

한국노년신경정신약물학회 워크숍

Statistical Analysis of Big data

(건강보험자료 분석 과정 및 프로그밍)

서울대병원
의학연구협력센터
이중엽

학습목표

- 건강보험자료의 특성을 이해하고, 임상 연구에 활용할 때 주의할 점을 설명할 수 있다.
- 건강보험자료의 구조를 이해하고 환자 기반으로 재구성하고 활용하는 과정을 설명할 수 있다.
- 건강보험자료를 활용한 성과연구의 과정을 이해하고 각 단계별로 필요한 조치가 무엇인지 설명할 수 있다.

건강보험 청구자료의 이해

- 국가에서 관장하는 전 국민이 통합된 국가 단일보험
- 적용 대상자는 국내에 거주하는 국민으로 가입자 및 피부양자
- 복지부(제도, 관리), 공단+심평원 (운영)
 - 행위별 수가제도(fee-for-service)를 기본으로 함
 - 질병군별 포괄수가제(diagnosis related group, DRG: 백내장수술, 맹장수술, 항문수술, 편도수술, 탈장수술, 자궁수술, 제왕절개분만 7가지의 수술)
 - 정액수가제(요양병원, 보건기관 등)을 병행

(별지 제10호 서식)

안과

서식번호	G I O 2				요양급여비용명세서		요양기관	
	종류번호		상행번호		구분	내역	내역	내역
가입자성명	상행번호				상행구분	요양기관명		내역
수진자성명	상행번호				상행구분	요양기관명		내역
상행번호	내분기	수술	신원	의안	기	의안	의안	의안
본인 부담액 (I)								
본인 부담액 (II)								
1. 진찰료 (의안) (보험)	①초진	원						
	②회진	원						
2. 입원료	③의약품	원						
	④유합	원						
	⑤일반	원						
	⑥중환자실	원						
	⑦진단실	원						
	⑧기타	원						
	⑨가산실	원						
3. 투약료 및 처방료	①처방	원						
	②의약품	원						
4. 주사료	③④⑤⑥⑦⑧⑨⑩⑪⑫⑬⑭⑮⑯⑰⑱⑲⑳㉑㉒㉓㉔㉕㉖㉗㉘㉙㉚㉛㉜㉝㉞㉟㊱㊲㊳㊴㊵㊶㊷㊸㊹㊺㊻㊼㊽㊾㊿	원						
	①	원						
	②	원						
	③	원						
5. 마취료	④	원						
6. 이환요별료	⑤	원						
7. 경신요별료	⑥	원						
8. 처치 및 수술료	⑦⑧⑨⑩⑪⑫⑬⑭⑮⑯⑰⑱⑲⑳㉑㉒㉓㉔㉕㉖㉗㉘㉙㉚㉛㉜㉝㉞㉟㊱㊲㊳㊴㊵㊶㊷㊸㊹㊺㊻㊼㊽㊾㊿	원						
9. 검사료	①②③④⑤⑥⑦⑧⑨⑩⑪⑫⑬⑭⑮⑯⑰⑱⑲⑳㉑㉒㉓㉔㉕㉖㉗㉘㉙㉚㉛㉜㉝㉞㉟㊱㊲㊳㊴㊵㊶㊷㊸㊹㊺㊻㊼㊽㊾㊿	원						
10. 영상진단 및 방사선 치료료	①②③④⑤⑥⑦⑧⑨⑩⑪⑫⑬⑭⑮⑯⑰⑱⑲⑳㉑㉒㉓㉔㉕㉖㉗㉘㉙㉚㉛㉜㉝㉞㉟㊱㊲㊳㊴㊵㊶㊷㊸㊹㊺㊻㊼㊽㊾㊿	원						
S. 특수장비	①CT	원						
	②MRI	원						
	③PET	원						
A. 100분의100미만 본인부담 1	①의약품	원						
	②치료재료	원						
	③진료비용	원						
B. 100분의100미만 본인부담 2	①의약품	원						
	②치료재료	원						
	③진료비용	원						
U. 건강보험100분의100본인부담	①의약품	원						
	②치료재료	원						
	③진료비용	원						
V. 보험 등 100분의100 본인부담	①의약품	원						
	②치료재료	원						
	③진료비용	원						
W. 비급여	①의약품	원						
	②치료재료	원						
	③진료비용	원						
U. 건강보험100분의100본인부담	①의약품	원						
	②치료재료	원						
	③진료비용	원						
V. 보험 등 100분의100 본인부담	①의약품	원						
	②치료재료	원						
	③진료비용	원						
W. 비급여	①의약품	원						
	②치료재료	원						
	③진료비용	원						
특수장비총액	원							
	원							
19. 청구액	원							

(별지 제11호 서식)

안과

서식번호	G I O 3				요양급여비용명세서		요양기관	
	종류번호		상행번호		구분	내역	내역	내역
가입자성명	상행번호				상행구분	요양기관명		내역
수진자성명	상행번호				상행구분	요양기관명		내역
상행번호	내분기	수술	신원	의안	기	의안	의안	의안
본인 부담액 (I)								
본인 부담액 (II)								
1. 진찰료 (의안) (보험)	①초진	원						
	②회진	원						
3. 투약료 및 처방료	③의약품	원						
	④유합	원						
	⑤일반	원						
	⑥중환자실	원						
	⑦진단실	원						
	⑧기타	원						
	⑨가산실	원						
4. 주사료	①처방	원						
	②의약품	원						
5. 마취료	③④⑤⑥⑦⑧⑨⑩⑪⑫⑬⑭⑮⑯⑰⑱⑲⑳㉑㉒㉓㉔㉕㉖㉗㉘㉙㉚㉛㉜㉝㉞㉟㊱㊲㊳㊴㊵㊶㊷㊸㊹㊺㊻㊼㊽㊾㊿	원						
	①	원						
	②	원						
	③	원						
6. 이환요별료	④	원						
7. 경신요별료	⑤	원						
8. 처치 및 수술료	⑥⑦⑧⑨⑩⑪⑫⑬⑭⑮⑯⑰⑱⑲⑳㉑㉒㉓㉔㉕㉖㉗㉘㉙㉚㉛㉜㉝㉞㉟㊱㊲㊳㊴㊵㊶㊷㊸㊹㊺㊻㊼㊽㊾㊿	원						
9. 검사료	①②③④⑤⑥⑦⑧⑨⑩⑪⑫⑬⑭⑮⑯⑰⑱⑲⑳㉑㉒㉓㉔㉕㉖㉗㉘㉙㉚㉛㉜㉝㉞㉟㊱㊲㊳㊴㊵㊶㊷㊸㊹㊺㊻㊼㊽㊾㊿	원						
10. 영상진단 및 방사선 치료료	①②③④⑤⑥⑦⑧⑨⑩⑪⑫⑬⑭⑮⑯⑰⑱⑲⑳㉑㉒㉓㉔㉕㉖㉗㉘㉙㉚㉛㉜㉝㉞㉟㊱㊲㊳㊴㊵㊶㊷㊸㊹㊺㊻㊼㊽㊾㊿	원						
S. 특수장비	①CT	원						
	②MRI	원						
	③PET	원						
A. 100분의100미만 본인부담 1	①의약품	원						
	②치료재료	원						
	③진료비용	원						
B. 100분의100미만 본인부담 2	①의약품	원						
	②치료재료	원						
	③진료비용	원						
U. 건강보험100분의100본인부담	①의약품	원						
	②치료재료	원						
	③진료비용	원						
V. 보험 등 100분의100 본인부담	①의약품	원						
	②치료재료	원						
	③진료비용	원						
W. 비급여	①의약품	원						
	②치료재료	원						
	③진료비용	원						
U. 건강보험100분의100본인부담	①의약품	원						
	②치료재료	원						
	③진료비용	원						
V. 보험 등 100분의100 본인부담	①의약품	원						
	②치료재료	원						
	③진료비용	원						
W. 비급여	①의약품	원						
	②치료재료	원						
	③진료비용	원						
특수장비총액	원							
	원							
17. 처방금	원							
18. 장애인료	원							
19. 청구액	원							
20. 본인부담총액(보험)	원							
21. 요양급여비용총액 2, 진료비용	원							
22. 보험청구액	원							
23. 건강보험 100분의100본인부담총액	원							

