

## 임상연구를 위한 데이터베이스의 개발

연세대학교 의과대학 외과학교실 및 장기이식연구소

김 명 수 · 김 유 선

### Development of Database for Clinical Transplantation Research

Myoung Soo Kim, M.D. and Yu Seun Kim, M.D.

Department of Surgery and the Research Institute for Transplantation, Yonsei University College of Medicine, Seoul, Korea

The database that is characterized by real-time accessibility and concurrent sharing by multiple users is indispensable to large population-based clinical research. Because initial design determines the effectiveness of database, delicate and complete development is required. Understanding of research and communication between user and programmer is the first step of database development. According to the purpose of research item, the entity and attribute are determined in detail that can minimize data redundancy and maintain data consistency. The programmer should analyse the demand of final user before logical design. Also the user should provide unit information and old data to programmer. After interaction between programmer and user, the database system flow is designed by real organizational unit flow diagram. The user's view is classified by subtitle and grouped by similar data. The flow of view is concordant with real flow diagram. The active participation of final user in early phase is essential for successful development of database. (J Korean Soc Transplant 2005;19:107-118)

**Key Words:** Database, Clinical research  
**중심 단어:** 데이터베이스, 임상연구

1990년대 이후 국내의 이식건수가 급격하게 증가하면서 기관별 누적이식건수가 1,000예를 상회하는 기관이

책임저자: 김유선, 서울시 서대문구 신촌동 134번지  
연세대학교 의과대학 외과학교실, 120-752  
Tel: 02-2228-2115, Fax: 02-313-8289  
E-mail: yukim@yumc.yonsei.ac.kr

본 연구는 2005년도 연세대학교 의과대학 장기이식연구소의 연구비 지원으로 이루어졌음.

점차 증가하는 추세이다.(1) 아울러 1990년대 중반 이후 국내장기이식현황조사 등 전국적인 자료수집과 이에 대한 통계처리가 늘어나면서 이식에 관련된 통계량도 양적으로는 물론 질적으로도 향상이 요구되는 실정이다. 이에 따라서 이식 데이터에 대한 효과적인 데이터 수집과 운영을 위한 데이터베이스가 필요하게 되었다. 더구나 1990년대 이후 새로운 면역억제제를 대상으로 한 다기관 조사나 임상 3상, 4상의 연구가 많이 시행되면서 데이터베이스의 필요성은 더욱 증가하고 있다.

그럼에도 불구하고 임상영역에서의 데이터베이스의 활용은 아직은 광범위하게 적용되지 못하고 있는데, 이는 실제 프로그램의 최종 사용자(user)라 할 수 있는 의료시행자가 데이터베이스의 설계단계부터 적극적으로 참여하여 프로그램을 개발하거나 설계하지 못하기 때문이다. 흔히 우리들은 데이터베이스를 데이터의 수집 혹은 데이터의 추출만으로 그 역할을 이해하고 있으나, 데이터를 이용한 다양한 정보를 추출해 내고, 이에 따라서 유의한 결과 내지는 결정사항을 얻기 위하여서는 데이터베이스의 설계단계부터 이용자의 요구사항과 원하는 정보의 양과 질에 따라서 데이터의 수준을 결정하고 데이터베이스를 설계하여야 한다.

이에 저자들은 다년간 이식관련 데이터베이스를 운영하면서 축적된 경험을 바탕으로, 데이터베이스의 설계단계에서 고려해야 할 사항과 이에 대한 예제를 함께 제시하고자 한다.

### 데이터(data)와 정보(information) 그리고 의사결정(decision-making)

관찰이나 측정 등의 수단을 통하여 수집한 사실이나 값을 데이터라고 하는데 이러한 데이터는 그 내용의 구조는 단순하지만 내용은 구체적이며 자주 바뀌는 동적인 특징을 갖는다. 이러한 데이터는 매우 상세한 실제 값을 의미하므로 어떠한 가공이나 조작이 없이는 데이터 내용의 일반화(generalization)가 어렵다. 반면에 정보라 함은 데이터를 가공한 결과물로서 유용한 해석이나 데이터 상호간의 관계를 제공하는 등 내용의 일반화가 가능하며 기존의 교육이나 경험을 바탕으로 파생된 지식과 더불어 의사결정의 중요한 근거가 된다.(2) 예를 들어 전체 이식환자의 데이터를 수집한 상태에서 재이식의 경우만을 추출하여 조사한 임상연구를 가정해 보

면, 재이식군에서 급성거부반응의 빈도가 증가하였고 따라서 재이식의 경우에는 면역억제제를 다르게 적용하여야 한다는 결론을 내릴 수 있을 것이다. 이 경우에 있어서 재이식 여부와 급성거부반응의 경력은 데이터이며, 재이식 여부에 따른 급성거부반응의 빈도 등과 같은 재이식군에 관련된 자료는 정보이며, 이에 따라서 면역억제제가 달라져야 한다는 것은 이들 데이터와 정보에 의거한 의사결정이라 할 수 있다.

