

국민건강정보데이터 건강검진정보 사용자 메뉴얼

I. 국민건강정보데이터

□ 데이터 개방 개요

- (개념) 국민건강보험공단이 업무를 수행함에 따라 축적된 정보 중 활용도가 높은 진료내역정보, 의약품처방정보, 건강검진정보를 공공데이터 포털을 통해 국민 모두에게 개방
- (목적) ❶법률에 근거한 국민의 공공데이터 이용권 보장 및 알권리 충족, ❷보건의료 산업계에 새로운 형태의 사업기회 창출, ❸개인중심의 건강 정보 개방으로 의료기관의 고품질 의료서비스 제공 유도

□ 데이터 생성 형태

- (범위) 2002년부터 1년 단위로 국민건강보험 가입자 중 1백만 명을 무작위 선별하여 진료내역, 의약품처방내역, 건강검진결과에 대한 데이터 생성
- (갱신주기) 매년 하반기에 전년도 데이터를 구축·개방
- (형태) 데이터 자료 구조(포맷)는 공공데이터 포털에서 권장하는 CSV (Comma Separated Value)이며 연도별 분할된 형태로 제공

□ 안전한 데이터 개방

- (개인식별정보 제외) 개인을 식별할 수 있는 개인식별정보* 등은 국가 중점 개방데이터 항목에서 제외

* 고유식별정보(주민번호, 국민건강보험자격번호 등), 쉽게 개인을 식별할 수 있는 정보(이름, 전화번호, 주소, 생년월일, 사진 등)

- (개인정보 비식별화) 타 정보와 결합을 통해 개인 식별이 가능한 정보가 포함되어 있는지 여부를 사전 필터링하여 재식별 가능성이 존재할 경우 개방 대상에서 제외하거나 개별 항목에 적합한 비식별화 처리기법을 적용

표1. 국민건강정보데이터 설명

구 분	내 용
표본의 추출	• 각 연도별 수진 및 건강검진 수검 환자 1백만 명 무작위 추출
자료별 조합제한	• 동 자료는 단면자료로서 국민건강정보데이터별 개인일련번호와 청구일련번호를 다르게 하여 제공(연계 불가능)
개인식별자 제거	• 개인식별번호(8자리) → 개인일련번호(8자리)
범주화	• 연령 그룹화(5세 단위) - 85세 이상은 '85세 이상'으로 범주화
데이터 마스킹	• 인권침해 우려가 큰 질환 및 법정감염병 등 민감상병 3,086개 상병코드 대분류화(예:D_) • 주성분코드는 주성분 4자리와 투여경로 1자리를 합한 5자리로 제공하며, 주성분별 제약사가 1곳인 단독등재약품 주성분코드를 ATC코드로 변환(2022년부터 적용)
최상위 지역코드 제공	• 소규모 지역에 거주하는 표본의 인식을 고려하여 시도코드에 한 하여 제공(17개 시도단위 사용)

II. 건강검진정보

□ 개념 및 구성

- (개념) 20세 이상의 국민건강보험 직장가입자와 피부양자, 지역가입자 세대주 및 세대원 약 1백만 명의 일반건강검진 및 구강검진 결과로 구성
- (구성) 총 33개 변수로 가입자 일련번호와 수진자 기본정보*, 건강검진 결과 및 문진정보**로 구성