보건의료빅데이터개방시스템

- “공공데이터의 제공 및 이용 활성화에 관한 법률”
 - 공공데이터에 개인정보 등 비공개대상정보가 포함된 경우 (공공기관의 정보공개에 관한 법률) 제공 제한
 - 주민등록번호, 성명, 주소, 전화번호, 요양기관기호(명칭)등 개인정보와 개별 법인·단체 등의 정보가 식별 불가능한 형태로 제공
 - 개인정보 등 비공개대상정보를 기술적으로 분리할 수 있는 경우 개인정보 등 비공개대상정보를 제외한 공공데이터는 부분제공
 - 공공데이터의 정보가 다른 정보와 결합하여 개인 식별이 가능한 경우 제공 제한
- 제공자료 내역
 - 진료정보, 의약품정보, 치료재료정보, 의료자원정보

구분	주요 정보	DB명칭
진료정보	청구명세서 정보	청구명세서 DB
	의료행위 정의 및 환자 분류 정보	정보분석 DB
	수가마스터 정보	통계자료 DB
	질병군(DRG), 요양병원 수가 마스터정보	질병통계 DB
	의료행위별 심사기준 정보(보장 범위) 의료행위 18개 분류별 진료규모 정보(진료량, 금액) 질병군(DRG) 및 요양병원 진료규모 정보 질병정보(주상병) 및 질병단위 진료규모 정보	행위통계 DB
의약품 정보	급 여 의 약 품 마 스 터 정 보 급여의약품 사용 정보	의약품DB
치료재료 정보	치료재료 마스터 정보 치료재료별 사용정보 특수 재료관련 정보(복강경시술, 조영제 등)	치료재료 DB
의료자원 정보	요양기관 개폐업 정보	요양기관현황 DB
	의료기관 시설 정보(병상, 집중치료실, 수술실 등) 인력(의사, 간호사, 의료기사 등) 현황 장비 보유현황(장비별 이력관리)	요양기관 종합정보 DB

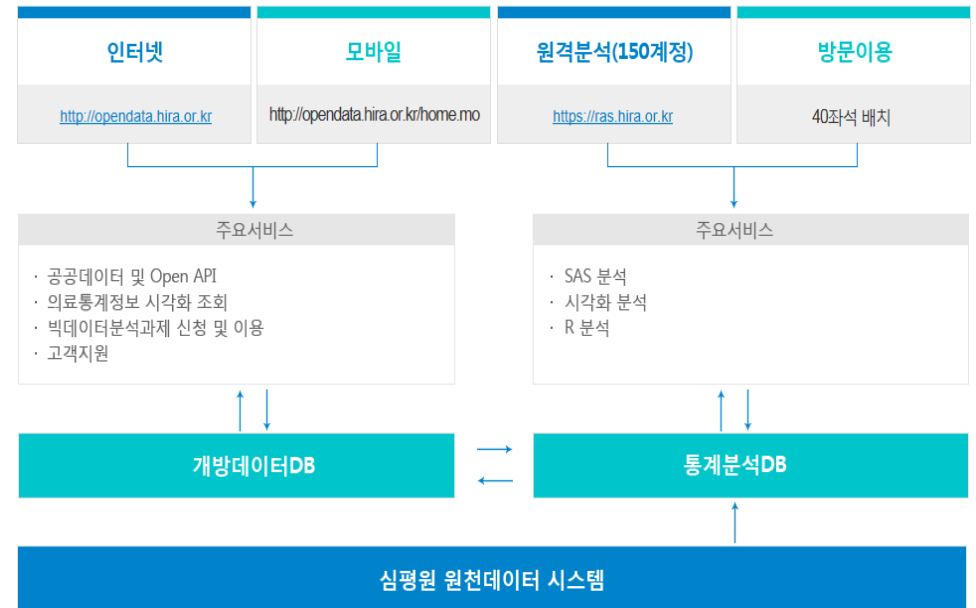
• 자료제공방법

- 환자데이터셋

구분	설명
입원 환자 데이터셋 (HIRA-NIS)	- 2009년, 2010년, 2011년, 2012년, 2013년, 2014년, 2015년 데이터 제공 - 입원환자 추출비율 13% (약 70만 명) - 외래환자 추출비율 1% (약 40만 명)
전체 환자 데이터셋 (HIRA-NPS)	- 2010년, 2011년, 2012년, 2013년, 2014년, 2015년 데이터 제공 - 전체환자 추출비율 3% (약 140만 명)
고령 환자 데이터셋 (HIRA-APS)	- 2010년, 2011년, 2012년, 2013년, 2014년, 2015년 데이터 제공 - 고령환자(65세 이상) 추출비율 20% (약 100만 명)
소아청소년환자데이터셋 (HIRA-PPS)	- 2010년, 2011년, 2012년, 2013년, 2014년, 2015년 데이터 제공 - 소아청소년환자(20세 미만) 추출비율 10% (약 110만 명)

• 보건의료빅데이터개방시스템

- 개방데이터 DB (인터넷 및 모바일을 통한 데이터 및 통계자료 조회 서비스)
- 통계분석 DB를 이용한 원격접속시스템 및 보건의료빅데이터센터(이하 '빅데이터센터')



자료신청과 분석단계

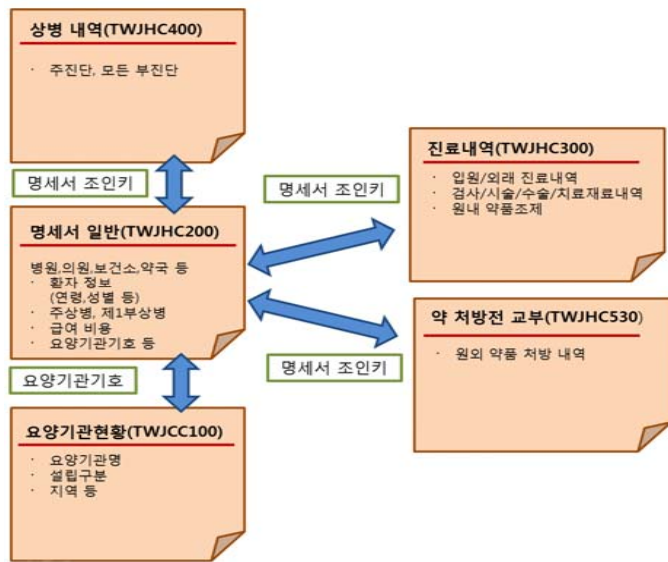


• 자료신청 주의사항

- 분석대상자에 대한 정의가 필요
- 각 자료원을 연구자가 직접 찾아야 함
 - 보건의료개방시스템 홈페이지
 - 산업계 위탁과제는 원칙적으로 자사 제품으로 한정함: 타사 자료 신청 동의를 구하는 경우 예외. 원칙적으로 주성분 코드가 4개 이상인 약품군에서 제공

구분	참고 파일
상병코드	상병 마스터파일
약품코드	약가 파일
행위코드	건강보험요양급여비용 파일
치료재료코드	치료재료 목록 파일

건강보험 청구자료의 구성



한글테이블명	영문테이블명	주요 조사내용
명세서일반내역 테이블	T200(TWJHC200)	환자 및 기관의 기본정보, 기본적인 상병(주상병명과 첫 번째 부상병명), 내원일수, 요양개시일자, 특정기호구분 등의 일반적 현황
진료내역 테이블	T300(TWJHC300)	요양기관 내에서의 처치, 수술, 원내 약처방, 청구금액 등 상세 의료서비스 현황
수진자상병내역 테이블	T400(TWJHC400)	T200의 주상병과 부상병을 포함한 추가 부상병명, 일련번호 '1' 은 T200의 주상병과 동일하며, 일련번호 '2' 는 T200의 부상병과 동일함
원외처방전상세내역 테이블	T530(TWJHC530)	원외로 처방되는 약에 대한 정보가 있는 테이블(원내 처방내역은 T300에 포함되어 있음). 처방된 약마다 한 줄로 구성 되어 KEY값 변수를 기준으로 동일명세서는 여러 줄로 되어있고 줄번호 변수가 있음. 이때 줄번호는 T300에 있는 줄번호와 다름
요양기관 테이블	YKIH0(TWJHC100)	요양기관의 지역, 종별, 설립구분 등 요양기관에 대한 일반적 정보

자료이용 시 주의사항

- 상병코드의 정확성
- 동일 명세서 내에서의 검사, 처치, 약물처방 등의 시간적 순서
- 동일입원의 분리청구
- 포괄수가, 정액제
- 분류비돌림: 처방받아도 청구되지 않는 경우, 처방되어도 사용되지 않는 경우 등
- 연구기간 중 급여 제도 변화
- 연구기간 중 수가, 약품 등 변화
- 분석자간 상호 교차검정