### 데이터베이스의 정의와 특성

일반적으로 특정조직내에서 다수의 사용자들이 공유할 수 있도록 통합시키고 컴퓨터 저장장치에 저장시킨 운영 데이터의 집합을 데이터베이스라 정의한다.(2,3) 데이터베이스는 손으로 기입하는 관리장부나 스프레드시트 혹은 파일 처리시스템과는 달리 다음과 같은 특성 혹은 장점을 가진다고 할 수 있다.

#### 1) 실시간 접근(real-time accessibility)과 공유(concurrent sharing)

데이터베이스에 수집되거나 저장된 데이터는 실시간 접근과 여러 명의 동시사용이 보장된다. 특히 이식분야의 데이터는 이식 전, 이식 직후 그리고 이식 후 등으로 순차적으로 발생하며, 이를 발생시마다 새로이 입력 혹은 갱신해야 하는 특성이 있다. 또한 발생하는 데이터의 내용이 단순입력자료부터 의학적 판단이 필요한 고도의 판단행위까지 다양하므로 데이터베이스의 사용자가 일반사무원, 장기이식 코디네이터, 실험실 연구원 및 의료진까지 다양할 수 밖에 없다. 따라서 데이터의 입력과 추출이 실시간으로 가능하며, 여러 명의 사용자가 동시 사용할 수 있어야 한다.

#### 2) 계속적인 변화

데이터베이스 내의 데이터는 계속적으로 변화하는 동적인 특성을 지닌다. 즉 데이터의 생성 및 소실은 물론 데이터의 갱신(update)이 수시로 일어날 수 있다. 이식관련 데이터 중 이식결과(result or outcome)에 해당하는 데이터는 추적기간이 증가할수록 갱신이 필요하다. 특히 이식장기 생존여부, 환자 생존여부, 이식장기 기능 정도, 합병증의 발생여부 등은 동적인 자료로서 사건(accident)의 발생 순간마다 해당 데이터가 갱신되어야 한다.

#### 3) 데이터의 중복성(redundancy)을 최소화하며 데이터의 일관성(consistency)을 유지한다

기존의 데이터 관리는 각 운영부서마다 별도의 데이터 파일이 필요하다. 예를 들어 장기이식코디네이터는 장기이식대기자 명단이 필요하고, 검사실에서는 검사자의 명단이 필요하고, 의료진은 이식한 환자의 명단이 필요하다. 그러나 실제로 각 부서에서 보관 중인 데이터는 일정부분이 서

로 중복되고 순차적으로 발생하는 특성을 가진다. 데이터베이스는 이러한 중복을 최소화할 수 있다. 즉 장기이식대기자 명단이 생성되면 다른 부서에서는 장기이식대기자 데이터에다가 검사 시행여부나 이식수술여부를 새로운 데이터로 입력하기만 하면 장기이식대기자로 등록된 데이터는 그대로 검사실이나 의료진에게도 제공될 수가 있다. 따라서 데이터의 중복성을 최소화함은 물론 각 부서마다 사용하는 데이터의 일관성을 유지할 수 있다.

#### 4) 데이터의 무결성(integrity)

데이터의 무결성이란 데이터베이스에 정확한 데이터가 유지되고 있음을 보장하는 것으로 수기로 작성하는 장부와는 달리 입력단계에서 데이터의 정확성을 유지시켜 준다. 만일 이식일자를 입력함에 있어서 '2005-11-31'이라는 일자는 존재할 수가 없으므로 입력이 거부되고, 숫자로 설정되어 있는 데이터에서 텍스트(text)가 거부되는 것은 데이터베이스만의 특징이라 할 수 있다. 특히 검사결과를 입력함에 있어서 참조치(reference value)를 벗어난 데이터에 대하여서는 입력을 제어할 수 있는 기능은 올바른 데이터의 수집을 위하여서는 반드시 필요하다.

#### 5) 데이터의 표준화

데이터베이스의 데이터는 공유하는 사용자 간의 데이터의 교환과 처리의 효율성을 위하여 데이터를 표준화하여 사용한다. 각 데이터는 데이터의 내용에 따라서 데이터의 종류, 형식, 크기, 기본값 등을 지정한다. 따라서 다양한 연구과제가 진행되는 경우 표준화된 데이터에 대하여서는 각 연구과제마다 일정한 형식의 데이터가 지속적으로 제공될 수 있는 것이다.

