

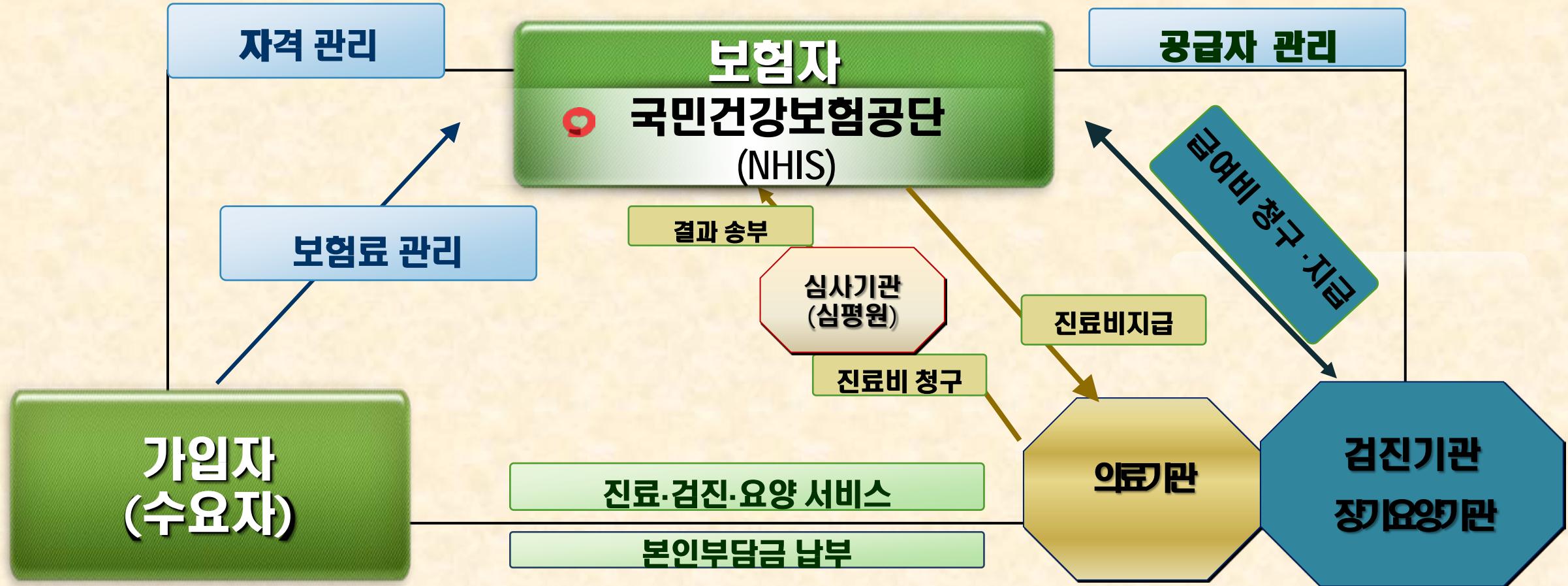
# 건강보험빅데이터 구성과 활용방안

2015. 2. 25.

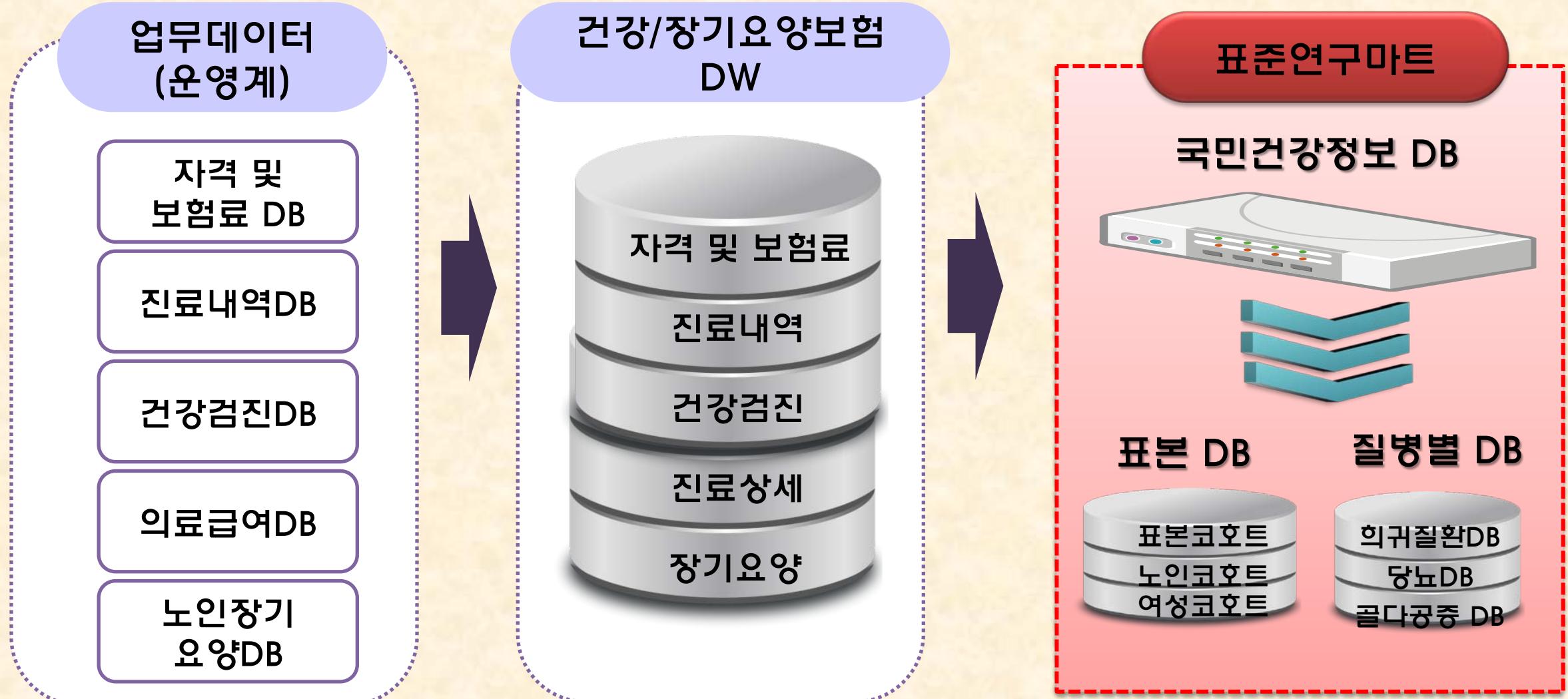
국민건강보험공단 빅데이터운영실 박종현

# 건강보험빅데이터 구성

# 건강보험 업무흐름도



# 데이터 수집체계



# 데이터 구성

## 자격 및 보험료 DB

- 생년월일, 성별, 나이, 거주지역, 사업자/가입자, 해외출입국자료 등
- 종합소득, 연금소득, 전월세, 재산세·자동차세 등 과세자료, 보수월액
- 국가유공자 명부, 장애인 등록자료(장애등급, 장애유형 등)

## 진료내역DB

- 명세서일반 : 요양기관, 상병명, 내원일수, 요양개시일, 진료과목 등
- 진료내역 : 병원 내의 처치 및 수술, 원내 약처방, 투여량, 진료비 등
- 수진자 상병내역 : 부상병을 추가한 상세
- 처방전교부 상세내역 : 원외 처방약, 투약량, 투여일수, 약물정보 등

## 건강검진DB

- 일반건강검진, 생애전환기 건강진단, 5대 암검진(위암/대장암/폐암/간암/유방암), 영유아 건강검진, 구강검진
- 문진자료(생활습관, 가족력, 기왕력 등)
- 검진결과 실측 데이터

## 의료급여DB

- 생년월일, 성별, 나이, 보장기관(시도·시군구), 의료급여 유형 및 거주 지역 등 자격 자료
- 요양기관, 상병명, 내원일수, 요양일수, 요양개시일, 진료비 등 명세서 내역

## 노인장기 요양DB

- 장기요양 신청내역, 노인상태상 등 인정조사자료, 등급판정자료 등
- 희망급여종류, 시설·재가급여, 특별현금급여, 급여제공기록지 등
- 의사소견서, 장기요양급여 청구자료, 심사자료, 기관 평가자료 등

# 분석가능변수: 인구사회학적 요인

성	남 / 녀	보험 종류	직장/지역/의료급여;
연령	수진 시점 / 연말 시점		가입자/피부양자
출생	출생일자	보험료	보험료 납부 금액
사망	사망일자		산정기준변수
장애	15개 장애유형 1~6 장애등급		보험료 20분위
	장애등록일자		피부양자 수 조정 보험료
지역	수진자 읍면동 단위 요양기관 읍면동 단위	사업장	직종/업종(대분류, 세분류)
	사업장 읍면동 단위		규모

# 분석가능변수: 요양기관 및 검진기관 요인

인력	요양기관인원현황	기본	검진기관기본현황
	요양기관관리의사현황		검진기관위탁항목
	진료과목별 전문의현황		검진기관검진유형
	물리치료 인력	인력	검진기관인력
시설현황	설립구분		검진기관인력교육
	요양기관종별		출장검진인력
	진료표방과목	장비	검진기관장비확인
	병실 및 병상 수		
	산재지정요양기관		
	물리치료기관 일반현황		
장비	요양기관 장비상세내역		