수가코드		
스텐트 시술 코드	경피적 관상동맥 스텐트 삽입술	M6561,M6562,M6563,M6564
치료재료코드	금속스텐트	J5083001,J5083011,J5083016,J5083029,J8083033,J5083040,J5083046,J5083070,J5083094,J5083101,J5083106,J5083107,J5083137,J5083140,J5083173,J5083201,J5083206,J5083213,J5083240,J5083273,J5083301,J5083401,J5083501,J5084040,J5084140,J8083006,J8083073,J8083089,J8083129,J8083340,J8083373,J8083473
	약물방출스텐트	J5231024,J5231061,J5231067,J5231071,J5231081,J5231089,J5231103,J5231107,J5231194,J5231221,J5231294,J5231306,J5231313,J5231525,J5231625,J8231061,J5231313,J5231525,J5231625,J8231061

수가코드		치료재료코드	
		무	유
스텐트 시술 코드	무	135,866,525	181
	유	206,978	1,692

정신과 연구 주의점

- 민감상병 제공 제한: 표본자료
- 서식구분코드 중 [정신과] 해석 주의
 - 서식구분코드 중 정신과라 함은 의료급여 정신과 정액을 의미함
 - 2008년 정신과 수가제도가 개선됨에 따라, 2009년 1월 진료분부터 의료급여 정신질환 정액수가에 수가코드(09:의료급여 정신과정액 낮병동, 10:의료급여 정신과정액 입원, 11:의료급여 정신과정액 외래)를 부여하고 명세서를 별도 작성하여 청구하도록 개정됨

명세서 id 의 활용

- 자료 분석을 위해서는 아래 테이블을 처방전 단위로, 또는 환자 단위로 연계가 필요
 - 1) 처치 (30)
 - 2) 진단 (40)
 - 3) 처방 (53)
- 20 명세서 테이블은 나머지 테이블과 연계되어 분석을 가능하게 해줌

spec_id_sno	MSIC_CD	SSICK_CD
9801	E1150	1109

=>

SPEC_ID_SNO	DIV_CD
9801	AA254030
9801	AA254
9801	AH200
9801	AL801
9801	KK010
9801	C3710

SPEC_ID_SNO	SICK_CD
9801	M1399
9801	E782
9801	C1699
9801	J459
9801	J329

SPEC_ID_SNO	GNL_NM_CD
9801	125201AC
9801	268000AT
9801	438901AT
9801	614100AS

질병 변수 생성

질환	KCD 상병코드
전체 암(Cancer)	C00-C97
당뇨병(Diabetes melitus)	E10-E14
고혈압(Hypertension)	I10-I15
천식(Asthma)	J45-J46
관절증(Arthrosis)	M15-M19
류마티스 관절염(Rheumatoid arthritis)	M05-M06
신부전증(Renal failure)	N17-N19
뇌혈관 질환(Cerebrovascular diseases)	I60-I69
이상지질혈증(Dislipidemia)	E78
허혈성 심장질환(Ischaemic heart diseases)	I20-I25

T200

SPEC_ID_SNO	MSICK_CD	SSICK_CD
10039919	C189	K550

↓
주상병과 제1 부상병

T400

SPEC_ID_SNO	SICK_CD
10039919	C189
10039919	K550
10039919	I109
10039919	K650
10039919	N179

↓
주상병과 모든 부상병

```

/* T400 테이블에서 질병 여부 변수 생성*/
data t40_1;
set hira.t40;
/*T40의 모든 상병코드에 대해 대상질병 변수 생성*/
if 'C00' <= substr(SICK_CD,1,3) <= 'C97' then cancer=1; /* 전체 암 */
if substr(SICK_CD,1,3) in ('E10', 'E11', 'E12', 'E13', 'E14') then diabetes=1; /* 당뇨병 */
if substr(SICK_CD,1,3) in ('I10', 'I11', 'I12', 'I13', 'I14', 'I15') then hypertension=1; /* 고혈압 */
if substr(SICK_CD,1,3) in ('J45', 'J46') then asthma=1; /* 천식 */
if substr(SICK_CD,1,3) in ('M15', 'M16', 'M17', 'M18', 'M19')
then arthrosis=1; /* 관절증 */
if substr(SICK_CD,1,3) in ('M05', 'M06')
then rheumatoid_arthritis=1; /* 류마티스 관절염 */
if substr(SICK_CD,1,3) in ('N17', 'N18', 'N19')
then renal_failure=1; /* 신부전증 */
if substr(SICK_CD,1,3) in ('I60', 'I61', 'I62', 'I63', 'I64', 'I65',
'I66', 'I67', 'I68', 'I69') then Cerebrovascular=1; /* 뇌혈관 질환 */
if substr(SICK_CD,1,3) in ('E78') then dislipidemia=1; /* 이상지질혈증 */
if substr(SICK_CD,1,3) in ('I20', 'I21', 'I22', 'I23', 'I24', 'I25')
then Ischemic=1; /* 허혈성 심장질환 */

```

```
run;
```

	SPEC_ID_SNO	SICK_SNO	SICK_DGSBJT_CD	SICK_CD	DWD_DGSBJT_CD	cancer	diabetes	hypertension	asthma	arthrosis
29798339	17358194	1	01	R05	01
29798340	17358194	6	01	J459	01	.	.	.	1	.
29798341	17358194	4	01	J42	01
29798342	17358194	5	01	M79108	01
29798343	17358194	3	01	K297	01
29798344	17358194	2	01	J304	01
29798345	17358197	1	05	S9348	05
29798346	17358197	2	05	M7977	05
29798347	17361613	1	01	E785	01
29798348	17361614	4	01	I748	01
29798349	17361614	1	01	E119	01	.	1	.	.	.
29798350	17361614	3	01	E785	01
29798351	17361614	2	01	K769	01
29798352	17361615	1	01	I109	01	.	.	1	.	.
29798353	17362069	2	05	M5456	05
29798354	17362069	3	05	Z5180	05
29798355	17362069	1	01	R521	01

의약품 사용 변수 생성

T300

SPEC_ID_SNO	DIV_CD	AMT	GNL_NM_CD
10000007	AA254011	12190	\$
10000007	AL825	1810	\$
10000007	\$	1035	111402AT
10000007	\$	4050	459801AT
10000007	\$	2678	235502AT

↓

원내 세부 진료 내역

T530

SPEC_ID_SNO	AMT	GNL_NM_CD
10000008	2440	111001AT
10000008	19520	185701AT

↓

원외처방전상세내역

```

data drug_table_30;
set hira.t30;
in_out=1; /*원내외처방 구분변수 1=원내, 2=원외*/
where gnl_nm_cd ne '$'; /*약처방 외의 추가행위를 제외하기 위해 주성분코드가$ 인 줄 제거*/
TOT_USE=TOT_USE_QTY_OR_EXEC_FQ; /*T300의 사용량 변수와 T530의 사용량 변수의 변수 명 이 다르기 때문에 하나의 변수로 통합*/
keep spec_id_sno gnl_nm_cd amt TOT_USE in_out;
run;

/*원외 의약품 사용 변수 생성 - 약물 및 실적(사용량, 금액) 추출 */
data drug_table_53;
set hira.t53;
in_out=2; /*원내외처방 구분변수 1=원내, 2=원외*/
TOT_USE=TOT_USE_QTY_EXEC_FQ;
keep spec_id_sno gnl_nm_cd amt TOT_USE in_out;
run;
    
```

T300
명세서 조인키(SPEC_ID_SNO), 주성분코드(GNL_NM_CD), 사용량(TOT_USE_QTY_OR_EXEC_FQ), 금액(AMT)

↑ set

T530
명세서 조인키(SPEC_ID_SNO), 주성분코드(GNL_NM_CD), 사용량(TOT_USE_QTY_EXEC_FQ), 금액(AMT)

⇒

의약품 테이블
명세서 조인키(SPEC_ID_SNO), 주성분코드(GNL_NM_CD), 사용량(TOT_USE), 금액(AMT)

/*의약품 사용 변수 생성 - 의약품 테이블 생성*/

```

data drug_table;
set drug_table_30 drug_table_53;
run;
    
```

	SPEC_ID_SNO	AMT	GNL_NM_CD	in_out	TOT_USE
1	9801	30	131801AT	1	15
2	9801	266	242303BI	1	1
3	9803	30	131801AT	1	15
4	10095	266	242303BI	1	1
5	10097	18	131801AT	1	9
6	10098	266	242303BI	1	1
7	13558	12	131801AT	1	6
8	13564	12	131801AT	1	6
9	13568	30	131801AT	1	15
10	14513	313	332700CO	1	1
11	14541	147	131802BI	1	1
12	14542	499	223802BI	1	1
13	14552	499	223802BI	1	1

질병 및 의약품 사용 통합 테이블 생성

t2040_yk
명세서 조인키(SPEC_ID_SNO), 요양기관대체키(YID), 개인식별대체키(JID), 성별(SEX_TP_CD), 연령군(AGE_CAT), 요양개시일자(RECU_FR_DD), 질병여부변수(CANCER, ...)