#### 6) 데이터의 독립성(independency)

데이터의 독립성이란 데이터의 내용과 구조가 운영프로그램과는 별개로 독립되어 존재한다는 것을 의미한다. 과거 스프레드시트의 경우 프로그램과 데이터가 같은 파일로 구성되고 있으므로 파일이 손상되는 경우 데이터도 같이 손상된다. 데이터베이스에서는 데이터의 내용과 구조는 운영프로그램과 독립적으로 유지되므로 프로그램의 표현방법이나 운영방법에 따라서 데이터의 내용이 독립적으로 유지될 수 있다. 이러한 독립성은 데이터의 보관(back-up)이나 보안성을 유지하는데 필요하다.

### 데이터베이스의 개발

데이터베이스의 개발은 요구사항 분석단계, 논리적 설계단계, 물리적 설계단계, 데이터베이스 구축단계로 나눌 수 있다. 일반적으로 임상연구를 위한 데이터베이스의 개발에 있어서 실패하는 가장 큰 원인은 사용자의 요구사항을 무

	1988	WBC	Hb.	Hct.	PLT	BUN	Cr.	Sugar
Pre	5840	10.7	31.8	181000	32.1	6.1	84	
1 wk	8900	10.9	30.6	259000	37	2	114	
1 mo	5170	9	27.1	201000	20.6	1.6	90	
3 mo	7800	10.5	30.8		29.6	1.6	104	
6 mo	7500	10.9	31.6		27.3	1.6	103	
9 mo	7700	11.9	36.5		20.4	1.3		
1 yr	6900	12.7	36.7	209000	22.3	1.4	99	
2 yr	6200	12.4	35.6	204000	24.1	1.4	108	
3 yr	6800	12.8	36.9		24.1	1.3		
4 yr	6950	12.7	38.8	253000	20.4	1.2	95	
5 yr								

Fig. 1. Exhibition of laboratory result according to the post-transplant timing (in YUMC kidney database).

시하고 데이터베이스를 개발하기 때문이다. 눈으로 보이는 view 자료와 운영프로그램의 개발을 전문 개발자 단독 혹은 최종 사용자의 의견수렴 없이 이루어지는 경우 데이터베이스를 통한 데이터의 수집이 업무흐름과 불일치하거나 오류가 발생하여 데이터의 수집이 어려질 수 있으며, 어렵게 수집된 데이터도 최종 이용자의 의도와는 다른 데이터가 수집될 수 있다.

1) 요구사항 분석

임상연구를 위한 데이터베이스의 개발에서 첫걸음은 무슨 목적으로 데이터베이스를 구축하고자 하는지 최종 사용자의 의도를 파악하는 것이다. 데이터베이스가 이식환자의 경과관찰이 목적인지 아니면 특정 임상연구를 위한 데이터

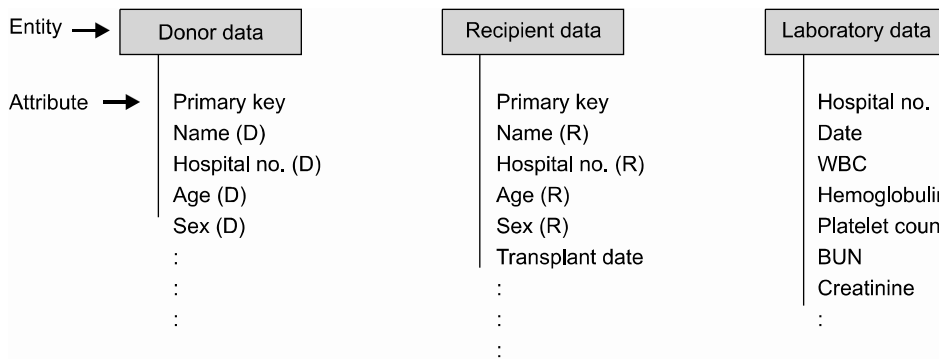


Fig. 2. Design of entity and attribute.

Table 1. Definition of entity and attribute

Original Table				Descriptive Table				
Table name	No.	Field name	Field type (Default) [size]	Description	Class	Table name (code+(Field name))	Code description ((Field name)+mane)	Reference
R_data								
(ID)	1	mTxno	Long	Transplant serial No.	L			
	2	mTxdate	Date	Transplantation date	A			
	1.1	Key	Long (1)	Waiting status	Code_type		1 Waiting-living donor 2 Waiting-Deceased donor 3 Transplant done 4 Cancel 5 Death during waiting time	
	3	Konos	Text [20]	KONOS registration No.	B			
	4	Konodate	Date	KONOS registration date	B			
	5	Hosno	Text [12]	Recipient Hospital Number (Link field for Tx candidate)	L			
	6	IMdr	Text [10]	Hepatologist	A			
	7	Surgeon	Text [10]	Transplant surgeon	A			
	8	Coordi	Text [10]	Coordinator	A			
	9	NameK	Text [20]	Korean Recipient name	A			
	10	NameE	Text [40]	English Recipient name	B			
	11	Sex	Long	Recipient sex	A	Code_sex_r	1 Male recipient	
	12	Age	Long	Recipient age	A		2 Female recipient	