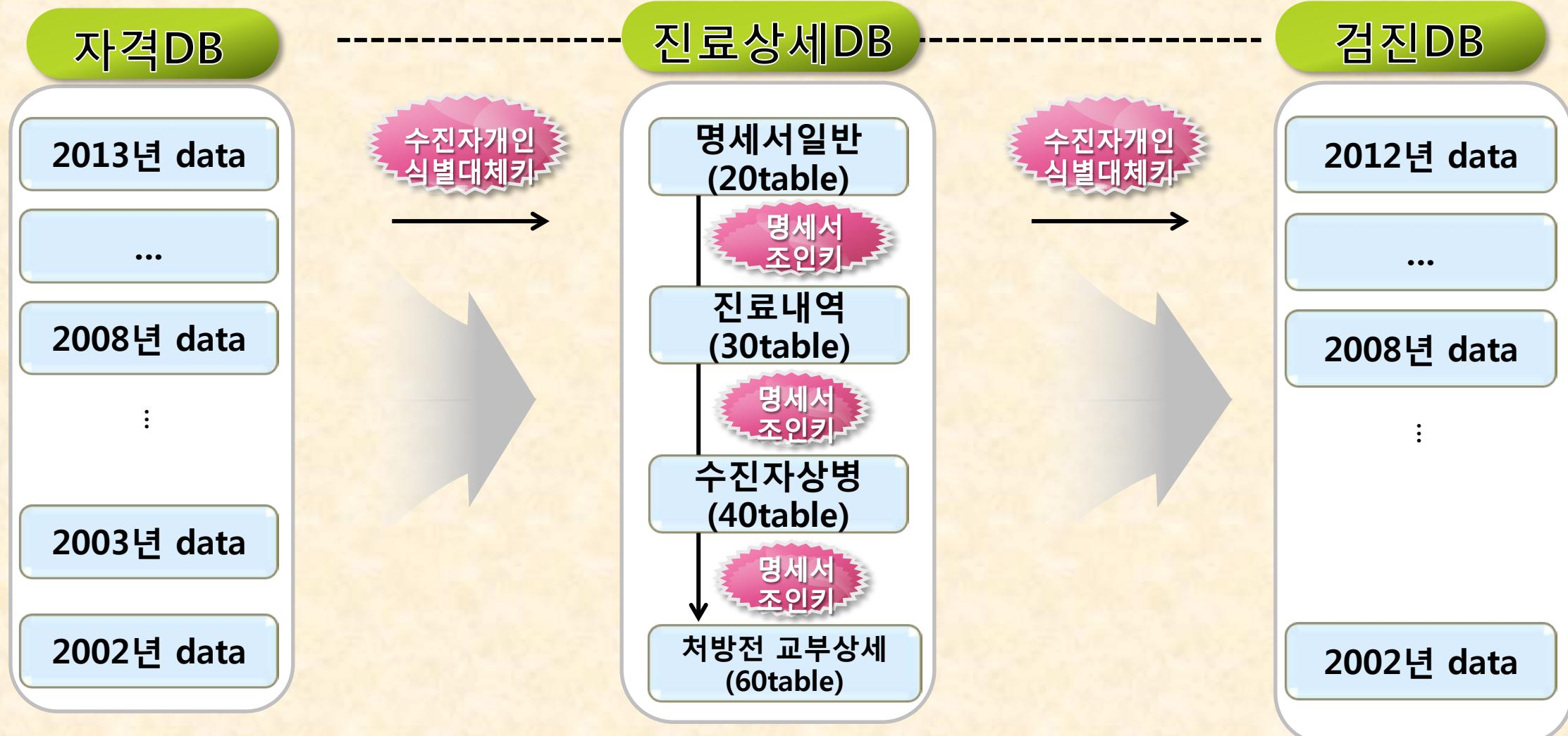
# 분석가능변수: 건강행태 및 LAB DATA

흡연	현재흡연여부	신체계측	키/몸무게/BMI/허리둘레
	흡연량/기간		시력/청력
음주	현재음주여부	혈압	수축기/이완기
	음주량/기간	혈당	공복혈당
신체활동	걷기/중등도/고강도	혈지질	총콜레스테롤/HDL/TG/LDL
과거력	고/당/지/심뇌/암 문진	간/신기능	간/신기능검사/B형간염(생애)
	고/당/지/심뇌/암 진료내역	골밀도	골밀도검사(생애)
가족력	고/당/지/심뇌/암 문진	암검사	내시경/Bx
인지기능	KDSQ		
우울감			
ADL			

# 분석가능변수: 질환 검사 및 치료행위

질환명	KCD6; 주상병 및 부상병
의료이용	입원 및 외래, 응급실 경유
	입내원일수/처방일수
	진료과목코드
	심결비용
상세내역	행위, 치료재료, 약제
	1회투약량/1일투약량
	총 투여일수/실시횟수

# 자료간 연계



# 건강보험빅데이터 활용

# 빅데이터 활용 로드맵

AS - IS

## 건보공단 플랫폼(폐쇄, 개방)

폐쇄형  
(식별)  
개방형  
(의명화)

건강보험 플랫폼  
(My Health Bank)  
국민건강정보 DB  
(표본 코호트 DB)

TO - BE

## 향후 추가

영상자료(CT...)

Life-log 자료

사망원인 자료

예방접종 자료

학생검진 자료

근로자특수검진

기상 자료

대기오염 자료

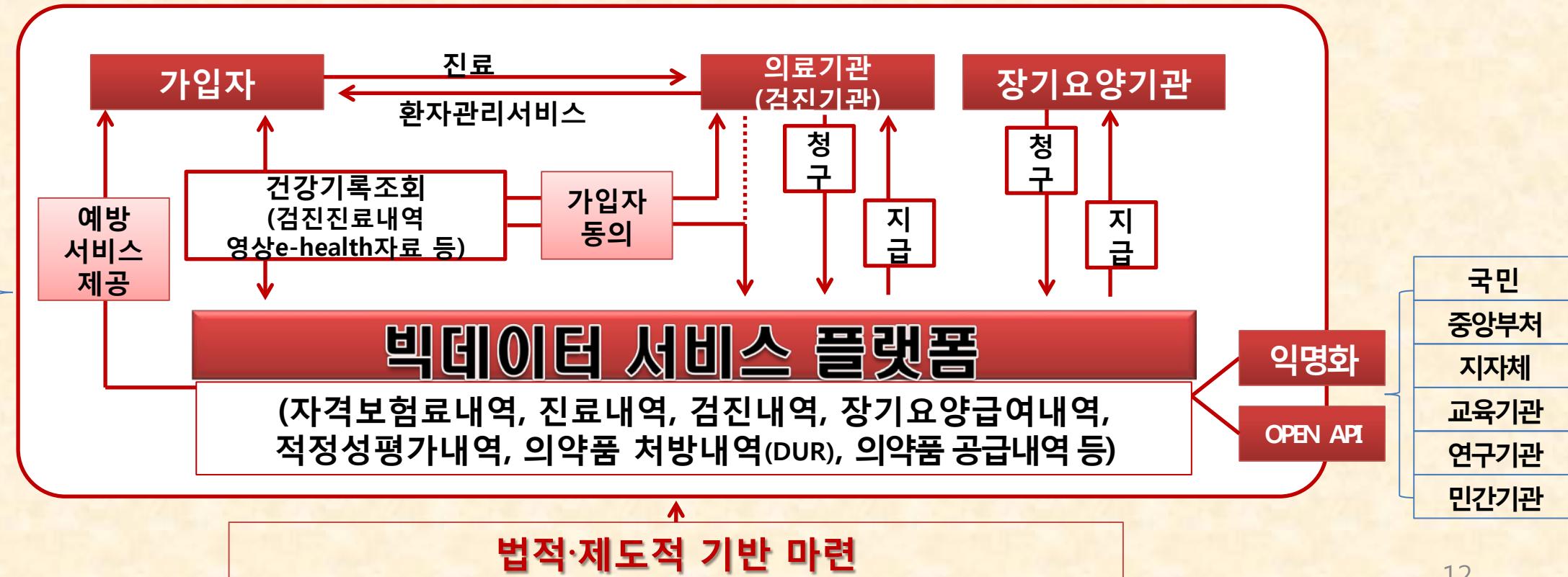
비정형 자료

## 데이터 분석

개인별 위험도  
예측모형 개발  
근거생산 연구

## 제공 서비스

검진결과 및 진료내역 확인  
맞춤형 예방 및 관리서비스 제공  
외부 공개를 통한 연구활성화(14년~)



# 학회 및 기관 연구 지원

문제의 크기, 질병부담, 발병 위험도 예측 연구, 인과관계 추론 모델링  
→ 건강보험 빅데이터를 보건의료분야의 연구 및 정책수립에 활용



- 국내 당뇨 관리 현황 및 질병부담



- 골절 위험도 예측모형 개발



- 재난 역학분야 연구  
- 희귀질환 역학조사,  
- 흡연 · 비만 등 건강영향평가



- 노인증후군 낙상예측 모형 개발



- 의약품 안전성 분석 및 평가 등



- 환경성 질병 예측모형 개발  
- 질환 발생 및 질병부담 평가



- 암 관리지표 및 암검진 효과  
- 발생 및 생존 관련 예측 모형  
- 암치료 패턴 및 질병부담



- 일차의료 서비스 질 평가



- 뇌졸중 중증도별 의료비 추산  
- 뇌졸중 환자 의료이용 경로분석  
- 뇌졸중 발생 위험요인 분석

# 학회 및 기관 연구 지원

## ■ 학회의 전문성 및 자료의 공익성을 반영한 연구

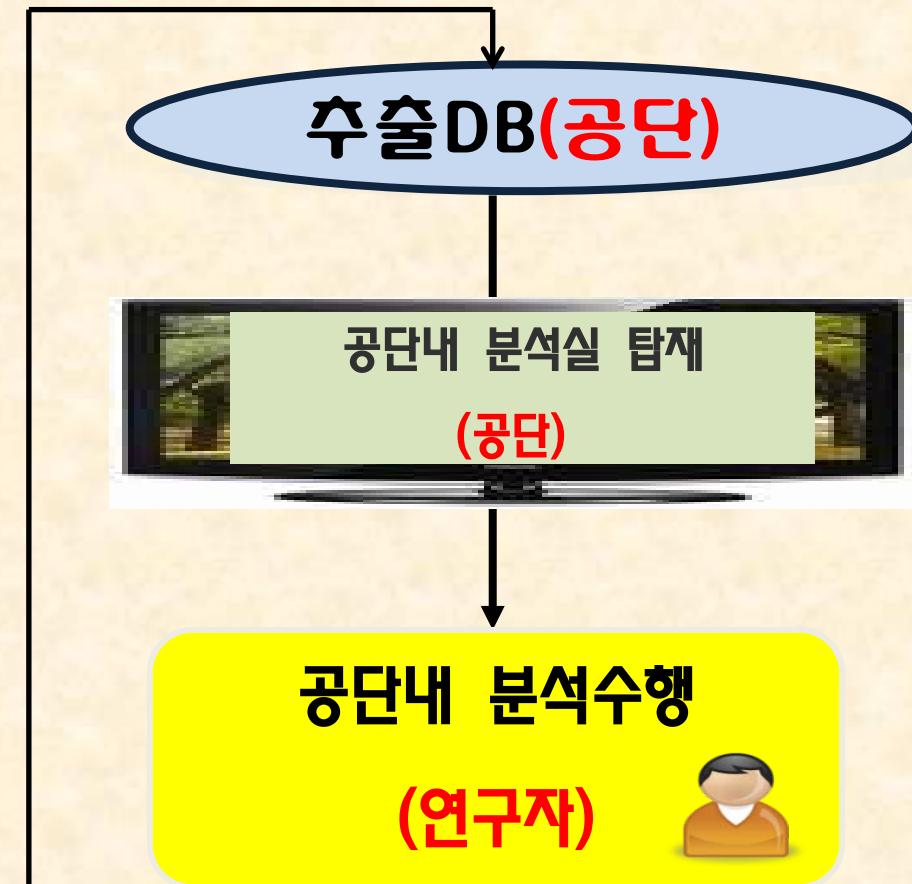
- › 대표성이 있는 학회 주관으로 해당 질환에 대한 국가 통계 생산
- › 사회적 이슈에 대한 건강영향 평가 (비만, 흡연 등)
- › 학회 및 회원들의 관심 주제 등

## ■ 청구자료의 품질과 활용 제고 연구

- › 해당 질병별로 청구자료 분석기준에 따라 통계값 차이 발생  
→ 이를 임상현장을 반영한 청구자료 분석 기준 마련(유병률 등)
- › 청구자료 활용의 가이드라인 마련 및 배포(워크숍, 매뉴얼 등)
- › 빅데이터 분석방법론 개발 : 전수자료의 p값 의미, 전수와 하위 비교
- › 질환별 DB 구축 및 활용을 위한 아이디어 제공

# 학회 및 기관 연구 지원

1.
  - 연구계획서 작성(주제 선정)
  - (학외)
2.
  - 공동연구 협약서 체결(MOU)
  - (공단-학외)
3.
  - 연구과제별 세부 수행개요서 작성
  - (학외)
4.
  - 자료구축
  - (공단- 학외)



# 학회 및 기관 연구 지원



- 원칙
  - 개인정보보호
  - 소량 데이터셋 생성
- Data summarizing
  - 분석대상자별 1개의 line으로 청구내역이 계산된 가상의 0/1 코드 혹은 sum 값이 병렬된 데이터셋
  - 자격, 검진, 진료내역이 연계된 연도별 1개의 데이터셋

# 학회 및 기관 연구 지원

- 유병
  - 1년간 치료 혹은 상병 혹은 검진검사기준 이상 여부
  - 0/1 coding
- 질병 burden
  - 1년간 입원일수 합, 외래 방문건수 합, 투약일수 합
  - 1년간 진료비 합,
  - 대상자별 Sum 변수
- 치료순응 여부
  - 외래방문 규칙성 혹은 투약일수 합 기준 이상
  - 연속 혹은 0/1 coding
- 조기치료 여부
  - 검진 이후 첫 의료기관 방문일수 차이
  - 연속 혹은 0/1 coding
- 사망 여부
  - 1년간 0/1 coding