← left join
(명세서조인키)

drug_table
명세서조인키(SPEC_ID_SNO), 주성분명(GNL_NM), 금액(AMT) 사용량(TOT_USE)

```

proc sql;
create table t2040_yk_drug as
select a.* , b.spec_id_sno, b.gnl_nm, b.amt, b.tot_use
from T2040_yk as a
left join hira.drug_table as b
on (a.spec_id_sno = b.spec_id_sno);
quit;
    
```

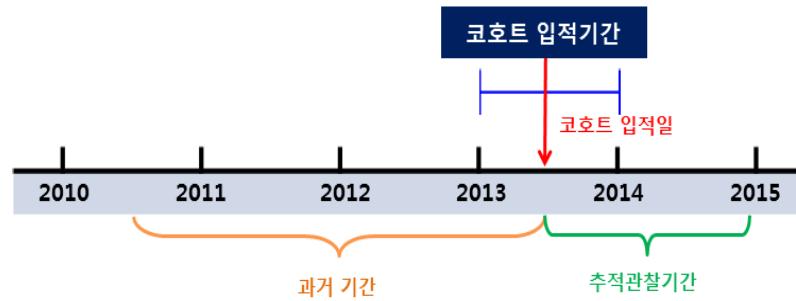
	SPEC_ID_SNO	JID	SEX_TP_CD	MSICK_CD	SSICK_CD	DGSBJT_CD	RECU_FR_DD	YID	SIDO_CD	CL_CD	AMT	gnl_nm	TOT_USE	cancer	diabetes
1	1	1063294	2	D259	K660	10	20151223	6150	37	11
2	10	460288	1	K352		04	20151221	2765	38	11
3	100	870968	2	E119	I109	01	20151207	9819	21	31	1
4	1000	1024172	1	H2592		12	20151231	32523	23	31
5	10000	685389	2	I100		01	20150916	45080	34	31
6	100000	1370594	2	N72	N762	15	20151217	50452	11	31
7	1000000	627510	1	J029		01	20151212	51434	26	31
8	10000000	36014	1	J209	J042	23	20150203	34938	11	31
9	10000001	1266320	1	N185	E211	01	20150516	19294	33	31
10	10000002	1266320	1	N185	E211	01	20150518	19294	33	31
11	10000003	1266320	1	N185		01	20150521	19294	33	31
12	10000004	1266320	1	N185	E211	01	20150523	19294	33	31
13	10000005	1266320	1	N185	E211	01	20150525	19294	33	31
14	10000006	1266320	1	N185	B353	01	20150528	19294	33	31
15	10000007	1266320	1	N185	E211	01	20150530	19294	33	31	6240	olmesartan	15	.	.
16	10000008	122345	1	I109	I259	01	20150526	19294	33	31	19520	losartan	40	.	.

청구자료를 이용한 성과연구

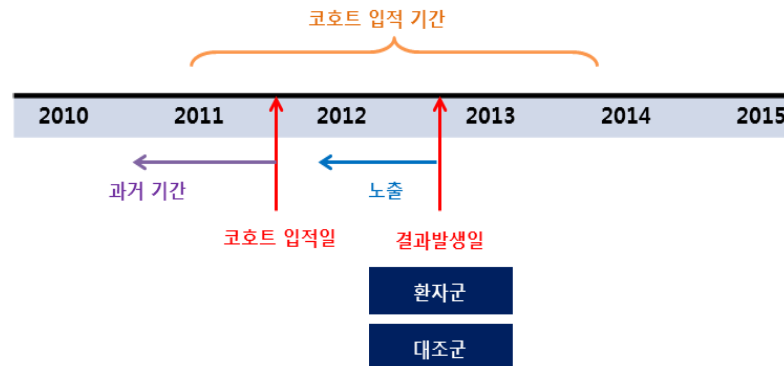
- 연구에서 관심있는 약물 또는 의료기술이 모두 급여대상
- 임상적 효과를 비교할 수 있는 의미있는 결과변수가 청구자료에서 정의 가능
- 이 때 비교하고자 하는 약물 또는 의료기술간의 비교가능성을 확보할 수 있으면 real world 에서 효과를 검증 가능함
 - 주로 임상적 효과를 비교할 때 바이어스를 최소화하기 위해 인과성 (causal effect)을 판단할 수 있는 코호트 연구 또는 환자-대조군 연구 등의 잘 설계된 연구설계 방법을 적용
 - 청구자료에서 파악할 수 있는 관찰된 교란요인들(confounders)을 보정하기 위해 적절한 통계분석 방법을 적용

연구설계

- 코호트 연구



- 환자-대조군 연구



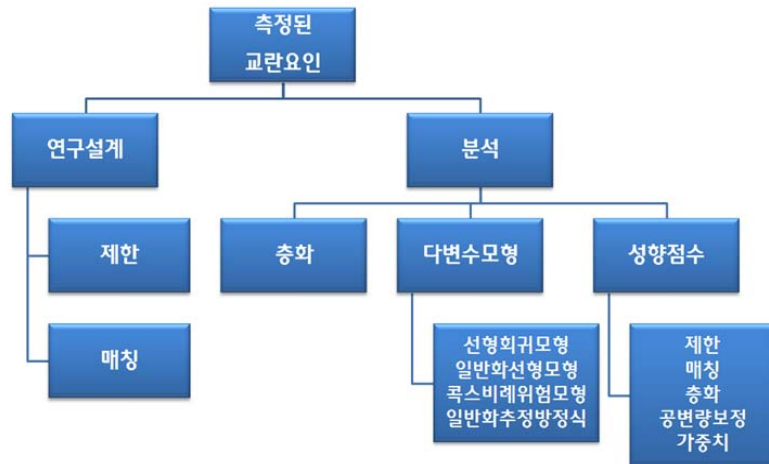
연구대상자 선정

- 청구자료를 대상으로 분석대상 치료법과 비교대상 치료법의 적응증을 가지는 연구대상자를 선정하기 위해서 선정/제외 기준 (inclusion/exclusion criteria)을 정의
- 입적 기간 내에 치료법에 대한 청구가 있는 대상자를 선정하고, 사전에 정한 제외기준에 해당하는 대상자를 제외하여 치료법의 적응증을 가지는 연구대상자로 제한(restriction)
 - 치료법의 관심 결과를 과거에 경험한 대상자를 제외하여 치료법에 대한 신환자를 연구대상자로 정의하고자 할 경우 과거력 기간에 해당 치료법에 대한 청구가 있었던 경우를 과거력으로 정의하여 제외
 - 연구대상자의 선정 및 제외 기준은 적응증의 특성에 따라 상병코드 또는 처치코드, 약물코드 등으로 정의하고, 필요한 경우 이들 코드들의 조합으로 정의

변수 정의

- 분석대상 치료법 및 비교대상 치료법, 결과변수, 교란요인을 정의하기 위하여 상병코드, 약물코드, 처치코드, 재료코드 등을 이용하여 조작적 정의
 - 선행연구 및 전문가 의견을 참고, 근거를 설명필요
 - 분류 비뉘림에 유의
- 결과변수
 - 질병의 발생: 조작적 정의 가능
 - 병원내 사망인 경우 진료결과변수 또는 상병코드를 이용하여 정의
 - 의료기관이 아닌 장소에서 사망한 경우는 청구자료로 파악불가: 연계!
 - 추적관찰 종료 시점은 연구가설 및 결과변수 특성에 따라 마지막 방문일 또는 가용한 청구자료의 마지막 일자

통계분석



적용 가능한 상황	
설계적 측면	
제한	특성이 동질한 연구대상자를 선정하여 치료효과 추정
매칭	교란요인의 분포가 동일하거나 비슷한 대상자를 매칭하여 치료효과 추정
분석적 측면	
층화	교란요인 또는 효과변경인자 수준별로 치료효과 추정
다변수 분석	교란요인을 보정하여 치료효과 추정
성향점수	비교군들간 균형을 맞추어 치료효과 추정

결과 해석 시 고려할 점

- Large sample size effect
- 통계적 유의성 뿐 아니라 임상적인 의미에 대한 해석 필요

이차자료원을 이용한 연구 체크리스트

- STROBE(Strengthening the Reporting of Observational studies in Epidemiology)
- RECORD(REporting of studies Conducted using Observational Routinely collected health Data)

1. 데이터가 올바른가?
 - 정확성 : 데이터가 진실을 반영하는가?
2. 데이터가 완전한가?
 - 완전성 : 데이터에 수집된 모든 기록이 포함되는가?
 - 포괄성과 적용 범위 : 데이터가 목표한 연구대상자의 100%를 포함하는가?
3. 데이터를 신뢰할 수 있는가?
 - 신뢰성 : 데이터가 재현 가능한가?
 - 유효성 : 데이터가 의미가 있는가?
4. 데이터를 사용할 수 있는가?
 - 익명성 : 데이터가 사생활 보호, 절차 및 관행을 준수하는가?
 - 연결가능성 : 건강 관리 시스템의 복잡성을 반영하기 위하여 다른 데이터와 연결 할 수 있는가?
 - 적시 : 데이터 수집과 사용 사이에 약간의 시간차가 있는가?
 - 유용성 : 데이터를 쉽게 구성하고 액세스 할 수 있으며 쉽게 사용할 수 있는 형식으로 제공되는가?