수집인지에 대하여 최종 사용자와 충분한 협의를 하여야 한다. 사용자는 개발자에게 정확한 요구사항을 제공하여야 하며, 이에 관련된 기존의 자료나 업무내용 파일을 제공하여야 한다. 반면에 개발자는 사용자의 현실업무에 대하여 정확하게 파악하고 있어야 한다. 즉 업무흐름도나 각 조직간의 정보 생성과 공유관계를 감안하여 효과적으로 수행할 수 있는 시스템을 개발하여야 한다. 특히 의료분야의 시스템개발에서는 다른 분야와는 달리 새로운 업무시스템을 개발하기보다는 기존의 업무를 유지하는 시스템을 개발하여야 하기 때문에 개발자로서는 어려움이 많게 된다. 실제로 일반혈액검사결과가 기존의 병원업무에서 WBC-Hemoglobin-Hematocrit-Platelet count 순으로 보고되고 있다면 당연히 새로운 데이터베이스 시스템에서도 이에 준하여 입력 혹은 표시되어야 한다. 아울러 일반혈액검사결과가 순차적으로 발생하여 이의 변화추이를 한 화면에 구현되기를 원한다면 이를 설계과정에 반영하여야 한다(Fig. 1). 그리고 업무의 흐름에 따라서 같은 자료라도 사용자의 부서에 따라서 표현형식을 달리하여 데이터의 공유를 적절히 구현하여야 한다.

2) 논리적 설계

요구사항 분석단계에서 파악된 요구사항에 적절한 모델

을 만드는 작업으로 필요한 데이터의 속성과 내용을 정의하고 각 데이터의 공유와 데이터간의 관계를 정의하는 작업이다. 각 데이터베이스의 개체(entity)를 정의하고 각 개체가 가지는 세부적인 항목(attribute)을 설정하여야 한다(Fig. 2). 아울러 이러한 각 항목에 대하여서는 구체적인 데이터의 속성, 정의, 크기 등을 정의하여야 한다(Table 1). 각 개체간의 연관성을 관계라고 하는데 공여자의 자료와 수여자의 자료를 연결하려면 동일한 항목을 이용하여 두 개체간 관계(relation)를 설정하여야 한다(Fig. 3). 이러한 관계를 각각 독립적인 개체를 유기적으로 연결함으로써 사용자의 업무흐름에 적절한 시스템을 만들 수 있다. 예를 들어 수여자의 정보를 보여 주면서 바로 옆에 공여자의 정보를 보여 주는 Fig. 4는 수여자의 자료와 공여자의 자료를 기본 key로 관계를 설정해서 가능해진다.

3) 물리적 설계

설계를 결정된 데이터베이스로 구현될 수 있도록 구체적으로 설계하는 것으로 일반적으로 논리적 설계단계에서의 개체와 속성은 각각 저장 데이터파일로 설정한다. 예를 들어 액세스 프로그램에서는 테이블과 필드로 변환함과 동시에 데이터 형식, 제약조건 및 인덱스 여부 등을 통하여 각 필드의 속성을 규정한다(Fig. 5). 필드의 속성을 얻고자 하

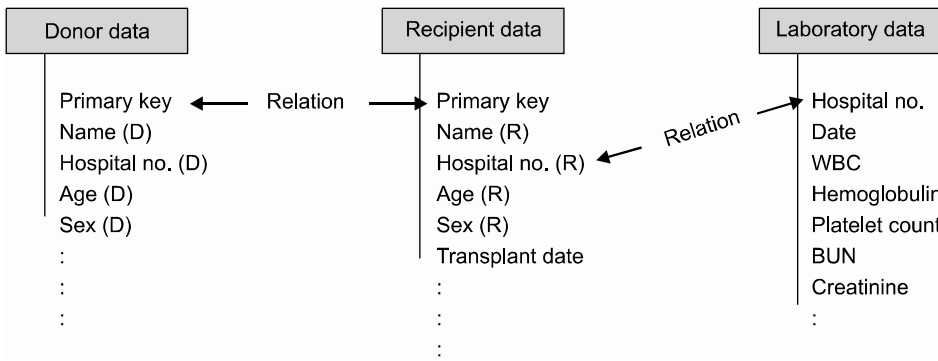


Fig. 3. Relation between entities.