# 학회 및 기관 연구 지원

RECU_FR_DT	MAIN_SICK	SUB_SICK	RECN	OFFC_INJ_TYPE	VSCN	MDC_DD_CNT	TOT_PRES_DD_CNT	EDEC_TRAMT	EDEC_SBRDN_AMT	EDEC_JBRDN_AMT	IDV_ID
20120122	R1049	K529		1 0	1	0	3	12890	3800	9090	44
20120103	I109	G441		1 0	1	0	30	11870	4100	7770	124
20120128	J0391	M79188		1 0	1	0	2	11880	4100	7780	124
20120209	J0390			1 0	1	0	3	10780	1500	9280	124
20120228	K140	R093		1 0	1	0	7	9210	1500	7710	124
20120202	K291	G441		1 0	1	0	30	11900	4100	7800	124
20120222	K140	R093		1 0	1	0	5	12890	1500	11390	124
20120219	R509	N10		1 0	1	0	0	564420	346600	217820	124
20120328	M5422	R51		1 0	1	0	5	14500	1500	13000	124
20120424	M4896	I100		1 0	1	0	30	12470	1500	10970	124
20120303	K140	R093		1 0	1	0	4	11250	1500	9750	124
20120419	M4896	K219		1 0	1	0	3	9210	1500	7710	124
20120416	J0390			1 0	1	0	3	14370	1500	12870	124
20120405	K291	M79198		1 0	1	0	3	11870	4100	7770	124
20120521	J00	M4896		1 0	1	0	3	10690	1500	9190	124
20120511	M4896	K219		1 0	1	0	3	9230	1500	7730	124
20120627	G479	M179		1 0	1	0	7	9210	1500	7710	124
20120613	K219	M179		1 0	1	0	5	10840	1500	9340	124
20120619	K219	M179		1 0	1	0	5	9250	1500	7750	124
20120626	K297	M179		1 0	1	0	3	10820	1500	9320	124
20120608	K219	I100		1 0	1	0	30	10910	1500	9410	124
20120601	K294	G459		1 0	1	0	7	11870	4100	7770	124
20120623	K297	M179		1 0	1	0	3	9230	1500	7730	124
20120621	H814	K219		1 0	1	0	3	10690	1500	9190	124
20120724	M5490	H814		1 0	1	0	0	10790	1500	9290	124
20120730	K219	H814		1 0	1	0	5	9240	1500	7740	124
20120710	K219	M5490		1 0	1	0	3	9210	1500	7710	124
20120725	K219	H814		1 0	1	0	5	9240	1500	7740	124
20120707	K297	F459		1 0	1	0	30	14530	1500	13030	124
20120711	M5490	M179		1 0	1	0	0	10790	1500	9290	124
20120709	G459	I109		1 0	1	0	7	11870	4100	7770	124
20120703	M5490	K297		1 0	1	0	3	10850	1500	9350	124
20120831	M5490	H814		1 0	1	0	5	9210	1500	7710	124
20120803	K219	K297		1 0	1	0	5	12900	1500	11400	124
20120818	K219	K297		1 0	1	0	7	9230	1500	7730	124
20120808	K219	K297		1 0	1	0	30	10890	1500	9390	124
20120920	M5490	K297		1 0	1	0	5	10820	1500	9320	124
20120926	M5490	K297		1 0	1	0	5	9250	1500	7750	124
20120927	G479	R35		1 0	1	0	5	11820	1500	10320	124
20120903	I100	M5490		1 0	1	0	30	10870	1500	9370	124

# 학회 및 기관 연구 지원

REC_U_FR_DT_DM	IDV_ID	S_VSCN_DM	S_RECN_DM	S_TOT_PRES_DD_DM	M_DM_2013
20131202	196	9	9	360	1
20130114	198	1	1	15	1
20131201	199	357	357	0	1
20130326	204	5	5	120	1
20131202	215	3	3	60	1
20131231	282	17	17	336	1
20131209	308	0	0	0	1
20131231	676	14	14	516	1
20131205	690	6	6	0	1
20131115	807	20	27	134	1
20130531	1157	1	1	0	1
20131227	1241	47	47	403	1
20131224	1421	6	97	256	1
20130215	2098	1	1	0	1
20130529	2201	1	1	0	1
20130926	2549	14	24	0	1
20131226	2582	6	6	360	1
20131104	2627	23	30	0	1
20131223	2704	14	14	360	1
20131216	2728	6	6	180	1
20131201	2735	14	14	0	1
20130413	2874	3	3	125	1
20131223	2998	13	13	375	1
20131230	7395	28	28	437	1
20131021	7780	5	5	367	1
20130705	8858	1	1	0	1
20130525	9832	4	10	0	1
20130713	10520	3	3	0	1
20131210	15165	12	12	390	1
20131017	15688	12	331	0	1
20131224	18871	13	13	360	1
20131220	19048	35	56	164	1
20130128	19502	1	1	30	1
20131126	20695	4	4	180	1
20131231	22428	4	4	281	1
20131220	24221	69	83	407	1
20131205	24633	11	11	270	1
20131231	26403	15	15	375	1
20131231	27943	17	17	385	1
20131008	28401	9	9	270	1
20130709	28448	7	7	210	1

# 학회 및 기관 연구 지원

Exclusion criteria	YEAR	종류
Those with past history of EXPOSURE(A) use during the past 2 years	2007-2008	진료
Those with past history of hospitalization due to myocardial infarction (I21-I24) and stroke (I60-I69) during the past 2 years (2006-2007)	2007-2008	진료
Those with past medical history of stroke and heart disease in health interviews	2007-2008	검진
Those with past history of hospitalization due to cancer (C00-D48) during the past 2 years	2007-2008	진료
Those with past medical history of cancer in health interviews	2007-2008	검진
Inclusion criteria		
Men and women aged 30-64	2008	자격
Those with baseline serum LDL cholesterol $\geq 150$ mg/dl (3.89 mmol/L) at the health exam in 2008	2008	검진
EXPOSURE(A)		
EXPOSURE(A): 약제 종류 명시		
EXPOSURE(A) user: those with any prescription for EXPOSURE(A) in 2008-2009	2008-2009	진료
Outcome measures between 2009-2012		
Any hospitalization or death due to all strokes (I60-I69)	2009-2012	진료
Any hospitalization or death due to myocardial infarction (I21-I24)	2009-2012	진료
Any hospitalization or death due to cancer (C00-D48)	2009-2012	진료
Confounders		
AGE, SEX, 소득수준(보험료 분위)	2008	자격
Health behavior (smoking, alcohol, physical activity)	2008	검진
Family history (HTN, DM, stroke)	2008	검진
Bioclinical profiles(BP, FSG, BMI)	2008	검진

# 학회 및 기관 연구 지원

Table 1. Characteristics of subjects and outcome measures

	Total no. of subjects	EXPOSURE(A) user	EXPOSURE(A) non-user
Demographic factors			
Age (years)			
Females (%)			
Bioclinical profiles			
SBP (mmHg)			
DBP (mmHg)			
Body mass index (kg/m <sup>2</sup> )			
Waist circumference (cm)			
Fast serum glucose (mmol/l)			
Gamma glutamyl transferase (U/l)			
Serum LDL cholesterol (mmol/L)			
Serum HDL cholesterol (mmol/L)			
Serum triglyceride (mmol/L)			
Serum total cholesterol (mmol/L)			
Behavioral factors			
Current smoking (%)			
Alcohol consumption			
Physical exercise (%)			
Family history			
Family history of stroke			
Family history of heart disease			
Family history of DM			
Family history of hypertension			
Socioeconomic factors			
Low income quintiles (%)			
High area deprivation (%)			
Outcome measures			
All strokes			
Hemorrhagic stroke			
Ischemic stroke			
Myocardial infarction			
Cancer			
DM			

Table 1

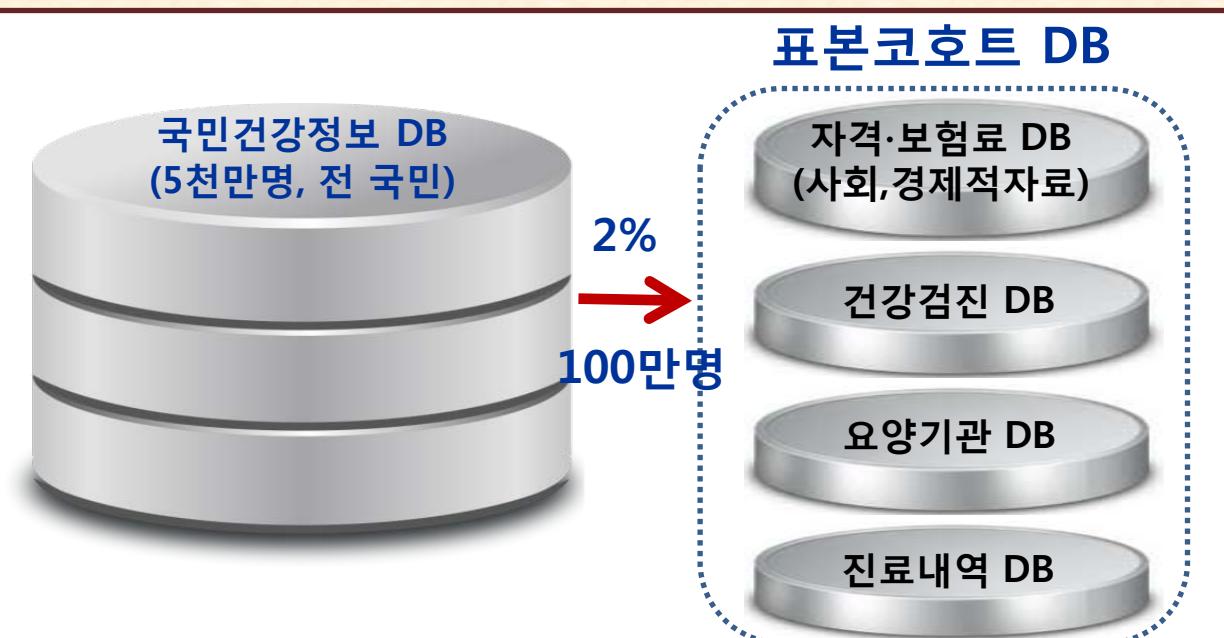
1. 양성 (Positive)					
2. 약물치료 필요여부					
예: 고위험군		예: 중등도위험군		예: 저위험군	
합계:					
합계:					
0.0-1.0년	건강->건강				
	건강 -> 관상동맥질환				
	건강 -> 사망				
	관상동맥질환 -> 관상동맥질환				
	관상동맥질환 -> 사망				
1.1-2.0년	건강->건강				
	건강 -> 관상동맥질환				
	건강 -> 사망				
	관상동맥질환 -> 관상동맥질환				
	관상동맥질환 -> 사망				
2.1-3.0년	건강->건강				
	건강 -> 관상동맥질환				
	건강 -> 사망				
	관상동맥질환 -> 관상동맥질환				
	관상동맥질환 -> 사망				
3.1-4.0년	건강->건강				
	건강 -> 관상동맥질환				
	건강 -> 사망				
	관상동맥질환 -> 관상동맥질환				
	관상동맥질환 -> 사망				
합계:					

# 공개용 표본코호트 DB 제공

- 국민건강정보 DB에서 대표성·신뢰성있는 표본(2%)을 추출하여 코호트로 구축완료
- 건강보험자료의 연구적 접근성 개선 및 근거 생산 활성화에 기여

## 표본코호트 DB

- 5천만 전 국민의 2%(100만명)의 '02~'10 (9년간)의 자격 및 보험료(사회경제적변수), 건강검진결과 및 병·의원 이용내역 등을 코호트 형식으로 구축
- 개인식별이 불가능한 일반 공개용 연구 DB
- 대표성·지속성·포괄성·완결성을 원칙으로 구축



