94	
16	총 콜레스테롤	TOT_CHOLE	<ul style="list-style-type: none"> 혈청 중의 에스텔형, 비에스테형(유리)콜레스테롤의 합 - 정상치는 150~250mg/dL 약 1/3이 비에스텔형(유리)콜레스테롤이며 나머지가 콜레스테롤에스테르 	N/mg /dL	164	
17	트리글리세라이드	TRIGLYCERIDE	<ul style="list-style-type: none"> 단순지질 혹은 중성지질을 뜻함 - 글리세롤에 3분자 지방산이 에스테르 합한 것으로서 자연계에서 찾아낼 수 있는 지방유도체 가운데 가장 분포가 넓음 - 정상치는 30~135mg/dL(0.34~1.52-mmol/l) ❖ 트리글리세라이드항목은 2008년부터 건강검진 문진항목으로 추가되었기 때문에, 기준년도가 2002년부터 2007년까지 인 경우 해당 항목 값이 결측 처리되어 제공됨 	N/mg/dL	94	
18	HDL 콜레스테롤	HDL_CHOLE	<ul style="list-style-type: none"> HDL(고밀도 리포단백질)에 포함되는 콜레스테롤 - 작은 입자의 콜레스테롤로 세포에 이끌려간 콜레스테롤을 간으로 돌려주고 혈관 벽에 쌓인 나쁜 콜레스테롤을 없애는 역할을 하는 성분 - 정상치는 30~65mg/dL ❖ HDL콜레스테롤 항목은 2008년부터 건강검진 문진항목으로 추가되었기 때문에, 기준년도가 2002년부터 2007년까지 인 경우 해당 항목 값이 결측 처리되어 제공됨 	N/mg/dL	45	
19	LDL 콜레스테롤	LDL_CHOLE	<ul style="list-style-type: none"> LDL(저밀도 리포단백질)에 함유된 콜레스테롤 - 입자가 매우 큰 콜레스테롤로 양이 과도하게 증가할 경우, 혈관벽에 쌓여서 동맥경화나 각종 질병을 야기 하는 성분 - 170mg/dL 이상일 경우 일반적으로 고LDL혈증으로 봄 ❖ LDL콜레스테롤 항목은 2008년부터 건강검진 문진항목으로 추가되었기 때문에, 기준년도가 2002년부터 2007년까지 인 경우 해당 항목 값이 결측 처리되어 제공됨 	N/mg/dL	예시 150 예시 2130	
20	혈색소	HMG	<ul style="list-style-type: none"> 혈액이나 혈구 속에 존재하는 색소단백으로 글로빈(globin)과 엠(heme)으로 구성되며 혈중의 산소운반체로서의 역할 수행 	N/g/dL	15.38 43.5	
21	요단백	OLIG_POTE_CD	<ul style="list-style-type: none"> 소변에 단백질이 섞여 나오는 것 - 1(-), 2(±), 3(+1), 4(+2), 5(+3), 6(+4)로 표기됨 	N	1	
22	혈청크레아티닌	CREATININE	<ul style="list-style-type: none"> 크레이티닌은 크레이틴의 탈수물로 내인성 단백대사의 종말산물로서 신장에서 배설되고 그 증감은 음식물에 관계없이 근육의 발육과 운동에 관계함 	N/mg/dL	0.8	