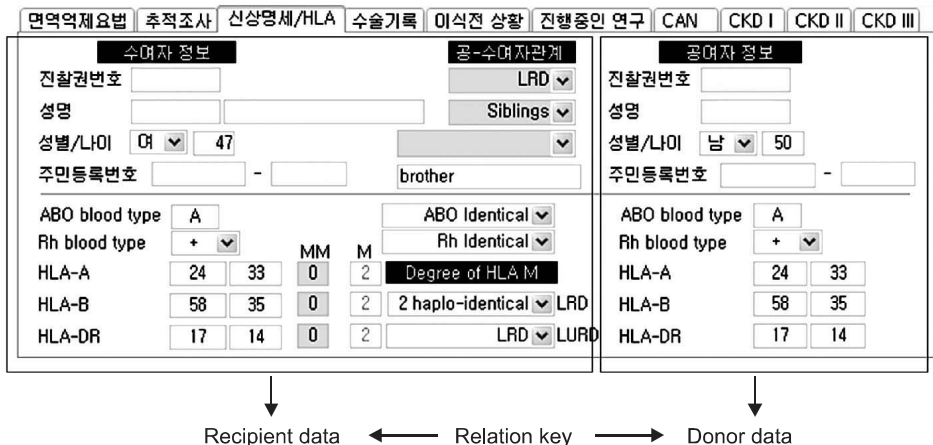


Fig. 4. Relation between the independent tables (in YUMC kidney database).

Main : 테이블			
필드 이름	데이터 형식	설명	
TXNO	숫자		
HOSPITAL	숫자		
SURGEON	숫자	박기일 / 김유선 / 김순일 / 김명수 각각 1, 2, 3, 4	
NEPHRO	숫자	장기식 / 이호영 / 최규현 / 강신욱 각각 1, 2, 3, 4	
KNAME_R	텍스트		
ENAME_R	텍스트		
UNITNO_R	숫자	병록번호	
TXDATE	날짜/시간	이식날짜	
TXYEAR	숫자	year of transplatation	
SEX_R	숫자	1:남자 / 2:여자	
AGE_R	숫자	이식 당시 나이	
BIRTH_R	날짜/시간	생년월일	
IDNO_R	텍스트		
SEX_D	숫자	기증자성별 1:남자 / 2:여자	
AGE_D	숫자	기증당시나이	
REL	텍스트	Donor-recipient relationship	
REL_CODE	숫자	Kind of Donor 0:빈칸 / 1: LRD / 2: LURD / 3:CAD* 기본설정 0:빈칸	
LRD	숫자	0:빈칸 / 1: parent / 2: offspring / 3:sibling* 기본설정 0:빈칸	
LURD	숫자	0:빈칸 / 1: distant relatives / 2: spouse / 3:exchange donor / 4: friend / 5: others*	
필드 속성			
일반 조회			
필드 크기	정수(Long)		
형식	자동		
소수 자릿수			
입력 마스크			
캡션			
기본값			
유효성 검사 규칙		필드에 입력할 수 있는 값을 제한하는 식, 유효성 검 <F1> 키를 누르면	
유효성 검사 텍스트			
필수	아니요		
인덱스	아니요		
스마트 태그			

Fig. 5. Table structure of kidney database at Yonsei University Medical Center.

는 데이터의 내용과 일치시키는 것이 무엇보다 중요하다. 숫자형 필드를 설정한다면 입력 값이 정수인지 혹은 실수인지에 대하여 고려하여야 하며, 허용범위에 대하여 구역을 설정하는 것이 좋다. 예를 들어 신장(height)을 입력함에 있어서 150 cm를 50 cm로 입력하는 오류는 입력단계부터 방지하여 주어야 한다. 단순한 판단형인 데이터(발생여부 혹은 약제의 투여여부)는 Yes와 No의 체크형태로 설정하면 편리하다.

#### 4) 데이터베이스 구축

마지막 단계는 물리적 설계를 바탕으로 데이터베이스를 생성하는 단계로 이 과정에는 데이터베이스 view를 개발하는 것도 포함된다. 사용자 뷰를 개발하는 단계는 사용자의 요구사항에 충실하게 제작이 되어야 하며 업무흐름도에 따라서 순차적으로 입력이 가능하며 소 그룹끼리 묶어서 하나의 타이틀로 묶어서 표현하는 것이 적절하다(Fig. 6). 즉 Fig. 6에서 보면 신이식 데이터베이스에서 화면의 우측에는 검사결과를 이식 후 기간에 따라서 순차적으로 배열하고, 좌측 상단에는 추적조사 및 임상연구에 관한 소그룹이, 그리고 좌측 하단에는 이식 후 발생하는 임상경과를 발생시에만 입력하게끔 배치되어 있다. 아울러 타이틀 내에서의 이해와 업무흐름에 필요하다면 하나의 데이터를 화면의 타이틀마다 반복하여 표시하여야 한다. 예를 들어 이식 전 몸무게는 이식 후 체중변화의 기준 값으로 이식 후 검사결과

에 표시될 수도 있으나, 수여자의 이식 전 체격조건을 나타내는 이식 전 자료항목으로도 표현될 수 있다.