# 공개용 표본코호트 DB 제공

구분	세부내용
추출 방식	- 환자 중심의 표본코호트(2002년말 가입자 자격기준)
총화	- 성별(2), 연령(18), 소득수준(41)을 총화변수로 1,476개층 구성
표본 수	<ul style="list-style-type: none"><li>- 전 국민 중 약 100만 명</li><li>- 층별 비례배분을 적용해 최종 결정된 2002년 표본크기는 1,025,340명임 (이는 모집단 인원수(약 4,600만명) 대비 2.2%에 해당함)</li></ul>
장점	<ul style="list-style-type: none"><li>- 층별 총의료비의 변동계수를 고려해 자료의 대표성 확보</li><li>- 많은 수의 층(1,476개)을 사용해 기타 변수들에 대한 대표성 보장</li><li>- 의료급여자 포함된 전국민 자료</li><li>- 건강검진 효과 평가 가능</li></ul>
단점	<ul style="list-style-type: none"><li>- 9개년 간의 제한적 코호트</li><li>- 표본 코호트 DB 자체의 규모도 커서 사용자 편의성 낮아짐</li><li>- 비급여 부재</li></ul>

The timeline diagram illustrates the progression of the project from January 2013 to July 2014. It features a horizontal blue line with five dark blue circular markers. Above each marker is a date: '13.1월', '13.5월', '13.12월', '14.6월', and '14.7월'. Blue arrows point upwards from four boxes below the timeline to these markers. The first box, labeled '표본 코호트 DB 구축' (Construction of Sample Cohort DB), points to the first marker ('13.1월') and contains the text '개인정보 대체 및 익명화' (Replacement and anonymization of personal information). The second box, labeled '시범연구 실시(16개 기관)' (Implementation of pilot study (16 institutions)), points to the third marker ('13.12월'). The third box, labeled '표본 코호트 세미나' (Sample cohort seminar), points to the fourth marker ('14.6월'). The fourth box, labeled '공개' (Release), points to the fifth marker ('14.7월').

# 공개용 표본코호트 DB 제공

## 가이드

GUIDE

보다 빠르게 정보를 찾으실 수 있도록  
도움을 드리겠습니다.

### 이의신청 안내

- ▣ 이의신청 자주묻는 질문
- ▣ 건강보험 이의신청
- ₩ 온라인 복지부 심판청구
- ▣ 고액 상습체납자 인적사항 공개

### 신고센터

- 부패신고 (사이버감사실)
- 악명신고 (스마트퀴즈)
- 부당청구요양기관신고포상
- 예산낭비신고
- 국민권익위원회 공익신고

### 웹 이용안내

- Smart Navigation 분할된 각종 서비스를 한눈에!
- 통합포털도움센터 쉽고 빠른 고객서비스 제공
- 시스템에러도우미 로그인 안될때 따라해보세요!

## 건강이야기

HEALTH CONTENTS

국민의 건강증진을 위한  
다양한 정보를 제공합니다.

### 건강 iN

다양하고 올바른 건강정보 건강IN에서  
제공하는 건강정보를 만나보세요!

### SNS 소통채널

SNS를 통해 언제 어디서나 편리하게  
국민건강보험공단과 소통하세요!

### 웹진

평생건강 지킴이! 건강보험 웹진!  
- H-well 예방 프로젝트  
- 이달의 건강예보



NHISS National Health Insurance Sharing Service

로그인 사이트맵 ENGLISH

소개 데이터신청 성과공유 통계 고객센터

**국민건강보험자료 공유서비스**  
근거 중심의 보건의료 정책 및 학술연구 지원을 선도합니다.

**표본연구DB**  
국민건강정보자료를 활용하여 구축한  
표본코호트DB를 제공합니다.

표본코호트DB  
제공절차  
신청하기

**데모용DB 다운로드**  
표본코호트DB를 체험해 보실 수 있는  
데모 자료를 제공 합니다.

다운받기

**건강·질병지표**  
만성질환 발생 전후를 관리할 수 있는  
모니터링 지표입니다.

이용안내

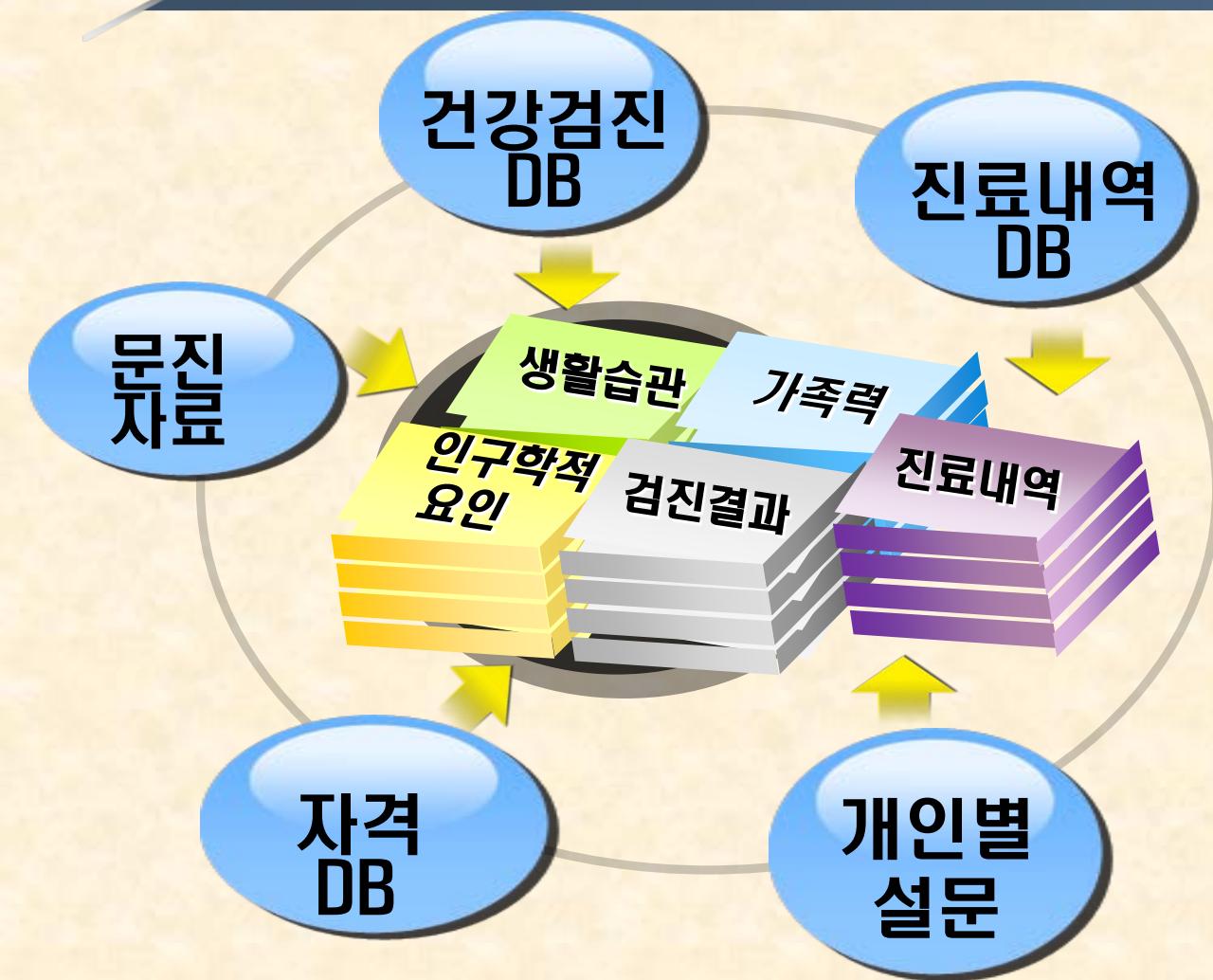
**질병통계**  
관심있는 질병의 진료현황을 주제·  
지역·요양기관종별로 비교할 수 있습니다.

**NHISS**  
National Health Insurance Sharing Service

제공안내  
신청하기  
소개보기

# 개인 맞춤형 건강서비스

## 맞춤형 기본 로직



# 개인 맞춤형 건강서비스



1  
**개인건강기록서  
비스(PHR)**  
건강검진결과,  
진료 및 투약내역, 문  
진결과 조회 및  
건강관리를 위한 입  
력정보란 제공

2  
**건강나이  
알아보기(HRA)**  
실제나이, 현재나이, 목  
표나이 제시 및 연령별  
위험요인 등 개선을 위  
한 행동 지침 및 처방  
메세지 제공

3  
**대사증후군  
맞춤정보**  
고혈압, 당뇨, 이상지질  
혈증 등 대사성 질환을  
중심으로 건강검진 결  
과와 연계한 처방 정  
보 제공

4  
**뇌졸중위험도예측  
프로그램**  
위험요인, 생활습관, 가족  
력, 환경요인 등을 기초  
로 10년 내  
뇌졸중 위험도 및  
처방 메시지 제공

5  
**비만개선  
프로그램**  
운동, 식사, 행동요  
법의 개별적 설문을  
통해 개인에게 최적  
화된 처방 메세지  
제공



건강 iN



“신뢰할 수 있는 건강정보 전문사이트

# 개인 맞춤형 건강서비스



## Personal Health Record(PHR), called My Health Bank

- **My Health Bank**에 저장된 개인건강정보를 안전하고 편리하게 조회 및 건강관심도 제고
- 건강N → 공인인증서 접속 → 나의맞춤건강정보 → 나의건강기록서비스 → 서비스신청

건강검진안내      나의건강정보      건강실천정보      질병의약정보      건강자료실      병원 및 검진기관안내

◆ | 나의건강정보 | 나의 건강정보 | 나의 건강정보란?

**나의 건강정보 (My Health Bank)**

나의 건강정보      나의 건강정보란?

나의 건강정보란?      서비스 개요      서비스목적 및 필요성      개인건강정보 사용범위      개인정보 보호지침

‘나의 건강정보’는  
산재되어있는 자신의 건강정보를 한 곳에 모아 언제, 어디에서든 필요할 때마다 나의 건  
강정보를 확인하고 관리하는 건강관리서비스입니다.  
‘나의 건강정보’는 은행을 통해 자신의 자산을 관리하듯 나의 건강정보를 안전하고 편리  
하게 저장하여 관리한다는 의미를 담고 있습니다.