NO	제공항목			속성정보		비고
	표준 항목명	영문명	설명	표현형식/단위	예시	
			<ul style="list-style-type: none"> - 혈청크레아티닌 농도는 신기능장애에 의해 증량함 - 정상치 0.8~1.7mg/dL ❖ 혈청크레아티닌 항목은 2008년부터 건강검진 문진항목으로 추가되었기 때문에, 기준년도가 2002년부터 2007년까지 인 경우 해당 항목 값이 결측 처리되어 제공됨 			
23	(혈청) AST	SGOT_A ST	<ul style="list-style-type: none"> • 간 기능을 나타내는 혈액검사상의 수치, 간세포 이외에 심장, 신장, 뇌, 근육 등에도 존재하는 효소로 이러한 세포들이 손상을 받는 경우 농도가 증가함 - 정상치 0~40IU/L 	N/mg/dL	12	
24	(혈청) ALT	SGPT_A LT	<ul style="list-style-type: none"> • 간 기능을 나타내는 혈액검사상의 수치, ALT는 주로 간세포 안에 존재하는 효소로, 간세포가 손상을 받는 경우 농도가 증가함 - 정상치 0~40IU/L 	N/U/L	15	
25	감마 지티피	GAMMA_GTP	<ul style="list-style-type: none"> • 간 기능을 나타내는 혈액검사상의 수치, 간 내의 쓸개관(담관)에 존재하는 효소로 글루타민산을 외부에 펩티드나 아미노산 등으로 옮기는 작용을 함, 쓸개즙(담즙) 배설 장애, 간세포 장애 발생 시 혈중에 증가하게 됨 - 정상치 남성 11~63IU/L, 여성 8~35IU/L 	N/U/L	예시 119 예시 2114	
26	흡연상태	SMK_ST AT_TYPE_CD	<ul style="list-style-type: none"> • 해당 수검자의 흡연 상태 여부 <ul style="list-style-type: none"> - 0(피우지 않는다), 1(이전에 피웠으나 끊었다), 3(현재도 피우고 있다) 	N	1	
27	음주여부	DRK_YN	<ul style="list-style-type: none"> • 해당 수검자의 음주 상태 여부 <ul style="list-style-type: none"> - 0(마시지 않는다), 1(마신다) - N(마시지 않는다), Y(마신다) 	N	1	
28	구강검진 수검여부	HCHK_O E_INSPE C_YN	<ul style="list-style-type: none"> • 해당 검진자가 구강검진을 선택하여 검진하였는지 여부에 대한 항목 <ul style="list-style-type: none"> - 0(미수검), 1(수검) - N(미수검), Y(수검) 	N	1	
29	치아우식증유무	CRS_YN	<ul style="list-style-type: none"> • 해당 수검자의 치아우식증 유무에 대한 항목 <ul style="list-style-type: none"> - 0(없음), 1(있음) 	N	1	
30	결손치 유무	TTH_MS S_YN	<ul style="list-style-type: none"> • 해당 수검자의 결손치(다양한 원인(우식치주병, 발치 등)으로 인하여 치열에서 탈락한 치아) 존재 유무에 대한 항목 <ul style="list-style-type: none"> - 0(없음), 1(있음) 	N	1	2002~2013 에만 있음
31	치아마모증유무	ODT_TR B_YN	<ul style="list-style-type: none"> • 해당 수검자의 치아마모증 유무에 대한 항목 <ul style="list-style-type: none"> - 0(없음), 1(있음) 	N	1	
32	제3대구치 (사랑니) 이상	WSDM_DIS_YN	<ul style="list-style-type: none"> • 해당 수검자의 제3대구치(사랑니)에 대한 이상 유무 <ul style="list-style-type: none"> - 0(없음), 1(있음) 	N	1	
33	치석	TTR_YN	<ul style="list-style-type: none"> • 해당 수검자의 치석 여부 <ul style="list-style-type: none"> - 0(없음), 1(있음) - N(없음), Y(있음) 	N	1	
34	데이터	DATA_S	<ul style="list-style-type: none"> • 데이터 작성 기준일자 	YYYY	2016	

◎ 치아우식증이란?
- 일반적으로 충치로 알려져 있으며, 원인은 섭취한 음식물을 구강 내 미생물들이 섭취하고 배설하여 산을 만들고, 그 산들이 치아의 성분인 칼슘을 부식 시키는 질환

◎ 치아마모증이란?
- 치근 부위의 잇몸이 내려가 노출된 치아표면이 마모된 것으로 잇몸이 내려앉아 시멘트질 안의 상아질이 노출된 상태

◎ 치석이란?
- 치면세균막이 타액과 치은열구(치아와 잇몸 사이의 공간) 내의 칼슘, 인 등의 무기질이 침착되어 만들어진 것

NO	제공항목			속성정보		비고
	표준 항목명	영문명	설명	표현형식/단위	예시	
	공개일자	TD_DT		MM DD	1220	

* 비고란에 '●' 표시가 되어 있는 제공항목의 경우 개별 코드 값의 명칭 제공 시 데이터량의 증가에 따른 제공 서버의 부하 및 다운로드 속도 저하 등의 문제가 발생하여 코드 값만 제공함, 해당 코드의 명칭은 항목 설명을 참조하거나, 관련된 웹사이트를 통하여 확인이 가능함