#### 임상연구를 위한 이식 데이터베이스의 설계경험

데이터를 모으기 위하여서는 먼저 수집된 데이터를 가지고 무엇을 할 것인가를 고려하여야 한다. 즉 데이터를 모은 다음에 그 데이터를 가지고 무슨 결과를 낼 것인가를 생각하는 것보다는 무슨 결과가 필요하므로 어떠한 데이터가 필요하다는 생각을 하여야 한다. 추출하고자 하는 정보나 임상연구 결과를 염두에 두고 데이터를 수집하기 위하여서는 수집하고자 하는 데이터의 내용과 형식에 대하여 이미 설명한 대로 충분한 설계를 사전에 시행하여야 한다. 연구의 목적에 합당하게 데이터의 내용이 설계되어야만 데이터 수집 이후에 추출과정이나 통계처리과정에서 추가로 데이터를 수집 또는 수정하여야 하는 수고를 덜 수가 있다.(4)

#### 1) 임상 연구용 데이터베이스는 전자차트와는 다르다

데이터베이스는 이식환자에 대한 모든 데이터를 저장 혹은 관리한다는 개념으로 진행되어서는 안 된다. 흔히 전산화된 의무기록(전자차트, electronic medical record)과 비슷한 수준의 데이터베이스를 요구하는 경우가 많은데, 데이터베이스의 데이터는 연구목적이나 얻고자 하는 결과에 따라서 일정한 형식과 원칙에 따라서 수집되어야 한다.

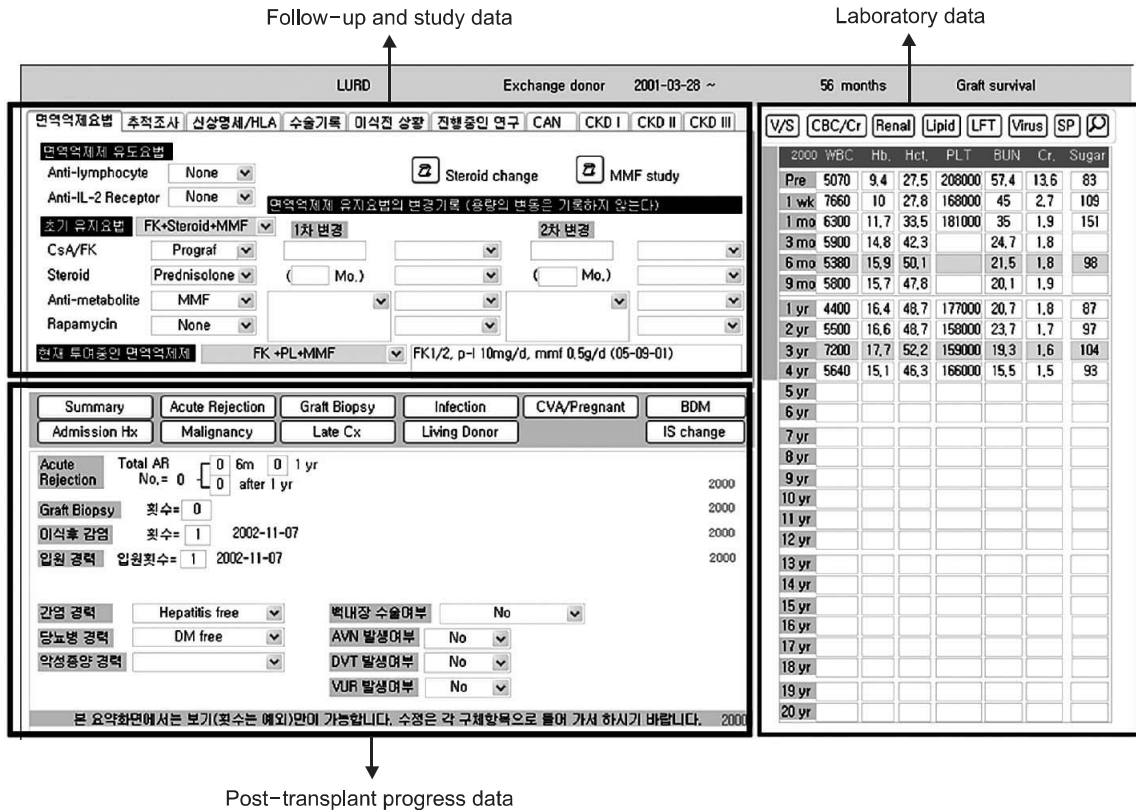


Fig. 6. User view of kidney database (in YUMC database).