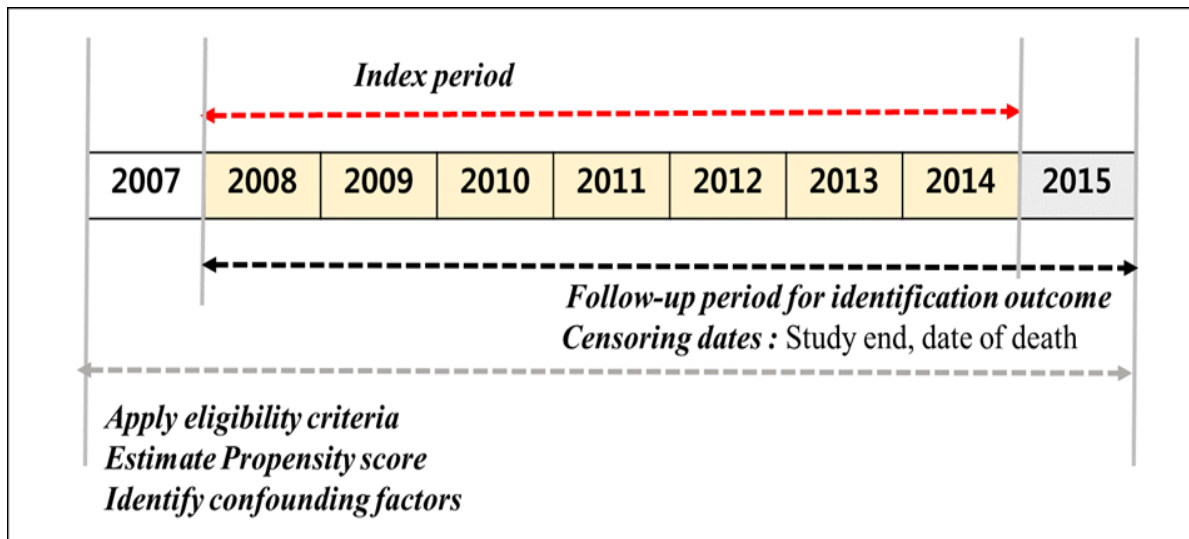
분석예시_1 RA and Osteoporotic fx

- 류마티스 관절염 (Rheumatoid Arthritis; RA)환자에서 전통적인 RA 치료제 (Disease Modifying Anti-Rheumatic Drug; DMARD)와 생물학적인 DMARD인 종양괴사인자-알파 억제제 (Tumor necrosis factor-alpha inhibitor; TNFi)에 사용에 따른 고관절 골절의 발생율을 비교
- 참고논문: Effects of disease-modifying antirheumatic drugs on nonvertebral fracture risk in rheumatoid arthritis: a population-based cohort study. Journal of bone and mineral research : the official journal of the American Society for Bone and Mineral Research. 2012;27(4):789-96. Epub 2011/12/14.

자료명세

구분	내용
진료기간	2007.01.01 ~ 2015.12.31 (9개년도)
요양기관 종별	상급종합병원, 종합병원, 병원, 요양병원, 의원, 정신요양병원, 보건소, 보건지소, 보건의료원
보험자 종별	건강보험, 의료급여
대상자	RA 환자
청구자료 상세내역	
테이블	내용
T200 명세서일반내역	명세서 요약정보
T300 진료내역	세부처치 및 원내 처방내역
T400 수진자 상병내역	상병정보
T530 원외처방전 상세내역	외래 처방 약제 정보

연구설계



연구대상

구분	내용
선정기준	<ul style="list-style-type: none">• 2007-2015 RA 환자 중 2008-2014 기간 중 처음으로 DMARD 처방이 확인된 환자
제외기준	<ul style="list-style-type: none">• tDMARD로 monotherapy 처방만 이루어진 환자• 전체 코호트 기간 중 종양, 간부전, 신부전, 후천성 면역결핍증의 진단으로 인한 청구가 있는 환자• 장기 이식을 시행 받은 환자• cyclophosphamide를 투여받은 환자• 코호트 입적일 이전에 대퇴골절로 인한 청구가 있는 환자

약물 목록

No	DMARD	ATC code
1	adalimumab	L04AB04
2	etanercept	L04AB01
3	infliximab	L04AB02
4	golimumab	L04AB06
5	mycophenolate mofetil	L04AA06
6	azathioprine	L04AX01
7	bucillamine	M01CC02
8	d-penicillamine	M01CC01
9	hydroxychloroquine	P01BA02
10	leflunomide	L04AA13
11	minocycline	A01AB23
12	mizoribine	L04AX
13	auranofin	M01CB03
14	sulfasalazine	A07EC01
15	thioguanine	L01BB03
16	methotrexate	L04AX03
--	-	-----

결과변수

- S72.0 fracture of neck of the femur
- S72.1 pertrochanteric fracture
- S72.2 subtrochanteric fracture
- 사망
 - 진료결과구분코드가 사망
 - 급성 심장사로 기술된 것(I461)
 - 원인 미상의 기타 급사(R96)
 - 순간적 사망(R960)
 - 증상의 발생으로부터 24시간 이내에 일어난 달리 설명되지 않는 사망(R961)
 - 지켜본 사람이 없었던 사망(R98)
 - 기타 불명확하고 상세불명의 사망원인(R99)

공변량

구분	내용
Demographic factors	Age sex Osteoporosis diagnosis (M80, M81) Use of bisphosphonate
Osteoporosis related factors	Prior fracture at forearms, upper arms, pelvis or vertebra (S32, S42, S43) Parkinson's disease (G20) Demetia (F00-F03) Cumulative dose of oral glucocorticoids, mean (SD)
Medication use	Anticonvulsants Benzodiazepines SSRIs Proton pump inhibitors Beta blockers Opioids
Health care utilization	Number of all prescription drugs, mean (SD) Hospitalizations Number of physician visits, mean (SD)

구분	내용
Bisphosphonate	etidronic acid , clodronic acid, pamidronic acid, alendronic acid , ibandronic acid, risedronic acid, zoledronic acid
Barbiturates and derivatives	methylphenobarbital, phenobarbital, primidone, barbexalone, metharbital ethotoin, phenytoin,
Hydantoin derivatives	amino(diphenylhydantoin) valeric acid, mephenytoin, fosphenytoin
Oxazolidine derivatives	paramethadione. trimethadione. ethadione
Succinimide derivatives	ethosuximide, phenosuximide, mesuximide
Benzodiazepine derivatives	clonazepam
Anticonvulsants	Carboxamide derivatives Fatty acid derivatives
	carbamazepine, oxcarbazepine, rufinamide, eslicarbazepine valproic acid, valpromide, aminobutyric acid, vigabatrin, progabide, tiagabine
	sultiam, phenacemide, lamotrigine, felbamate, topiramate, gabapentin, pheneturide, levetiracetam, zonisamide, pregabalin, stiripentol, lacosamide, carisbamate, retigabine, peramppanel, brivaracetam, beclamide
	morphine, opium, hydromorphone, nicomorphine, oxycodone, dihydrocodeine, papaveretum
	Natural opium alkaloids
	Phenylpiperidine derivatives
	ketobemidone, pethidine, fentanyl dextromoramide, piritramide, dextropropoxyphene, bezitramide, methadone
Opioids	Diphenylpropylamine derivatives Benzomorphan derivatives Oripavine derivatives Morphinan derivatives
	Other opioids
	flurazepam, nitrazepam, flunitrazepam, estazolam, triazolam, lormetazepam, temazepam, midazolam, brotizolam, quazepam, lorazepam, doxefazepam, cinolazepam
Benzodiazepines	zimidine, fluoxetine, citalopram, paroxetine, sertraline, alaproclate, fluvoxamine, etoperidone, escitalopram
Selective serotonin reuptake inhibitors	esomeprazole, pantoprazole, rabeprazole, omeprazole, lansoprazole
Proton pump inhibitors	amosulalol, atenolol, betaxolol, bisoprolol, evantolol, carteolol, celiprolol, metoprolol, nebivolol
Beta blockers	