다. 건강검진정보 이용 시 유의사항

- 행정구역이 연도별로 각각 코드화되어 있으므로, 변경사항을 충분히 검토하여 활용하여야 한다. 예컨대 세종특별자치시의 경우 2012년부터 신규로 편입됨에 따라, 2011년까지의 데이터에는 해당 항목이 존재하지 않는다.
- 건강검진정보는 해당년도에 수검자에 대한 1차 일반건강검진 및 생애전환기건강진단(2008년~) 주요 결과 및 문진응답 자료로서 의료급여수급권자의 일반건강검진 자료는 미포함 된다.
- 2009년 건강검진제도 개편으로 인해 주요 검진 항목이 변경된바 연도별 건강검진 제도, 건강검진 항목, 문진 항목 및 유형, 변수명 등의 변경사항을 충분히 검토하여 활용하여야 한다.

구분	2002~2007년	2008년	2009~2015년
신장			
체중			
허리둘레			
수축기혈압			
이완기혈압			
식전혈당(공복혈당)			
총콜레스테롤			
트리글리세라이드 (총성지방)			
HDL콜레스테롤			
LDL콜레스테롤			
혈색소			
요단백			
혈청크레아티닌			
(혈청지오티)AST			
(혈청지오티)ALT			
감마지티피			

* 트리글리세라이드,
HDL콜레스테롤,
LDL콜레스테롤,
혈청크레아티닌 자료는
2008년 생애부터 수록
* 2012년부터
LDL콜레스테롤의 계산
값과 측정치 값이 통합됨.
총콜레스테롤,
HDL콜레스테롤,
트리글리세라이드 수치로
계산하되,
트리글리세라이드
측정값이 400mg/dl
이상인 경우 실측정한
값이 들어감

제4절 국가중점 개방데이터 다운로드 방법

공공데이터 개방서비스(<http://nhiss.nhis.or.kr/op/it/index.do>)에 접속하여 편리하게 국가중점 개방데이터를 다운로드 받을 수 있다.

1. 공공데이터 개방서비스 활용 방법 소개

- 1단계 : 국가중점 개방데이터 개방을 위한 웹사이트에 접속하여 첫 화면에서 「다운로드」메뉴 선택

The screenshot shows the NHISS Open Data Service homepage. At the top, there is a navigation bar with tabs: 소개 (Introduction), 공지사항 (Announcements), **다운로드** (Download) [highlighted with a red dashed box], 활용가이드 (Usage Guide), and 활용사례 (Case Studies). Below the navigation bar, there is a large banner for 'Open Data Service'. To the right of the banner, there are two boxes: 'Download' and 'Usage Examples'. The 'Download' box contains text about downloading data related to medical insurance, pharmaceuticals, and health checks. The 'Usage Examples' box shows icons of a computer monitor and a magnifying glass. At the bottom of the page, there are sections for 'Announcements' and 'Usage Guide' with a list of items and their dates.

공공데이터 개방서비스 첫 화면에서 다운로드 메뉴 선택

- 2단계 : 진료내역, 의약품처방, 건강검진 정보 중 다운로드를 원하는 정보와 연도 선택

The screenshot shows the 'Medical Record Information' download selection screen. At the top, there is a navigation bar with tabs: 소개 (Introduction), 공지사항 (Announcements), **다운로드** (Download) [highlighted with a red dashed box], 활용가이드 (Usage Guide), and 활용사례 (Case Studies). Below the navigation bar, there is a section titled '1 다운로드' with a list of categories: 진료내역정보, 의약품처방정보, and 건강검진정보. A green arrow points to the '진료내역정보' button. Below this, there is a section titled '2 연도선택' with a list of years from 2002 to 2013. A green arrow points to the '2008' button. At the bottom, there is a '3 다음단계' (Next Step) button. A blue box at the bottom indicates that the user has selected the information and year for download.

다운로드를 원하는 정보와 연도를 선택하고 다음단계 버튼을 클릭함

3) 3단계 : 연령대, 성별, 사용자 거주 지역(시도 단위), 사용목적 등 간략한 사용자 정보 입력(국가중점 개방데이터 사용자의 주된 사용자층과 이용목적 등을 수집하여 향후 데이터 제공과 관련된 정책 수립에 반영하고자 함)

● 홈 > 다운로드 > **진료내역정보**

원하는 데이터를 선택한 후 다음단계를 누르시면 간단한 설문 후 다운로드 받으실 수 있습니다.