# 개인 맞춤형 건강서비스

## 나의 건강기록서비스 종류(Personal Health Record, PHR)

### 건강검진 결과조회

- 최근 **5년간** 검진결과  
누적 조회

### 문진표 조회

- 전년도/당해년도** 검진  
기관에서 청구한 문진  
표 조회

### 진료 및 투약내역

- 최근 **1년간** 진료 및 투  
약기록 조회
- 약물정보 제공

### 본인건강 입력정보

- 본인이 관리하고  
싶은 건강정보를  
입력하여 누적관리



건강검진결  
과

문진표  
조회

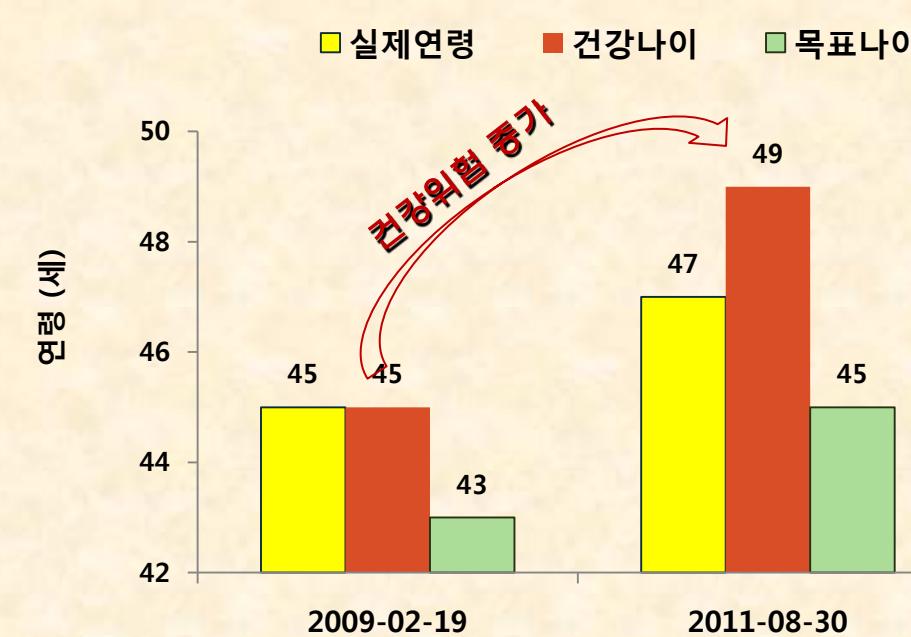
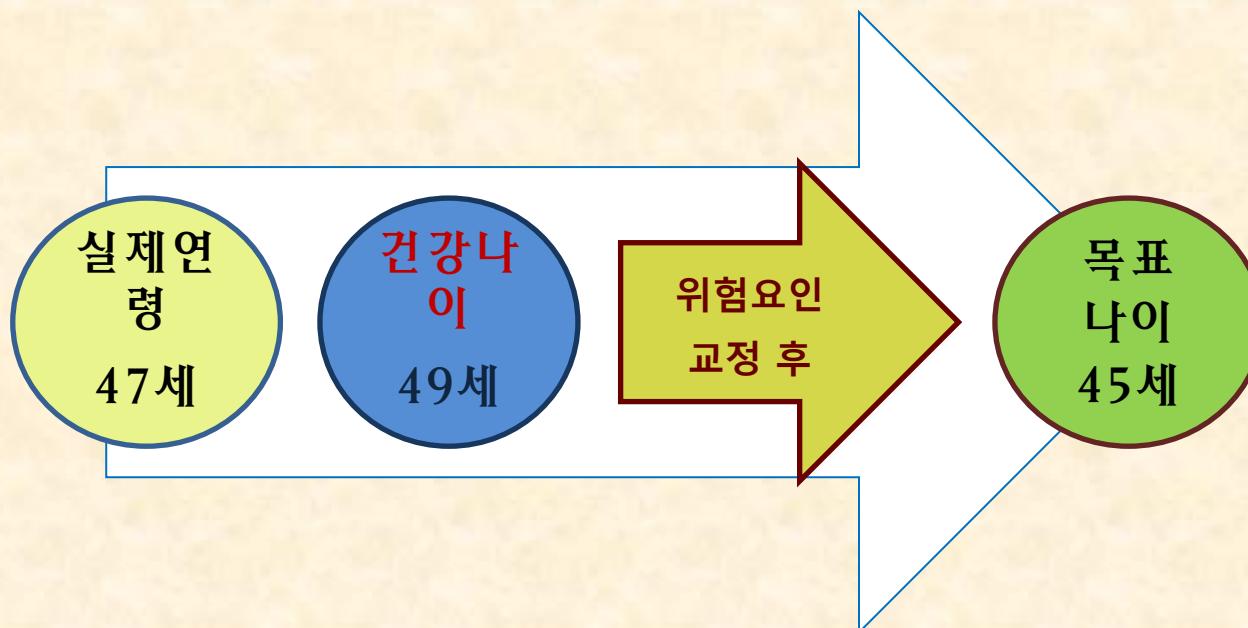
진료 및 투약내  
역

본인 건강  
입력정보

# 개인 맞춤형 건강서비스

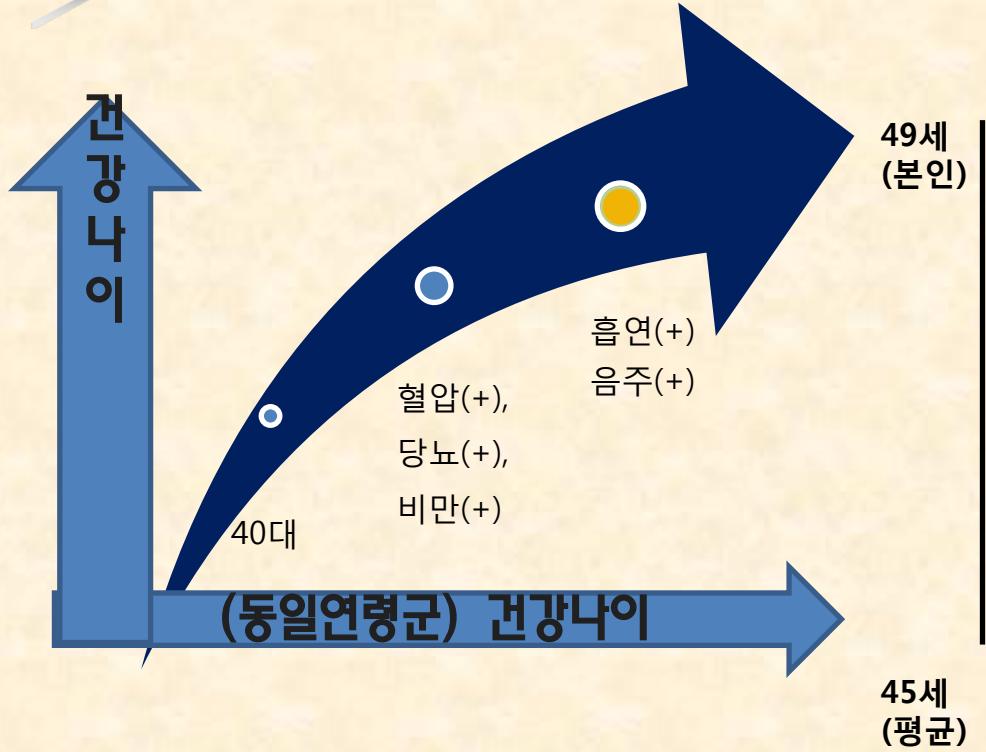
## 건강나이 알아보기(Health Risk Appraisal, HRA)

- 위험도 산출 구성요소: 비만, 흡연, 음주, 신체활동, 혈압, 혈당, 총 콜레스테롤, LDL/HDL 콜레스테롤
- 실제나이, 현재 건강나이, 목표나이 평가
- 성별, 연령별 위험요인 제시 - > 행동지침과 교육자료 제공

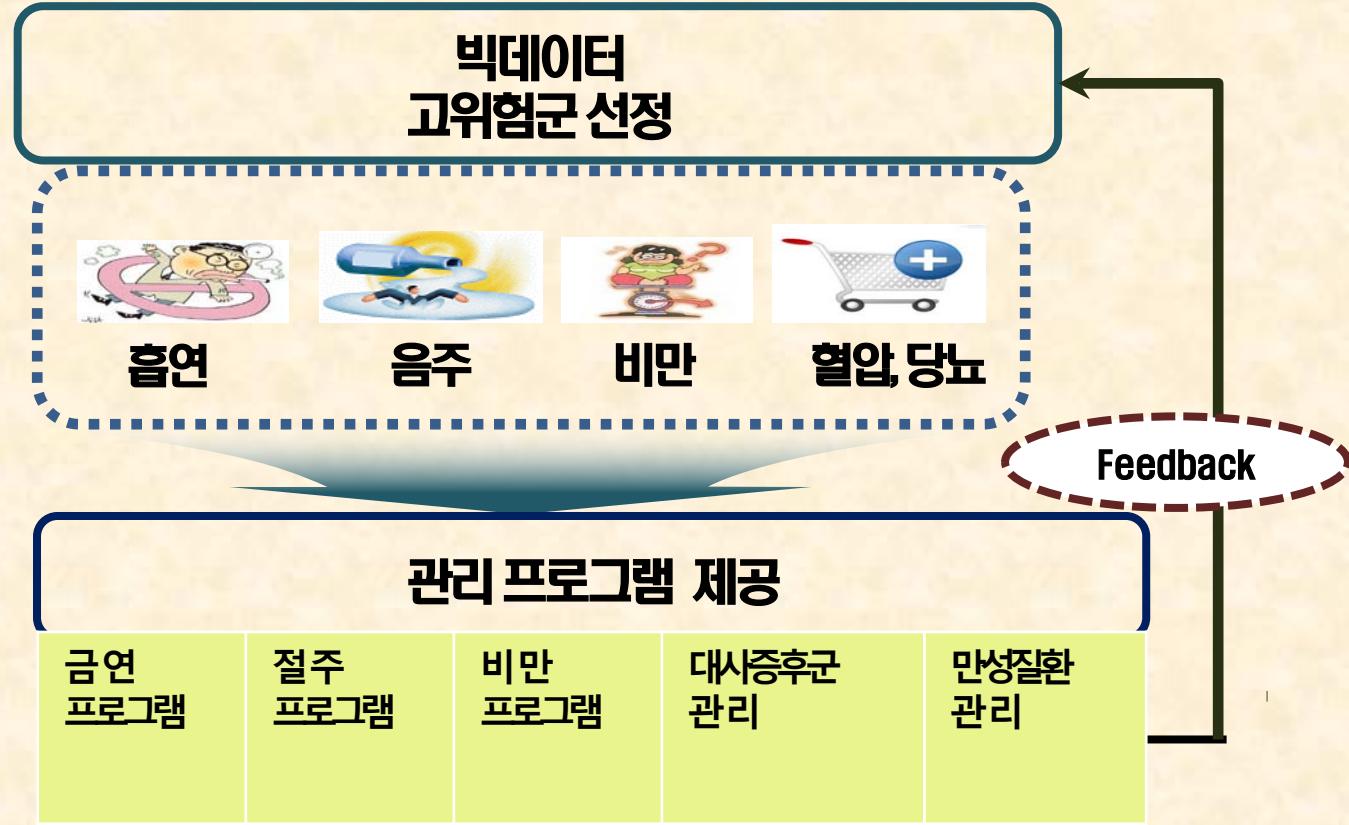


# 개인 맞춤형 건강서비스

## 건강나이 알아보기(Health Risk Appraisal, HRA)



- 동일 연령군 비교 및 시각화  
(홈페이지, 앱)