스테로이드 코드 목록 및 등가용량표

code	generic name	dose (mg)	Equivalent dose of hydrocortisone (mg)
316100BJ	betamethasone	9	270
116502BJ	betamethasone	4	120
296900ATB	betamethasone	0.25	7.5
116401ATB	betamethasone	0.5	15
116501ATB	betamethasone	0.5	15
141901ATB	dexamethasone	0.5	12.5
141903ATB	dexamethasone	0.75	18.75
141904ATB	dexamethasone	4	100
142001BJ	dexamethasone	4	100
142201BJ	dexamethasone	5	125
142202BJ	dexamethasone	4.37	109.25
142203BJ	dexamethasone	4	100
243201ATB	triamcinolone	1	5
243202ATB	triamcinolone	2	10
243203ATB	triamcinolone	4	20
243301BJ	triamcinolone	200	1000
243303BJ	triamcinolone	40	200
243306BJ	triamcinolone	50	250
217001ATB	prednisolone	5	20
217003ASY	prednisolone	1	4
217004ASY	prednisolone	3	12
217101BJ	prednisolone	110	440
217104BJ	prednisolone	100	400
217301BJ	prednisolone	1000	4000
217302BJ	prednisolone	250	1000
217401ATB	prednisolone	6.65	26.6
193301ATB	methylprednisolone	16	80
193302ATB	methylprednisolone	4	20
193303ATB	methylprednisolone	8	40
193304ATB	methylprednisolone	2	10
193501BJ	methylprednisolone	200	1000
193502BJ	methylprednisolone	40	200
193601BJ	methylprednisolone	125	625
193602BJ	methylprednisolone	250	1250
193603BJ	methylprednisolone	40	200
193604BJ	methylprednisolone	500	2500
170901ATB	hydrocortisone	10	10
170905ATB	hydrocortisone	20	20
170906ATB	hydrocortisone	5	5

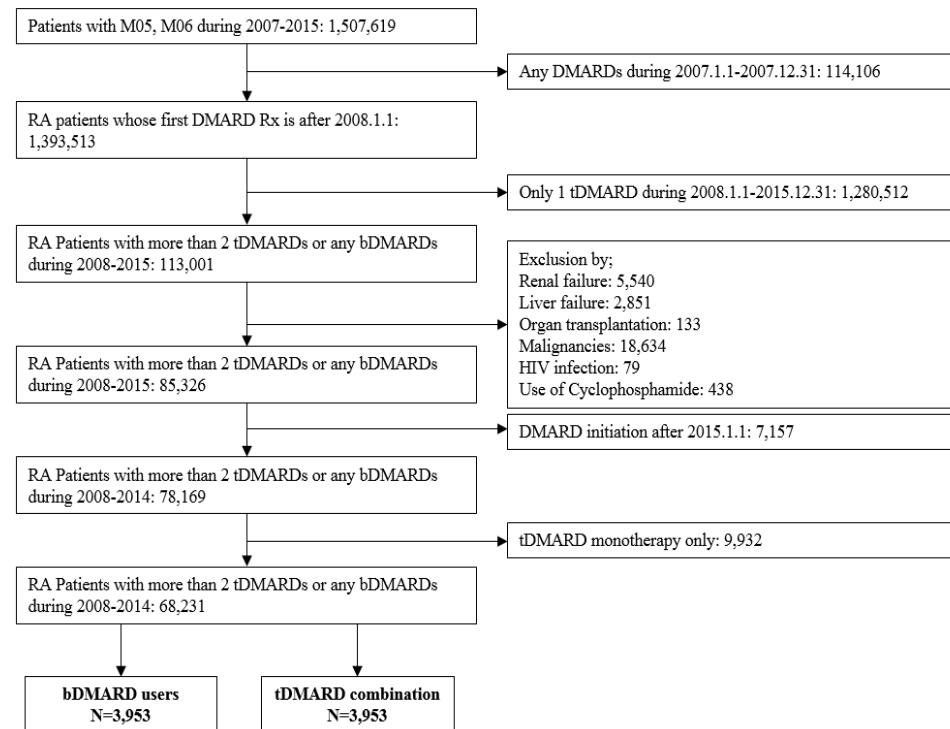
통계분석방법

- 성향점수 짝짓기법
 - 짝짓기비율: 1:1
 - 표준화 차이를 이용한 균형 평가
- 비례위험모형
- SAS® Enterprise Guide version 6.1 (SAS Institute, Cary, NC, USA)

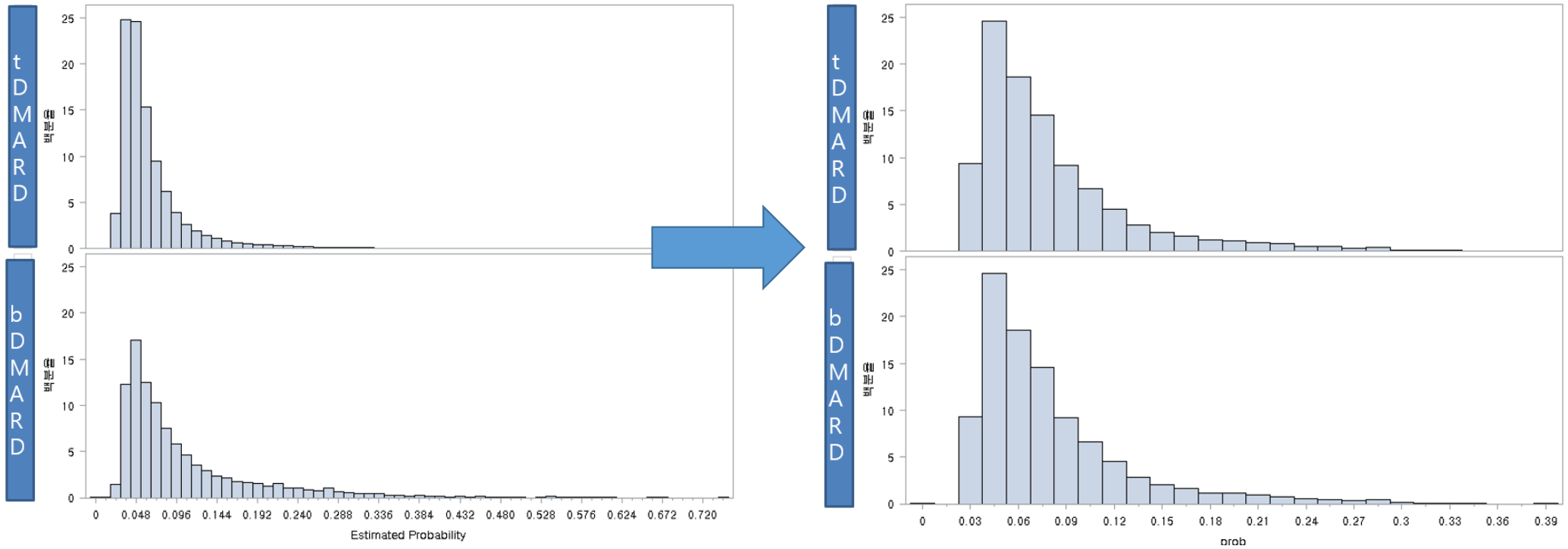
$$d = \frac{(\hat{p}_{bDMARD} - \hat{p}_{iDMARD})}{\sqrt{\frac{\hat{p}_{bDMARD}(1 - \hat{p}_{bDMARD}) + \hat{p}_{iDMARD}(1 - \hat{p}_{iDMARD})}{2}}}$$

$$d = \frac{(\bar{x}_{bDMARD} - \bar{x}_{iDMARD})}{\sqrt{\frac{(s_{bDMARD})^2 + (s_{iDMARD})^2}{2}}}$$

연구대상 선정



성향점수의 균형



결과

Treatment groups	HR	95% CI	p-value
tDMARD	Reference		
bDMARD by adjustment [†]	1.58	(1.16-2.16)	0.004
bDMARD by PSM [*]	1.40	(0.76-2.57)	0.29

1. 류마티스관절염 환자 선정

- T40 테이블에서 류마티스관절염 상병코드(M05, M06)를 가지는 행 선택

```
data t40_RA;  
  set hira.t40;  
  if substr(SICK_CD,1,3) in ('M05', 'M06') then RA=1; /* 류마티스 관절염 */  
run;
```

1. 류마티스관절염 환자 선정

- 각각의 명세서에 대한 류마티스관절염 상병변수 생성

proc sql;

create table t40_RA1 as

select spec_id_sno, sick_cd,
max(RA) as ra

from t40_ra

group by spec_id_sno;

quit;

1. 류마티스관절염 환자 선정

- T20 테이블에 T40에서 생성한 류마티스관절염 상병변수 조인

proc sql;

```
create table t2040_ra as
```

```
select a.*, b.*
```

```
from hira.t20 as a
```

```
left join t40_RA1 as b
```

```
on (a.spec_id_sno=b.spec_id_sno);
```

quit;

1. 류마티스관절염 환자 선정

- 류마티스 관절염(RA)환자들에 대한 명세서만 선택

```
data t2040_ra1;  
  set t2040_ra  
  where ra=1;  
run;
```

1. 류마티스관절염 환자 선정

위의 프로그램 결과 생성되는 테이블은 다음과 같다.