* 성, 연령, 거주지 시도코드와 같은 기본정보

** 신체, 몸무게, 허리둘레 등 신체사이즈 정보와 혈압, 혈당, 콜레스테롤, 요단백, 감마지티피와 같은 병리검사결과, 시력과 청력, 구강검사와 같은 진단검사결과, 그 외 음주와 흡연 여부에 대한 문진결과

표2. 건강검진정보 개방 항목

No.	항목명	설명	표현형식 /단위	예시																																																																								
1	기준년도	• 해당 정보의 기준년도	YYYY	2009																																																																								
2	가입자 일련번호	• 해당가입자에 부여한 일련번호(1 ~ 1,000,000)	N	1																																																																								
3	성별코드	• 해당 정보 대상자의 성별 - 성별 : 1(남자), 2(여자)	N	1																																																																								
4	연령대코드 (5세단위)	• 기준년도에 수진자의 나이를 5세 단위로 그룹화 (범주화)하여 구분한 코드 - 5세 단위 그룹화, 85세 이상은 85+로 그룹화 - 2002~2013년 까지 <table><tr><th>그룹</th><th>연령대</th><th>그룹</th><th>연령대</th></tr><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></table> - 2014년 이후 <table><tr><th>그룹</th><th>연령대</th><th>그룹</th><th>연령대</th></tr><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></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+	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+	N	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+																																																																									
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	• 해당 수진자 거주지의 시도코드 - 2012년부터 세종특별자치시가 신규로 편입됨에 따라, 2011년까지의 데이터에는 해당 항목이 존재하지 않음 <table><tr><th>코드</th><th>시도명</th><th>코드</th><th>시도명</th></tr><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></table>	코드	시도명	코드	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																																
코드	시도명	코드	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	-	-																																																																									

No.	항목명	설명	표현형식 /단위	예시
6	신장 (5cm단위)	• 검진자의 키(5cm 단위) 예) 100~104cm → 100cm	N/cm	140
7	체중 (5kg단위)	• 검진자의 몸무게(5kg 단위) 예) 25~29kg → 25kg	N/kg	45
8	허리둘레	• 검진자의 허리둘레 - 허리둘레 항목은 2008년부터 건강검진 문진항목으로 추가되었기 때문에 기준년도가 2002년~2007년까지인 경우 해당 항목 값이 결측 처리되어 제공됨	N/cm	82
9	시력(좌)	• 수검자의 좌측 눈의 시력 - 0.1~2.5 사이의 값으로 표기하며 0.1 이하의 시력은 0.1, 실명은 9.9로 표기	N	0.5
10	시력(우)	• 수검자의 우측 눈의 시력 - 0.1~2.5 사이의 값으로 표기하며 0.1 이하의 시력은 0.1, 실명은 9.9로 표기	N	0.5
11	청력(좌)	• 수검자의 좌측 귀의 청력 - 1(정상), 2(비정상)	N	1
12	청력(우)	• 수검자의 우측 귀의 청력 - 1(정상), 2(비정상)	N	1
13	수축기 혈압	• 검진자의 최고 혈압으로 심장이 수축해서 강한 힘으로 혈액을 동맥에 보낼 때의 혈관 내압	N/mmHg	140
14	이완기 혈압	• 검진자의 최저 혈압으로 심장 이완기 시의 혈압	N/mmHg	81
15	식전혈당 (공복혈당)	• 검진자 식사 전 혈당(혈액 100ml당 함유되어 있는 포도당의 농도) 수치	N/mg/dL	94
16	총 콜레스테롤	• 혈청 중의 에스텔형, 비에스텔형(유리)콜레스테롤의 합 - 정상치는 150~250mg/dL 약 1/3이 비에스텔형(유리) 콜레스테롤이며 나머지가 에스텔형	N/mg/dL	164
17	트리글리세 라이드	• 단순지질 혹은 중성지질을 뜻함 - 글리세롤에 3분자 지방산이 에스테르 결합한 것으로서 자연계에서 찾아낼 수 있는 지방유도체 가운데 가장 분포가 넓은 - 정상치는 30~135mg/dL(0.34~1.52-mmol/l) - 트리글리세라이드항목은 2008년부터 건강검진 문진 항목으로 추가되었기 때문에, 기준년도가 2002년~2007년 까지인 경우 해당 항목 값이 결측 처리되어 제공됨	N/mg/dL	94