# 개인 맞춤형 건강서비스



# 지역·사업장 단위의 건강·질병지표

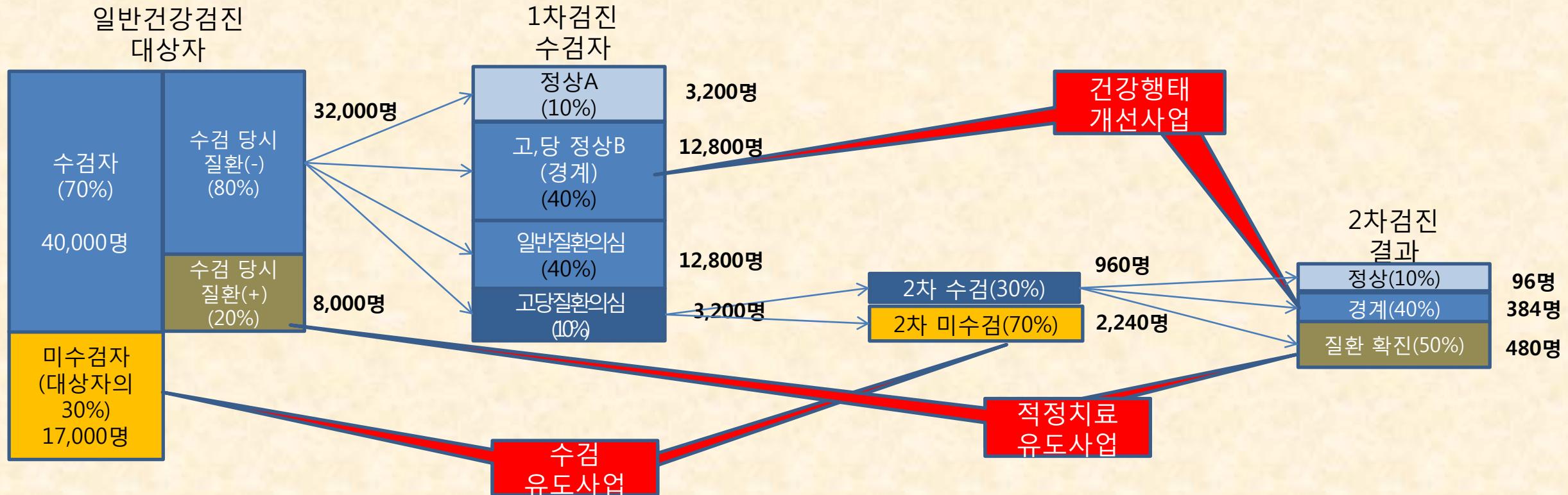
- ❖ 지역별(시도, 시군구), 사업장별 지표를 개발하여 유관기관(보건소·사업장 등)에 제공,  
지역별·사업장별 특성에 맞는 건강관리사업 지원 → 맞춤형 서비스 기반 마련



예) 고혈압/당뇨/이상지질/심뇌혈관질환 관리 지표



# 지역·사업장 단위의 건강·질병지표



## 수검유도사업

- 1차검진 미수검자
  - 다수(대중사업)
- 2차검진 미수검자
  - 소수(개인맞춤사업)

## 건강행태개선사업

- 다수(대중사업)
- 소수(개인맞춤사업)
- 우선순위 설정
  - 대사증후군 진단기준
  - 흡연, 음주, 신체활동
  - 경제적, 지리적 취약계층

## 적정치료유도사업

- 기존 유질환자
  - 다수(대중사업)
- 2차검진 신규 유질환자
  - 소수(개인맞춤사업)

# 지역·사업장 단위의 건강·질병지표

■ 건강·질병지표 제공(웹사이트) … 읍면동, 연령(10세단위 6군), 소득수준(보험료분위)별 등

지역별	시도별	시군구별	읍면동별	성별	연령별	연도별	
	<ul style="list-style-type: none"><li>시도 계</li><li>서울특별시</li><li>부산광역시</li><li>대구광역시</li><li>...</li></ul>	<ul style="list-style-type: none"><li>시군구 계</li><li>강남구</li><li>송파구</li><li>...</li><li>...</li></ul>	<ul style="list-style-type: none"><li>읍면동 계</li><li>논현동</li><li>청담동</li><li>...</li></ul>	<ul style="list-style-type: none"><li>전체</li><li>남</li><li>여</li></ul>	<ul style="list-style-type: none"><li>전체</li><li>20~29세</li><li>30~39세</li><li>...</li></ul>	<ul style="list-style-type: none"><li>2002</li><li>2003</li><li>2004</li><li>...</li></ul>	
사업장별/직종별	사업장별/직종별		질환별(15개)		지표별(100개)		기초통계
	<ul style="list-style-type: none"><li>사업장 전체</li><li>300명이하 사업장</li><li>500명이하 사업장</li><li>1,000명이하 사업장</li><li>...</li></ul>		<ul style="list-style-type: none"><li>업종별 전체</li><li>농업·수렵업</li><li>공업</li><li>숙박업</li><li>제조업</li><li>...</li></ul>		<ul style="list-style-type: none"><li>비만</li><li>고혈압</li><li>당뇨병</li><li>이상지질혈증</li><li>심뇌혈관질환</li><li>만성신장질환</li></ul>		<ul style="list-style-type: none"><li>총 인구수</li><li>지역별 인구수</li><li>연령별 인구수</li><li>사망자수</li><li>질환별 환자 수</li><li>질환별 사망자수</li></ul>

지역별 지표별

동일 및 유사지역  
지표별 비교분석

연도별 추이

지자체 및 사업장별로 건강서비스우선순위 설정 및 보건계획 수립

지자체 및 사업장간 비교평가 지표 활용

보건의료정책의 Feedback 및 사업실행에 대한 평가 활용

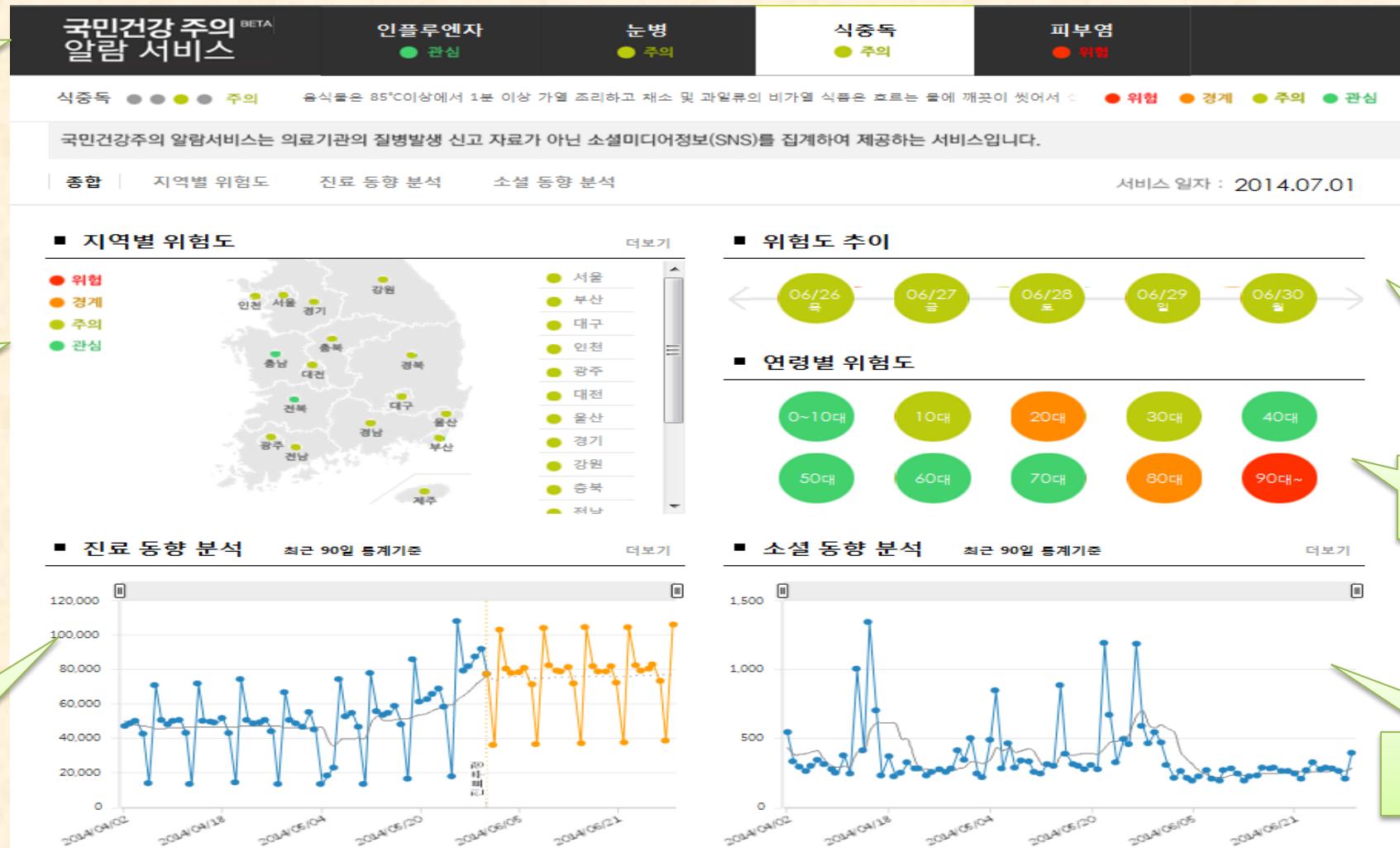
# 국민건강 주의 알람 서비스

- 건강 빅데이터와 SNS 정보를 융합한 질병 발생예측 모델



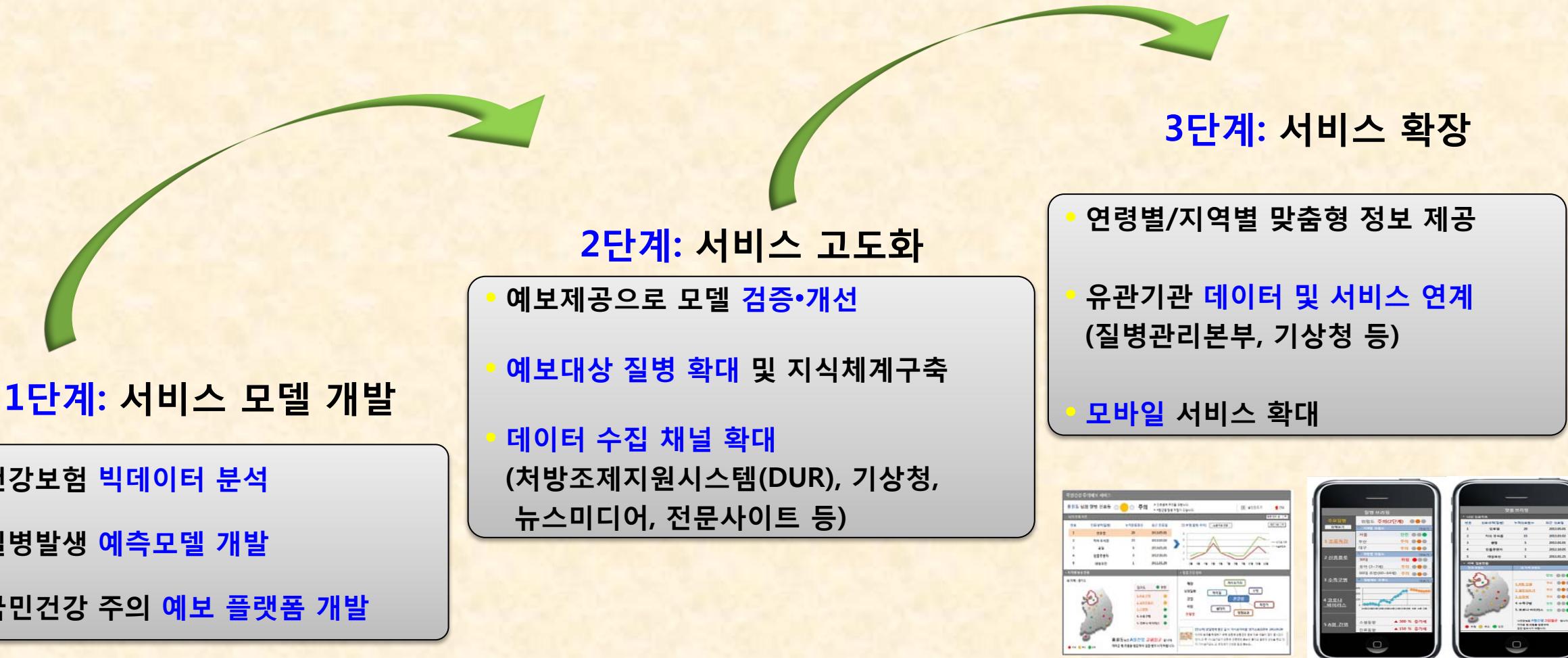
# 국민건강 주의 알람 서비스

질병별 예보



질병 관련  
진료 동향 분석

# 국민건강 주의 알람 서비스



[맞춤형 정보제공]