	SPEC_ID_SNO	ra	DMD_TP_CD	DMD_FRM_CD	FOM_TP_CD	JID	SEX_TP_CD	PAT_AGE
1	1000935	1 0	3	031	564944	1	41	
2	1000979	1 0	3	031	1412621	1	47	
3	1000980	1 0	3	031	1412621	1	48	
4	1000981	1 0	3	031	1412621	1	48	
5	1001417	1 0	3	031	103835	2	48	
6	1001418	1 0	3	031	103835	2	48	
7	1001529	1 0	3	031	145055	2	84	
8	1003570	1 0	3	031	183036	2	89	
9	1004473	1 0	3	031	359652	2	83	
10	1005098	1 0	3	031	915555	2	40	
11	1005099	1 0	3	031	915555	2	40	
12	1005890	1 0	3	031	988990	1	73	
13	1005891	1 0	3	031	1307912	2	62	

1. 류마티스관절염 환자 선정

※ 프로그램 Tip

① **if** substr(SICK_CD,1,3) in ('M05', 'M06') **then** RA=1

⇒ T40의 상병코드가 입력되어있는 'SICK_CD'변수에서 substr함수를 통해 해당 질병에 해당하는 코드가 있을 경우 해당 질병 변수에 1이라는 값을 넣어 변수를 생성

SICK_CD		SICK_CD		RA
M0609	substr(SICK_CD, 1, 3)	M06	then RA = 1	1
K602		K60		
M0590	⇒	M05	⇒	1
M0642		M06		1

1. 류마티스관절염 환자 선정

② **select** spec_id_sno, max(RA) as ra
group by spec_id_sno

⇒ 동일한 명세서 조인키를 가지는 경우 RA 변수가 가지는 값들 중 최대값을 RA 변수에 입력 하도록 하며, GROUP BY문을 이용해 동일한 여러 행(명세서)들을 모아 한줄로 요약

SPEC_ID_SNO	RA
41958	1
41958	1
9041	.
9041	.
9041	.

group by SPEC_ID_SNO

⇒

SPEC_ID_SNO	RA
9041	.
41958	1

2. 치료법 정의

- 치료법의 정의를 위하여 고려한 DMARD 목록에서 일부만 선택
(DMARD 처방은 T300 진료내역, T530 원외처방전 상세내역에서 DMARD 코드의 유무로 정의)

proc sql;

```
create table code.drug_master_RA as
```

```
select *
```

```
from drug_master
```

```
where
```

```
generic_nm like 'adalimumab%' or
```

```
generic_nm like 'etanercept%' or
```

```
generic_nm like 'infliximab%' or
```

```
generic_nm like 'golimumab%' ;
```

quit;

- T30테이블의 약품성분명 코드 중 일부 DMARD약물을 충족하는 명세서 추출

proc sql;

create table replace_30 **as**

select SPEC_ID_SNO,

sum(TOT_USE_QTY_OR_EXEC_FQ) **as** ra_drug_t30 */*동일 명세서 내 총사용량 합계 변수 생성*/*

from Hira.t30

where GNL_NM_CD **in** (**select** generic_code **from** code.drug_master_ra)

*/*DMARD약물을 충족하는 명세서 추출*/*

group by spec_id_sno;

quit;

- T53테이블의 약품성분명 코드 중 일부 DMARD약물을 충족하는 명세서 추출

proc sql;

```
create table replace_53 as
```

```
select SPEC_ID_SNO,
```

```
sum(TOT_USE_QTY_OR_EXEC_FQ) as ra_drug_t53 /*동일 명세서 내 총사용량 합계 변수 생성*/
```

```
from Hira.t53
```

```
where GNL_NM_CD in (select generic_code from code.drug_master_ra )
```

```
/* DMARD약물을 충족하는 명세서 추출*/
```

```
group by spec_id_sno;
```

quit;

- 위의 T30 T53테이블에서 추출한 명세서를 통합하여 DMARD약물을 투약한 명세서 추출

```
data drug_ra;
  merge replace_30 replace_53;
  by spec_id_sno;
  /*T30 T53 테이블의 약물 투약량을 sum하여 DMARD약물 투약량 변수 생성*/
  drug_ra_use = sum(ra_drug_t30, ra_drug_t53);
  if drug_ra_use > 0 then DMARD=1;
      /*DMARD약물 투약 실적이 있는 명세서 추출하여 RA치료법에 대한 테이블생성*/
  keep SPEC_ID_SNO DMARD;
run;
```

- 류마티스 (RA)환자에 대한 테이블에 DMARD변수 조인

proc sql;

```
create table t2040_ra_dmard as
```

```
select a.* , b.*
```

```
from t2040_ra1 as a left join drug_ra as b
```

```
/*t2040_ra1: 류마티스(RA)환자 테이블, drug_ra: DMARD 변수 테이블*/
```

```
on (a.spec_id_sno=b.spec_id_sno)
```

```
group by jid;
```

quit;

2. 치료법 정의

위의 프로그램 결과 생성되는 테이블은 다음과 같다.

	SPEC_ID_SNO	ra	DMARD	RECU_FR_DD	DMD_TP_CD	DMD_FRM_CD	FOM_TP_CD	JID
1	12719509	1	,	20150803	0	3	021	1000011
2	5959303	1	,	20150411	0	3	031	1000013
3	24937272	1	,	20151113	0	3	031	1000049
4	22914403	1	,	20150609	0	3	031	1000049
5	22914402	1	,	20150603	0	3	031	1000049
6	23829321	1	,	20150930	0	3	031	1000049
7	20806042	1	,	20150713	0	3	031	1000049
8	20806043	1	,	20150724	0	3	031	1000049
9	12535212	1	,	20150305	0	3	031	1000049
10	12535213	1	,	20150318	0	3	031	1000049

※ 프로그램 Tip

① **where**

generic_nm like 'adalimumab%' or

generic_nm like 'etanercept%' or

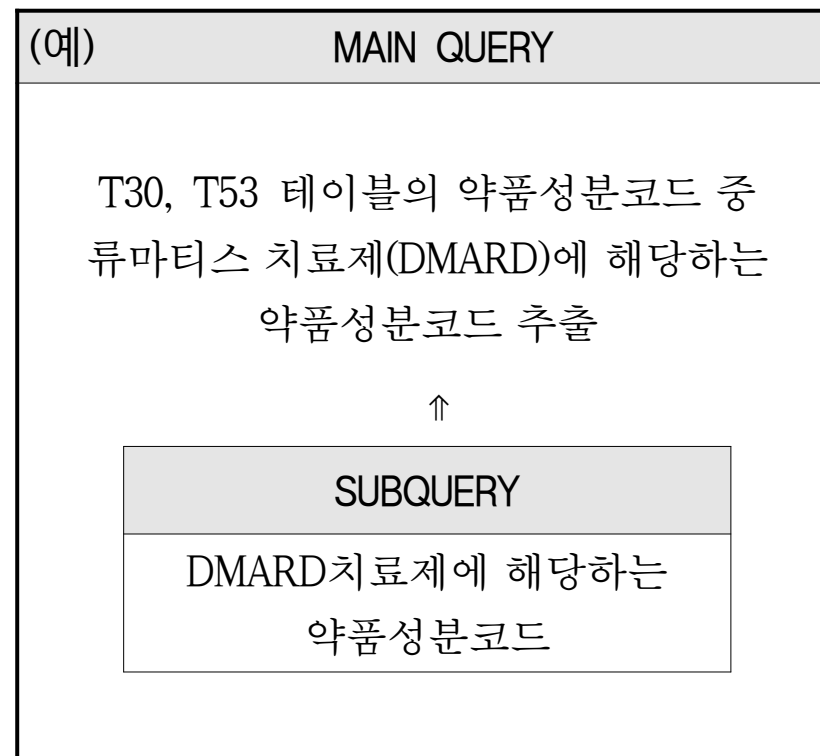
generic_nm like 'infliximab%' or

generic_nm like 'golimumab%'

⇒ like 'A %' 는 특정 자료에 대한 검색을 할 때 활용하는 함수이다. WHERE 이후에 대문자 A 로 시작하는 크기에 상관 없이 모든 문자를 검색하도록 한다.