No.	항목명	설명	표현형식 /단위	예시
18	HDL 콜레스테롤	<ul style="list-style-type: none"> HDL(고밀도 리포단백질)에 포함되는 콜레스테롤 <ul style="list-style-type: none"> 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, 기준연도가 2002년~2007년까지인 경우 해당 항목 값이 결측 처리되어 제공됨 	N/mg/dL	45
19	LDL 콜레스테롤	<ul style="list-style-type: none"> LDL(저밀도 리포단백질)에 함유된 콜레스테롤 <ul style="list-style-type: none"> 입자가 매우 큰 콜레스테롤로 양이 과도하게 증가할 경우, 혈관벽에 쌓여서 동맥경화나 각종 질병을 야기하는 성분 170mg/dL 이상일 경우 일반적으로 고LDL혈증으로 봄 LDL콜레스테롤 항목은 2008년부터 건강검진 문진항목으로 추가되었기 때문에, 기준연도가 2002년~2007년까지인 경우 해당 항목 값이 결측 처리되어 제공됨 	N/mg/dL	(예시1) 50 (예시2) 130
20	혈색소	<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	15.398 4
21	요단백	<ul style="list-style-type: none"> 소변에 단백질이 섞여 나오는 것 <ul style="list-style-type: none"> 1(-), 2(±), 3(+1), 4(+2), 5(+3), 6(+4)로 표기됨 	N	1
22	혈청크레아티닌	<ul style="list-style-type: none"> 크레아티닌은 크레아틴의 탈수물로 내인성 단백대사의 종말산물로서 신장에서 배설되고 그 증감은 음식물에 관계없이 근육의 발육과 운동에 관계함 <ul style="list-style-type: none"> 혈청크레아티닌 농도는 신기능장애에 의해 증량함 정상치 0.8~1.7mg/dL 혈청크레아티닌 항목은 2008년부터 건강검진 문진항목으로 추가되었기 때문에, 기준연도가 2002년~2007년까지인 경우 해당 항목 값이 결측 처리되어 제공됨 	N/ mg/dL	0.8
23	혈청지오티(AST)	<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 <ul style="list-style-type: none"> 정상치 0~40IU/L 	N/IU/L	12

No.	항목명	설명	표현형식 /단위	예시
24	혈청지피티 (ALT)	<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/IU/L	15
25	감마 지티피	<ul style="list-style-type: none"> 간 기능을 나타내는 혈액검사상의 수치, 간 내의 쓸개관(담관)에 존재하는 효소로 글루타민산을 외부에 펌프질하거나 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/IU/L	(예시1) 19 (예시2) 114
26	흡연상태	<ul style="list-style-type: none"> 해당 수검자의 흡연 상태 여부 - 1(피우지 않는다), 2(이전에 피웠으나 끊었다), 3(현재도 피우고 있다) 	N	1
27	음주여부	<ul style="list-style-type: none"> 해당 수검자의 음주 상태 여부 - 0(마시지 않는다), 1(마신다) 	N	1
28	구강검진 수검여부	<ul style="list-style-type: none"> 해당 검진자가 구강검진을 선택하여 검진하였는지 여부에 대한 항목 - 0(미수검), 1(수검) 	N	1
29	치아우식증 유무	<ul style="list-style-type: none"> 해당 수검자의 치아우식증 유무에 대한 항목 - 0(없음), 1(있음) 	N	1
30	결손치 유무	<ul style="list-style-type: none"> 해당 수검자의 결손치(다양한 원인(우식치주병, 발치 등)으로 인하여 치열에서 탈락한 치아) 존재 유무에 대한 항목 - 0(없음), 1(있음) ※ 2009년부터 검진항목 제외 	N	1
31	치아마모증 유무	<ul style="list-style-type: none"> 해당 수검자의 치아마모증 유무에 대한 항목 - 0(없음), 1(있음) ※ 2014년부터 검진항목 제외 	N	1
32	제3대구치 (사랑니)이상	<ul style="list-style-type: none"> 해당 수검자의 제3대구치(사랑니)에 대한 이상 유무 - 0(없음), 1(있음) ※ 2014년부터 검진항목 제외 	N	1
33	치석	<ul style="list-style-type: none"> 해당 수검자의 치석 여부 - 0(없음), 1(있음) 	N	1

□ 이용 시 유의사항

- 행정구역이 연도별로 각각 코드화되어 있어 변경사항의 충분한 검토 필요
 - 예를 들어, 세종특별자치시의 경우 2012년부터 신규로 편입됨에 따라, 2011년까지의 데이터에는 해당 항목이 존재하지 않음
- 건강검진정보는 해당년도에 수검자에 대한 1차 일반건강검진, 구강검진 및 생애전환기건강진단(2008년~) 주요 결과 및 문진응답 자료로서 2차 검진, 암검진 결과 등은 포함하지 않음
- 건강검진제도 개편 등으로 주요 검진항목이 자주 변경되므로 연도별 건강검진 ☐ 문진 항목 및 유형, 결과 기재방법 등 변경사항의 충분한 검토가 필요함 ... 「건강검진 실시기준(보건복지부 고시)」 참조