2) 유의한 데이터만을 수집하여야 한다

데이터의 무조건적인 수집은 데이터베이스에 아무런 도움이 주지 않는다. 예를 들어 이식 후 경과관찰을 위하여 외래추적기간 중의 모든 혈중 크레아티닌농도를 수집하는 것은 의미가 없다. 각 개별 환자별로 이식신의 상태를 파악하는 데는 모든 시기의 혈중 크레아티닌농도가 수집되는 것이 유리하지만, 반복되는 자료 중 이식신의 기능상태를 대표하는 데는 일부 데이터만이 필요하다. 이런 경우에는 특정시점의 이식신 기능을 대표할 수 있는 대표수치를 데이터로 수집하여야 한다.

3) 데이터는 원래의 형태로 수집되도록 설계되어야 한다

데이터의 속성은 원래의 데이터 속성을 유지하는 것이 좋다. 숫자를 텍스트형으로 변경하거나 날짜를 숫자나 텍스트형으로 바꾸는 일이 없도록 한다. 특히 날짜는 연월이나 이식 후 기간과 같은 데이터를 연산으로 추출할 수 있는 자료이므로 반드시 날짜속성을 유지하도록 설계되어야 한다. 데이터의 속성의 변경은 데이터의 변환이나 다른 데이터와의 연산을 불가능하게 할 수 있다.

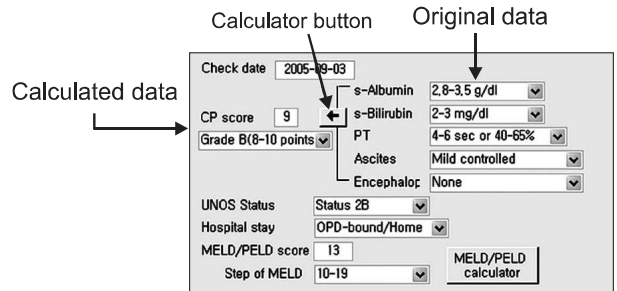


Fig. 7. Data conversion; original data and calculated data (in YUMC liver database).

4) 계산된 데이터는 반드시 원데이터를 같이 가지고 있어야 한다

Body mass index (BMI)와 같이 체중과 신장으로 계산되는 데이터는 계산된 BMI보다는 체중과 신장으로 데이터를 수집하여야 한다. 마찬가지로 Child-Pugh score도 각 항목별 점수를 별도로 가지고 있어야 한다. 계산된 데이터는 원데이터로 전환이 불가능한 반면 원데이터는 계산된 데이터로 언제든지 전환이 가능하다(Fig. 7).

5) 텍스트형의 데이터보다는 표준화된 데이터가 통계 처리에 유리하다

다양한 기술이 가능한 텍스트형태의 자료는 숫자로 표준화하여야 한다. 이는 구축된 데이터베이스로부터 임상연구 정보를 추출하거나 전체 데이터에 대한 통계처리를 위하여

서는 숫자화되어 있는 것이 유리하기 때문이다. 표준화과정에서는 반드시 참고문헌이나 교과서 내에서 널리 쓰이고 분류와 정의를 사용하여야 한다. 예를 들어 공수여자간의 관계를 표시할 때 단순히 쌍둥이, 형제, 자매, 부모자식간, 먼 친척, 순수기증자 등으로 표시하는 것보다 공여자의 종류에 따른 분류, 조직적합항원의 일치정도에 따른 구분, 혈연관계에 따른 구분으로 세분하여 표준화시키는 것이 타당하다.

Fig. 8. Appointment of required field (in YUMC liver database).

6) 간단하고 적절한 연산은 필요하다

단순연산이 필요한 데이터는 입력단계에서 연산이 가능하게 설정하는 것이 필요하다. Child-Pugh score는 각 세부 항목의 합으로 구성 되므로, 연산기능을 사용하면 기존의 데이터만으로도 자동으로 데이터를 입력할 수 있다(Fig. 7). 마찬가지로 BMI, 체표면적, MELD score, 몸무게 당 면적억 제제의 용량, 이식 후 기간, 생존기간 등은 연산으로 자동계 산이 가능하므로 이를 적극 활용하여야 한다.

7) 필수입력항목의 설정

내용이나 속성상 반드시 입력할 필요가 없는 필드도 있지만 기본 key로 사용하거나 다른 테이블과의 상관에 사용되는 key로 사용되는 필드는 반드시 입력하여야 한다. 이러한 필수입력필드는 기본 데이터로서 null값을 허용하여서는 안 되며 새 자료 등록과정에서 누락되는 일이 없도록 조치하여야 한다(Fig. 8).