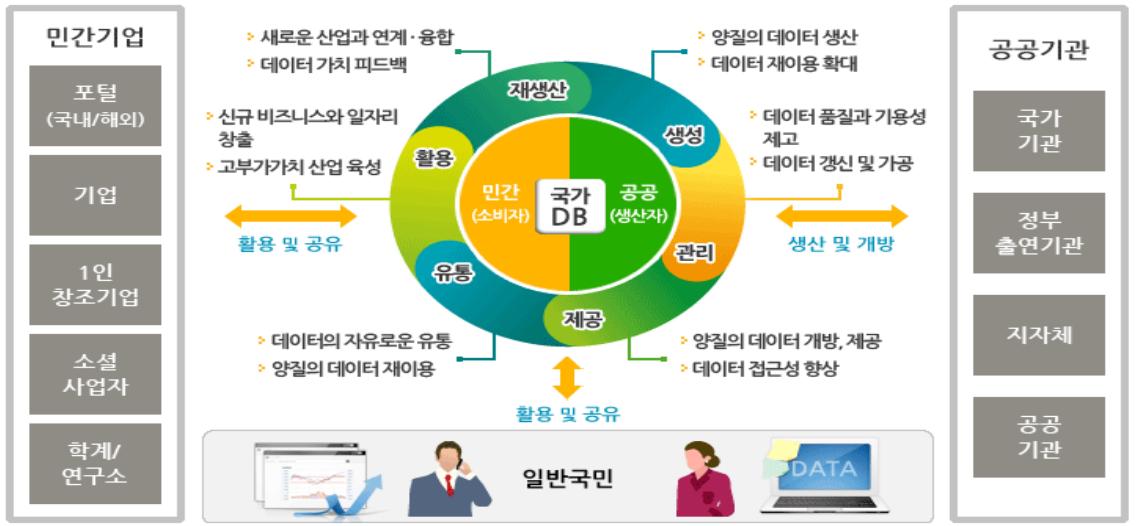


[모바일 서비스]

# 국가 DB 구축 사업 참여

## 국가DB 로드맵

### 데이터의 생성·관리·유통·활용의 선순화 체계선도



## 2014년도 국가DB 구축 사업 참여

- 사업명: 빅데이터를 활용한 지역 및 사업장 단위 건강 · 질병지표 DB구축
- 주관 기관: 미래창조과학부(한국정보화진흥원)
- 2014년 지원 예산: 600백만 원
- 내용: 흡연율, 비만율, 만성질환의 유병율, 사망률 등 상세 건강 질병지표를 지역 및 사업장 단위로 제공

## 국가DB 의의

국가나 민간에서 보유하고 있는 가치 있는 DB를 국가 DB로 지정하고 민간이나 공공기관에서 활용할 수 있도록 국가에서 지원하고 관리

## 향후 계획

공간정보와 결합한 융합DB구축  
☞ 3D공간지도를 통한 「맞춤형 의료이용지도」 서비스 제공

# 참조표준데이터 지정 추진

한국인의 특성이 반영된 보건의료분야(국가건강검진의실측데이터) 참조표준개발 · 개방

## | 참조표준데이터의 선정

### 한국인의 건강지수(korean Health Index)

건강에 대한 관심과 이해도를 높일 수 있도록  
국가건강검진의 검사항목을 분석하여  
한국인의 건강상태를 반영한 지수

\*건강데이터는 국가별로 체형이나 식생활이 달라 20% 이상 차이 발생

약물오남용 · 오진발생예방, 의료기기 개발 등 활용



유아의 월령별 체중 데이터 제공으로  
유아복 · 유모차 생산에 활용

## | 단계적 참조표준 보급계획

1단계(2015~2017)

<개발 · 확장>

참조DB개발 로드맵 수립  
참조 DB후보 도출

2단계(2018~2020)

<검증 · 응용기반>

참조DB확장 구축  
산업연계형 활용 로드맵 수립

3단계(2021~2024)

<응용 · 확산>

통합 플랫폼화를 통한  
건강지수 백본(Back bone) 구축 및 활용

**활용 전제조건:**  
**청구자료 → 분석용 자료**

# 데이터 정제

- 진료개시일자 기준 재구축
  - 최근 진료개시일 기준 1개 월 내 청구율 70%대
  - 최근 진료개시일 기준 2개 월 내 청구율 90%대
  - 진료개시일 이후 최대 3년 6개월까지 청구자료를 추적하여 진료개시일 기준으로 DB 구축
- 연구용 건강검진 DB 구축
  - 연도별, 검진 프로그램별 테이블 분리
  - 청구 등 행정절차에 필요 한 변수 제외
  - 새로운 변수명 부여
  - 판정, 자동계산값 검증
  - 주요 대상질환별 진료데이터 연계(예정)

# 데이터 정제

- 오입력값 처리
  - 흡연기간이 현재 나이보다 많을 경우
  - 과거력 진단여부 응답은 '미해당'인데, 약물치료여부에 '해당'이라고 응답한 경우
  - 흡연여부는 '과거흡연 혹은 비흡연'이라고 응답한자가 '현재 하루흡연량'에 값이 있는 경우 등
  - HDL, LDL이 총콜레스테롤보다 큰 경우
- 극단값 검토
  - 각 변수의 분포확인 후 상하위 0.05%를 임계점으로 설정
    - $Pn = (P99.95 - P0.05) / 99.9$
    - $P0 = \text{MAX}(P0.05 - (5 * Pn), \text{min})$
    - $P100 = \text{MIN}(P99.95 - (5 * Pn), \text{max})$
    - 실제 MIN, MAX값을 수정
  - 반복측정값의 표준편차를 이용하여 2SD이상의 값을 MISSING처리(2SD는 수정 가능)
  - 반복 수행
    - 적정 Percentile과 SD값은 임상의 와 상의 필요.

# 데이터 정제

- 텍스트 자료 코딩
  - 요양기관주소 동읍면 코딩(완료), 요양기관 및 환자주소 지오 코딩(올해 예정), 기타 검진 텍스트 항목 등

사용도구	FIND 함수 MID, LEFT 함수 IFERROR 함수																																																								
방법	=IFERROR(IFERROR(IFERROR(IFERROR(IFERROR(IFERROR(MID(J2,FIND("동",J2)-2,3),MID(I2,FIND("동",I2)-2,3)),MID(H2,FIND("동",H2)-2,3)),LEFT(G2,FIND("동",G2))),LEFT(G2,FIND("읍",G2))),LEFT(G2,FIND("면",G2))),F2)																																																								
그림설명	<p>FIND : 텍스트에서 "동", "읍", "면" 을 찾는 역할</p> <p>MID,LEFT : "동", "읍", "면" 을 찾은 텍스트에서 "동" 까지의 3글자, "읍", "면"을 찾은 경우는 그 앞의 텍스트를 추출하는 역할</p> <p>IFERROR : 함수를 해당하는 값이 없다면 다른 조건으로 값을 추출</p> <p>The screenshot shows a Microsoft Excel spreadsheet. On the left, there's a table with columns for province, city, and district. In the middle, there's a list of extracted place names. A red circle highlights the list of extracted place names.</p> <table border="1"><thead><tr><th>eun</th><th>전라북도</th><th>군산시</th><th>조촌동</th><th>777</th><th></th><th></th><th>조촌동</th></tr></thead><tbody><tr><td></td><td>충청남도</td><td>천안시</td><td>동남구</td><td>풍세면</td><td>630</td><td></td><td>풍세면</td></tr><tr><td></td><td>충청남도</td><td>천안시</td><td>서북구</td><td>쌍용1동</td><td>2층</td><td></td><td>쌍용1동</td></tr><tr><td></td><td>전라북도</td><td>익산시</td><td>어양동</td><td>665-1</td><td></td><td></td><td>어양동</td></tr><tr><td></td><td>경기도</td><td>용인시</td><td>수지구</td><td>죽전동</td><td>에미랄드빌딩301호</td><td></td><td>죽전동</td></tr><tr><td></td><td>충청남도</td><td>논산시</td><td>연무읍</td><td>죽령리</td><td>76-16</td><td></td><td>연무읍</td></tr><tr><td></td><td>대구광역시</td><td>수성구</td><td>지산동</td><td>지벌로</td><td></td><td></td><td>지산동</td></tr></tbody></table> <p>Extracted place names (highlighted by a red circle):</p> <ul style="list-style-type: none"><li>조촌동</li><li>풍세면</li><li>쌍용1동</li><li>어양동</li><li>죽전동</li><li>연무읍</li><li>지산동</li></ul>	eun	전라북도	군산시	조촌동	777			조촌동		충청남도	천안시	동남구	풍세면	630		풍세면		충청남도	천안시	서북구	쌍용1동	2층		쌍용1동		전라북도	익산시	어양동	665-1			어양동		경기도	용인시	수지구	죽전동	에미랄드빌딩301호		죽전동		충청남도	논산시	연무읍	죽령리	76-16		연무읍		대구광역시	수성구	지산동	지벌로			지산동
eun	전라북도	군산시	조촌동	777			조촌동																																																		
	충청남도	천안시	동남구	풍세면	630		풍세면																																																		
	충청남도	천안시	서북구	쌍용1동	2층		쌍용1동																																																		
	전라북도	익산시	어양동	665-1			어양동																																																		
	경기도	용인시	수지구	죽전동	에미랄드빌딩301호		죽전동																																																		
	충청남도	논산시	연무읍	죽령리	76-16		연무읍																																																		
	대구광역시	수성구	지산동	지벌로			지산동																																																		

# Operational Definition

- 상병코드 ≠ 환자
  - 주상병: 진료기간 중 치료나 검사 등에 대한 환자의 요구가 가장 컸던 상병
  - 부상병: 진료기간 중 주상병과 함께 있었거나 발생된 상병으로 환자 진료에 영향을 주었던 상병
  - 배제진단
    - 상병명을 최종 확진한 경우 이전에 고려되었던 R/O 상병 등
    - '최종상병이 확진된 경우 이전에 고려하였지만 배제된 상병'

# Operational Definition

- 고혈압

대상질환	상병코드 정의	논문 갯수
고혈압	I10	6
	I10-13, I15	5
	I10-13, I15, I27.0, I67.4, I70.1, G93.2, K76.6	1
	없음	6
	계	18

- 당뇨병

대상질환	상병코드 정의	논문 갯수
제2형 당뇨병	E11	2
당뇨병	E10-14	11
	없음	7
	계	20

# Operational Definition

- 뇌졸중

대상질환	상병코드 정의	논문 갯수
뇌경색	I63	3
	I63-69	1
	I63, I67-69	1
뇌출혈	I60-62	3
뇌졸중	I60-63	4
	I60-64	4
	I60-69	5
	I60-69, G45	2
	I60-63, I67-69, G45-46	1
	I60 (I60.8 제외)	1
	없음	5
계		30

- 심혈관질환

대상질환	상병코드 정의	논문 갯수
심혈관질환	I00-99	1
	I20-25	1
	I20-25, I60-63	1
	I20-25, I42, I50, I60-64	1
	I10-15, I20-25, I44-51, I60-69, I70-74, R96	2
	없음	1
계		7

# Operational Definition

- 고혈압

대상질환	환자 정의	논문 갯수
고혈압	상병	4
	상병 및 약제 1일 이상	2
	상병 및 약제 7일 이상	1
	상병 및 연간 2회 이상 외래 혹은 1회 이상 입원	2
	계	9

- 당뇨병

대상질환	환자 정의	논문 갯수
당뇨병	상병	2
	상병 및 약제 1일 이상	5
	상병 혹은 약제 1일 이상	3
	상병 혹은 약제 1일 이상 혹은 검진	1
	상병 및 연간 2회 이상 의료이용	1
	상병 및 연간 3회 이상 외래 혹은 1회 이상 입원	1
	계	13

# Operational Definition

## • 심뇌혈관질환

대상질환	환자 정의	논문 갯수
관상동맥질환	상병 1년 이상	1
	상병 연간 3회 이상 외래 혹은 2일 이상 입원 혹은 PCI나 CABG 시술	1
허혈성심질환	상병	1
	상병 및 입원	1
불안정협심증	상병	1
급성심근경색	상병	2
심근경색	상병	1
	상병 및 입원	1
뇌경색	상병 1년 이상	1
뇌출혈	상병 1년 이상	1
뇌졸중	상병	6
	상병 및 입원	2
	상병 및 연간 1회 이상 입원 혹은 3회 이상 외래	1
	상병 및 입원 혹은 응급실	1
	상병 및 14일 이상 입원	1
	상병 및 입원 및 뇌 CT 촬영	2
심혈관질환	상병	2
	상병 1년 이상	1
만성신부전	상병	1
계		48

# Operational Definition

- 주/부상병 중 당뇨병 청구 횟수(2012. 12.)