② **where GNL_NM_CD in (select generic_code from code.drug_master_ra)**

- 서브쿼리(subquery) : WHERE, FROM, HAVING절에 다른 SQL문장을 포함하는 형태이다. 서브쿼리를 사용하면 불필요한 테이블을 생성하지 않고도 바로 원하는 테이블을 생성할 수 있기 때문에 코드를 간소화할 수 있다. 서브쿼리 사용시, 문장을 괄호로 둘러싸야 하며, ORDER BY 절은 포함할 수 없다.



③ `drug_ra_use = sum(ra_drug_t30, ra_drug_t53)`

`if drug_ra_use > 0 then DMARD=1`

⇒ 류마티스 치료제(DMARD)약물을 처방받은 환자를 구별하기 위하여 T30 T53 테이블의 DMARD 약물의 총사용량을 모두 합하여 DMARD약물 변수를 생성함. 만약 약물 총사용량 변수값이 0보다 크면 이를 DMARD를 처방받은 환자로 볼 수 있다.

3. 류마티스관절염 환자 중 DMARD 사용 환자 선정

- 2015년 RA 환자 중 4-6월 기간 중 처음으로 DMARD 처방이 확인 된 환자

```
data index;  
  set T2040_ra_dmard;  
  if dmard=1;  
  recu_fr_dt = input(recu_fr_dd, yymmdd10.); /*숫자형태의 변수를 날짜형태로 바꿔주는 과정*/  
  if 4 <= month(recu_fr_dt) <= 6;  
  format recu_fr_dt yymmdd10. ;  
run;  
  
proc sort data=index;  
  by jid recu_fr_dt; /*환자 별로 요양개시일자 순서대로 정렬 */  
run;
```

3. 류마티스관절염 환자 중 DMARD 사용 환자 선정

- cohort 정의, index date 정의

```
data cohort; /* => eligible cohort */  
  set index;  
  by jid;  
  if first.jid; /* 개인식별대체키를 기준으로 처음으로 dmard를 처방받은  
                명세서만 선택하고 요양개시일자를 index date로 정의*/  
  rename recu_fr_dt=index_date;  
  keep jid sex_tp_cd pat_age recu_fr_dt ra dmard;  
run;
```

3. 류마티스관절염 환자 중 DMARD 사용 환자 선정

위의 프로그램 결과 생성되는 테이블은 다음과 같다.

	JID	SEX_TP_CD	PAT_AGE	ra	DMARD	index_date
1	1004278	2	75	1	1	2015-06-26
2	1012241	2	57	1	1	2015-04-17
3	101568	2	65	1	1	2015-04-02
4	1026937	1	42	1	1	2015-06-04
5	105498	2	44	1	1	2015-05-11
6	1055840	2	74	1	1	2015-06-26
7	1059694	2	40	1	1	2015-04-02
8	1063514	2	54	1	1	2015-05-07
9	1068210	2	53	1	1	2015-04-20
10	1077186	2	52	1	1	2015-04-15
11	1078690	1	34	1	1	2015-04-03
12	1080463	2	56	1	1	2015-05-19

3. 류마티스관절염 환자 중 DMARD 사용 환자 선정

※ 프로그램 Tip

① `recu_fr_dt = input(recu_fr_dd, yymmdd10.)`

format `recu_fr_dt yymmdd10.`

⇒ STEP1. input 함수로 값을 읽어 들여 날짜포맷으로 정의

STEP2. format을 이용해 출력할 포맷을 지정

FORMAT 형태	출력 형태
MMDDYY8.	07-04-15
DDMMYY10.	04-07-2015
DATE9.	04JUL2015
MONYY5.	JUL2015

- SAS 날짜 함수

사용법	기능/참고사항
<code>mdy(월, 일, 연도)</code>	지정한 날짜를 sas날짜로 변환, 1960/01/01~
<code>today() / date()</code>	오늘 날짜를 sas 날짜로 변환
<code>time()</code>	현재 시간을 sas 시간(초)으로 변환(0-86400), 지정 12시~
<code>day(sas날짜)</code>	입력한 sas 날짜의 '일'반환 (1~31)
<code>qtr(sas날짜)</code>	입력한 sas 날짜의 '분기'반환 (1~4)
<code>weekday(sas날짜)</code>	입력한 sas 날짜의 '요일'반환 (일=1, 월=2, ..., 토=7)
<code>month(sas날짜)</code>	입력한 sas 날짜의 '월'반환 (1~12)
<code>year(sas날짜)</code>	입력한 sas 날짜의 '연도'반환
<code>intnx('interval', from, increment, '<alignment>')</code>	from에서 increment만큼 구간을 이동한 값 반환 interval : year, semiyear, month day, qtr, week, weekday, tenday, dtweek, demonth, semimonth, hour, minute, second alignment : 'B' beginning, 'M' middle, 'S' sameday 기간의 초, 중, 말 설정 가능

3. 류마티스관절염 환자 중 DMARD 사용 환자 선정

② **if 4 <= month(recu_fr_dt) <= 6**
⇒ 4월부터 6월까지 명세서만 선택

4. 결과 발생 정의

- 고관절 골절 발생

/*T2040_ra_dmard테이블에 기준일자를 left join*/

proc sql;

create table t40_1 as

select a.*, b.index_date

from t2040_ra_dmard as a

left join cohort as b

on (a.jid=b.jid); /*개인식별대체키를 기준으로 통합*/

quit;

4. 결과 발생 정의

```
data t40_2;  
  set t40_1;  
  /*기준일자 + 6개월을 followup date로 정의*/  
  followup_date=intnx('month', index_date, 6, 'sameday');  
  if index_date <= input(recu_fr_dd, yymmdd10.) <= followup_date;  
  /*T40테이블에서 상병코드 중 고관절 골절 발생 변수 생성*/  
  if substr(SICK_CD,1,4) in ('S720','S721','S722') then fracture=1;  
  format followup_date yymmdd10.;;  
run;
```

4. 결과 발생 정의

proc sql;

create table t40_3 as

select jid, max(fracture) as fracture /*고관절 골절 발생*/

from t40_2

group by jid;

quit;

4. 결과 발생 정의

```
proc sql;  
  create table cohort1 as  
    select a.*, b.*  
  from cohort as a  
  left join t40_3 as b  
  on (a.jid=b.jid); /*개인식별대체키를 기준으로 통합*/  
quit;
```

4. 결과 발생 정의

위의 프로그램 결과 생성되는 t40_2, cohort1 테이블은 다음과 같다.

	SPEC_ID_SNO	ra	DMARD	index_date	followup_date	fracture	JID	DMD_TP_CD
1	13603405	1	1	2015-06-26	2015-12-26	.	1004278	0
2	13603405	1	1	2015-06-26	2015-12-26	.	1004278	0
3	13603405	1	1	2015-06-26	2015-12-26	.	1004278	0
4	13603405	1	1	2015-06-26	2015-12-26	.	1004278	0
5	15985410	1	.	2015-06-26	2015-12-26	.	1004278	0
6	15985411	1	1	2015-06-26	2015-12-26	.	1004278	0
7	15985412	1	1	2015-06-26	2015-12-26	.	1004278	0
8	15985410	1	.	2015-06-26	2015-12-26	.	1004278	0
9	15985411	1	1	2015-06-26	2015-12-26	.	1004278	0
10	15985410	1	.	2015-06-26	2015-12-26	.	1004278	0
11	15985412	1	1	2015-06-26	2015-12-26	.	1004278	0

	JID	SEX_TP_CD	PAT_AGE	ra	DMARD	index_date	fracture
1	1004278	2	75	1	1	2015-06-26	.
2	1012241	2	57	1	1	2015-04-17	.
3	101568	2	65	1	1	2015-04-02	.
4	1026937	1	42	1	1	2015-06-04	.
5	105498	2	44	1	1	2015-05-11	.
6	1055840	2	74	1	1	2015-06-26	.
7	1059694	2	40	1	1	2015-04-02	.
8	1063514	2	54	1	1	2015-05-07	.
9	1068210	2	53	1	1	2015-04-20	.
10	1077186	2	52	1	1	2015-04-15	.
11	1078690	1	34	1	1	2015-04-03	.

경청해 주셔서 감사합니다.