진료내역정보 → 설문작성 → 다운로드

연령	<input type="radio"/> 10대	<input type="radio"/> 20대	<input checked="" type="radio"/> 30대	<input type="radio"/> 40대	<input type="radio"/> 50대	<input type="radio"/> 60대 이상	
성별	<input type="radio"/> 남성	<input checked="" type="radio"/> 여성	• 연구목적 • 상품개발 • 기타				
지역	경상북도						
사용목적	선택						

이전단계 | 다음단계 ➤

연령, 성별, 지역, 사용목적 등 간단한 설문작성

4) 4단계 : 최종 다운로드 화면으로 이동하여 첨부파일을 다운로드 함, 첨부파일에는 사용자가 선택한 연도의 데이터 및 사용자 매뉴얼이 포함되어 있음

● 홈 > 다운로드 > **진료내역정보**

원하는 데이터를 선택한 후 다음단계를 누르시면 간단한 설문 후 다운로드 받으실 수 있습니다.

진료내역정보 → 설문작성 → **다운로드**

연도	2010년도
첨부파일	notice.txt

다운로드 | **처음으로**

첨부파일 내 선택한 연도의 데이터 및 사용자 매뉴얼이 포함됨

제5절 사용자 지원 사항

1. 국가중점 개방데이터 활용가이드 및 사례 제공

국가중점 개방데이터 개방을 위한 웹사이트를 통해 사용자가 자주 문의하는 사항들을 정리하여 활용 가이드로 제공한다. 그 밖에 국가중점 개방데이터를 활용하여 개발된 국내·외 다양한 사례를 제공함으로서 국가중점 개방데이터를 활용하고자 하는 국민들의 효과적인 데이터 활용을 지원한다.

3. 국가중점 개방데이터 개방을 위한 웹사이트를 활용한 온라인 문의기능 지원

국가중점 개방데이터 개방을 위한 웹사이트 내에 ‘문의하기’ 기능을 통해 국가중점 개방데이터 이용과 관련한 온라인 문의가 가능하다. 그 외 데이터 이용과 관련된 건의사항이 있거나 오류를 발견한 경우 관리자에게 신고 할 수 있도록 하는 기능을 제공하여 사용자와의 효과적인 소통이 가능케 한다.

4. 국가중점 개방데이터 다운로드 전에 데모자료 제공기능 지원

국가중점 개방데이터 다운로드 전에 사용자가 해당 정보의 데모자료를 확인할 수 있도록 함으로써 이용자가 자신이 필요로 하는 정보가 국가중점 개방데이터 내에 포함되어 있는지를 사전에 확인할 수 있어 사용자 편의성이 증가된다.

Q&A 묻고 답하기

1. 국가중점 개방데이터의 활용하고자 할 때 별도의 신청 및 심사 절차를 거쳐야 하나?

⇒ 아니다. 사용자 매뉴얼에 따라 국민건강보험자료 공유서비스(NHISS) 웹페이지에 접속하여 자유롭게 다운로드 받을 수 있으며 별도의 신청 및 심사 절차를 거칠 필요가 없다.

2. 국가중점 개방데이터를 활용하여 의료서비스 관련 어플이나 사이트 등을 개설하는 등 상업적인 용도로 활용할 수 있는가?

⇒ 가능하다. 국가중점 개방데이터는 민간이 자유롭게 활용할 수 있는 공공데이터로서 개별 법령(의료법 등)에 저촉되지 않는 범위 내에서 다양한 용도로 활용이 가능하다.

3. 국가중점 개방데이터의 활용 가능 범위가 궁금하다. 학술·연구용으로도 활용 가능한가?

⇒ 학술·연구용으로 활용하기 위해서는 다량의 표본에 대한 건강관련 데이터가 제공되어야 한다. 그러나 국가중점 개방데이터는 웹을 통해 정보가 제공되기 때문에 많은 양의 데이터를 제공하는 것이 불가능하다. 구축과정에서도 이러한 이유 때문에 데이터의 대한 부분적 편집이나 가공이 이루어졌다. 또한 일반국민을 대상으로 제공되는 데이터이기 때문에 학술 연구자들에 한하여 심사를 거쳐 제공되는 데이터에 비해 개방항목의 수도 적다. 결론적으로 국민건강보험공단이 학술·연구용으로 구축하여 제공하고 있는 표본연구DB 등을 활용하는 것이 더욱 효과적일 것이라 생각된다.