횟수	빈도	백분율(%)
1	3,105,364	96.31
2	108,882	3.38
3	9,217	0.29
4	696	0.02
5	20	0.00

	주상병	1 부상병	2 부상병	3 부상병	4 부상병
N	1,501,114	855,628	390,753	227,647	130,222
%	48.3%	27.6%	12.6%	7.3%	4.2%

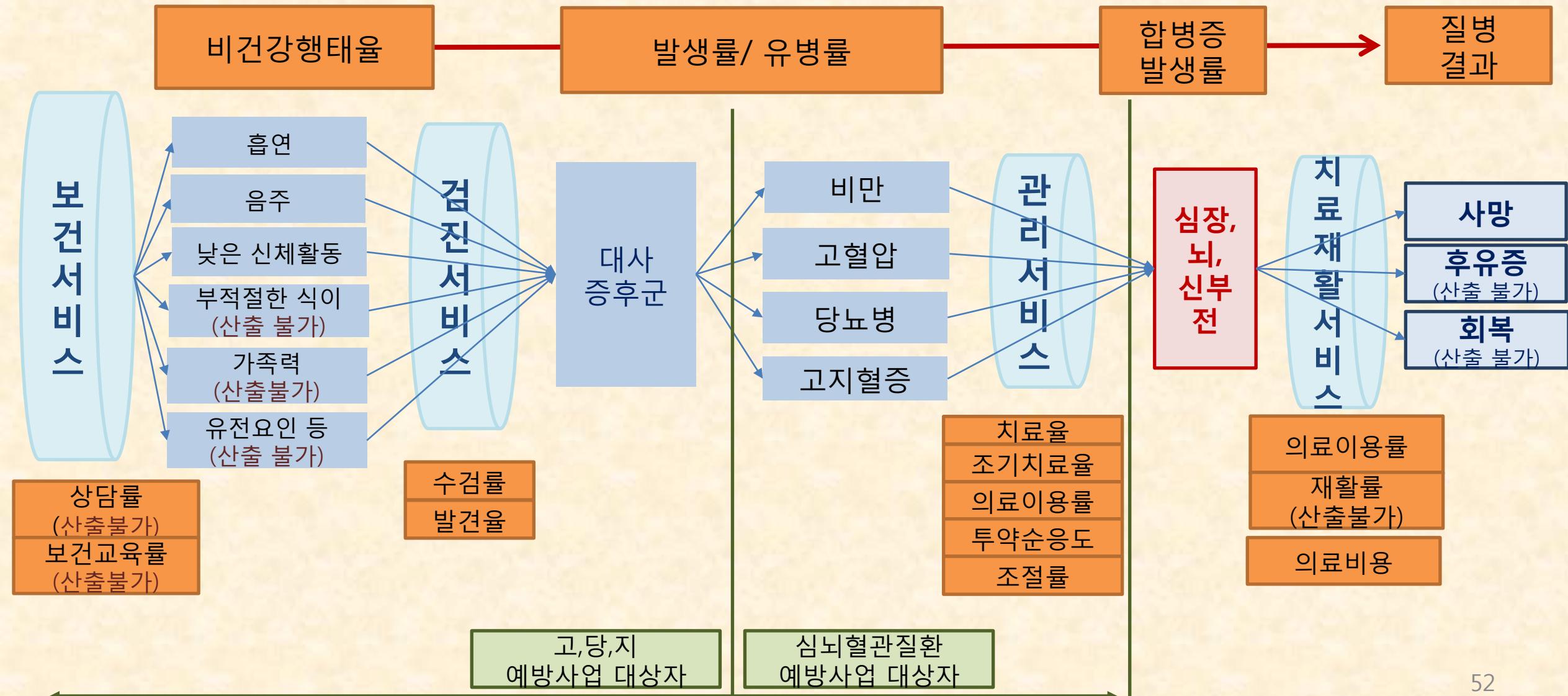
# Operational Definition

연도	상병명	상병기호	대상자 (주상병) <b>A</b>	대상자 (주부상병) <b>B</b>	A/B %
2010	자궁내막증	N80	1,387	1,942	71.4%
2010	바이러스성 간염	B16-19	8,447	11,672	72.4%
2010	알코올성 간질환 등	k70-77	24,184	44,025	54.9%
2004	천식	J45*	42,087	62,365	67.5%
2010	급성 충수염	K35	2,541	2,726	93.2%
2009	상세불명의 열	R50.9	4,874	10,944	44.5%
2004	악성신생물	C*	10,089	11,829	85.3%
2003	눈 합병증을 동반한 당뇨병 등	E10.3*, E11.3*, E12.3*, E13.3*, E14.3*	729	1,095	66.6%
2004	심방 잔떨림 및 된떨림	I48	988	1,988	49.7%

# Operational Definition

예) 당뇨병		주부상병+3개			주부상병			주상병		
		N	%	누적%	N	%	누적%	N	%	누적%
약물처방 미 고려	환자수	3,531,260			2,821,901			2,174,179		
		755,553	21.4	22.2	459,311	16.3	16.7	347,417	16.2	16.4
	진료횟수	261,395	7.4	29.6	194,805	6.9	23.6	173,092	8.1	24.5
		172,738	4.9	34.5	153,198	5.4	29.1	146,168	6.8	31.3
		173,996	4.9	39.4	163,954	5.8	34.9	156,985	7.3	38.6
		160,786	4.6	44.0	155,318	5.5	40.4	142,573	6.6	45.2
	환자수	2,332,358			2,174,179			1,627,361		
		680,523	29.2	29.1	352,723	16.4	16.6	270,823	16.8	16.9
		435,582	18.7	47.8	190,464	8.8	25.4	161,887	10.0	27.0
		189,038	8.1	55.9	156,654	7.3	32.7	137,042	8.5	35.5
		133,676	5.7	61.7	148,663	6.9	39.6	128,918	8.0	43.4
		113,533	4.9	66.5	134,573	6.2	45.8	111,473	6.9	50.4

# Operational Definition



# Operational Definition

지표명	분자 정의	분모 정의
고혈압유병률	해당년도 고혈압 코드(I10~15)가 명시된 모든 상병으로 의료이용한 사람 중 고혈압 약제처방을 받았던 사람 혹은 일반1차건강검진 혈압측정결과가 기준값 이상인 사람	해당 연도 자격자 전체
고혈압발생률	전년도 고혈압 유병자가 아니면서 해당년도 고혈압 유병자인 사람	해당 연도 자격자 전체
당뇨병유병률	해당년도 당뇨병 코드(E10~14)가 명시된 모든 상병으로 의료이용한 사람 중 당뇨병 약제처방을 받았던 사람 혹은 일반1차건강검진 혈당측정결과가 기준값 이상인 사람	해당 연도 자격자 전체
당뇨병발생률	전년도 당뇨병 유병자가 아니면서 해당년도 당뇨병 유병자인 사람	해당 연도 자격자 전체
허혈성심질환전체의료이용률	해당년도 허혈성심질환 코드(I20~25)가 명시된 모든 상병으로 의료이용한 사람	해당 연도 자격자 전체
허혈성심질환입원의료이용률	해당년도 허혈성심질환 코드(I20~25)가 명시된 모든 상병으로 입원이용한 사람	해당 연도 자격자 전체
뇌혈관질환전체의료이용률	해당년도 뇌혈관질환 코드(I60~69)가 명시된 모든 상병으로 의료이용한 사람	해당 연도 자격자 전체
뇌혈관질환입원의료이용률	해당년도 뇌혈관질환 코드(I60~69)가 명시된 모든 상병으로 입원이용한 사람	해당 연도 자격자 전체
대사증후군발견율	해당년도 일반1차검진 결과 대사증후군 기준에 해당하는 사람	해당년도 일반1차검진 수검자
흡연율	해당년도 일반1차검진에서 현재흡연 중으로 응답한 사람	해당년도 일반1차검진 수검자
음주율	해당년도 일반1차검진에서 주당 하루 이상으로 1일 기준, 4잔 이상 술을 마신다고 응답한 남성과 3잔 이상 술을 마신다고 응답한 여성	해당년도 일반1차검진 수검자
권장신체활동수행률	해당년도 일반1차검진에서 중등도신체활동을 5일 이상 혹은 고강도신체활동을 3일 이상 한다고 응답한 사람	해당년도 일반1차검진 수검자
비만율	해당년도 일반1차검진에서 체질량지수 25kg/m <sup>2</sup> 인 사람	해당년도 일반1차검진 수검자

# Operational Definition

지표명	분자 정의	분모 정의
1년후고혈압투약순응률	해당년도 다음 해에 고혈압 약제를 300일 이상 처방 받은 사람	해당년도 고혈압유병자
1년후고혈압정기방문율	해당년도 다음 해에 고혈압 코드(I10~15)가 명시된 모든 상병으로 분기별 1회 이상 의료기관을 방문한 사람	해당년도 고혈압유병자
1년후고혈압조절률	해당년도 다음 해 일반건강검진에서 혈압측정결과가 기준 값 이하인 사람	해당년도 고혈압유병자 중 다음 해 수검자
2년후고혈압조절률	해당년도 2년 후 일반건강검진에서 혈압측정결과가 기준 값 이하인 사람	해당년도 고혈압유병자 중 다음 해 수검자
고당환자의1년후정기방문율	해당년도 다음 해에 고혈압 혹은 당뇨병 코드가 명시된 모든 상병으로 분기별 1회 이상 의료기관을 방문한 사람	해당년도 고혈압 혹은 당뇨병 유병자
고당지환자의1년후정기방문율	해당년도 다음 해에 고혈압 혹은 당뇨병 혹은 이상지질혈증 코드가 명시된 모든 상병으로 분기별 1회 이상 의료기관을 방문한 사람	해당년도 고혈압 혹은 당뇨병 혹은 이상지질혈증 유병자
해당상병 유병자의 1년이내 심뇌신 입원발생률	해당년도 다음 해에 허혈성심질환 혹은 뇌혈관질환 혹은 만성신부전 코드가 명시된 모든 상병으로 입원이용한 사람	해당년도 해당상병 유병자
해당상병 유병자의 2년이내 심뇌신 입원발생률	해당년도 2년 후 허혈성심질환 혹은 뇌혈관질환 혹은 만성신부전 코드가 명시된 모든 상병으로 입원 이용한 사람	해당년도 해당상병 유병자
해당상병 유병자의 3년이내 심뇌신 입원발생률	해당년도 3년 후 허혈성심질환 혹은 뇌혈관질환 혹은 만성신부전 코드가 명시된 모든 상병으로 입원 이용한 사람	해당년도 해당상병 유병자