이식 데이터베이스 설계 예

이식 데이터베이스 설계에는 전통적인 임상연구의 과정은 물론 최신 임상연구의 과정이 반영되어야 한다. 따라서 다양한 임상연구과제를 근거로 한 임상연구를 위한 데이터베이스의 설계가 필요한데 이에 대한 최근 데이터 항목을

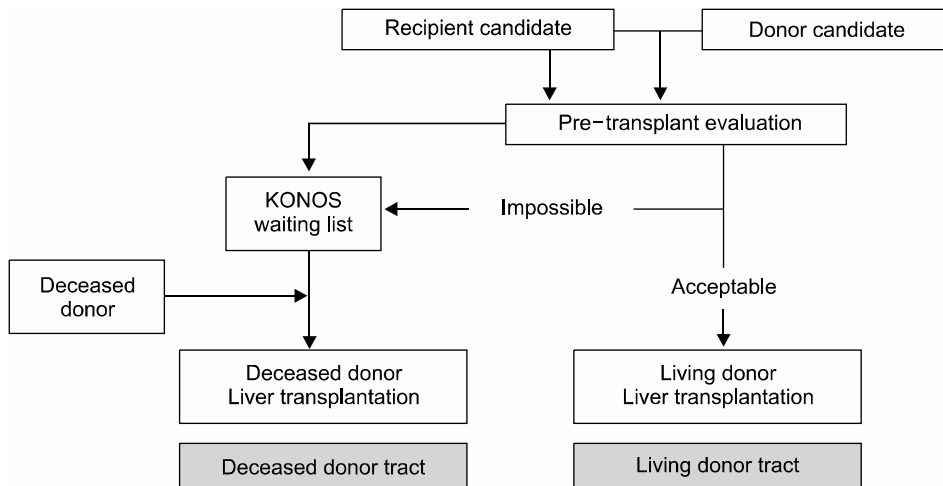


Fig. 9. The flow diagram of liver transplant recipient (in YUMC liver database).

Table 2. Basic data of kidney transplantation database

Title	Subtitle	Data	Calculated data	
Transplantation information	Yearly census Demographics (R/D)	Transplantation date		
		Birth date (R/D)		
	R-D relationship	Gender (R/D)	Gender relationship	
		Donor type by social relationship		
	Degree of HLA matching	Donor type by HLA matching degree		
		HLA-A mismatching No.	Matching or mismatching degree	
		HLA-B mismatching No.		
	ABO blood type matching	HLA-DR mismatching No.		
		ABO blood type (R/D)		
	Retransplantation Immunosuppressive regimen	Rh blood type (R/D)		
		Previous transplantation numbers		
		Main immunosuppressive agent		
		Steroid free or type		
Underlying disease of recipients	Triple regimen and type	Change of regimen		
		Original kidney disease		
	Diabetes mellitus	Hepatitis B virus status (R/D)		
		Hepatitis C virus status (R/D)		
		CMV status (R/D)		
Size mismatch	Weight and height (R/D)	Body mass index (R/D)		
	Graft (kidney) weight	Body surface (R/D)		
Survival analysis	Patient survival rate	Mortality date	Survival months	
		Cause of mortality	Survival status	
		Graft status at mortality date	Censored graft survival rate	
	Graft survival rate	Graft fail date		
		Cause of graft fail		
Follow-up loss	Last follow-up date			
Renal allograft outcome	Delayed graft function (DGF)	Onset of DGF		
		Recovery degree		
	Acute rejection	Diagnosis date	Numbers of rejection within 6 months	
		Diagnosis method		
		Anti-rejection therapy	within 1 year	
		Response rate	after 1 year	
	Chronic allograft nephropathy (CAN)	Diagnosis date	Post-rejection survival months	
		Pathologic findings	Serial change of CAN	
	Graft function monitoring	Anti-CAN management	Post-rejection survival months	
		Serum creatinine level	Measured GFR	Calculated GFR
			Urinary protein excretion	
			Urine analysis	
at discharge				
at post-transplant 3 months				
at post-transplant 1 year				
at post-transplant n year				
Chronic renal disease (CKD) criteria	Anemia	CKD grade		
	Diabetes			
	Dyslipidemia			
	Bone disease			
	Hypertension			
	Cardiovascular disease			
	Onset			
Diagnosis				
Management				
Progress of disease				



Table 2. (Continued)

Title	Subtitle	Data	Calculated data
Deceased donor information	Donor condition	Age	
		Sex	
		Cause of brain death	
		History of cardiac arrest	
		Duration of hypotension	
	Organ harvest	History of Cardiogenic agent	
		Multi-organ harvest	
		Cold ischemic time	
	Cold ischemic time	Allocation distance	
		Organ preservation	Type of preservation solution

R = recipient, D = donor.

Table 3. Basic data of liver transplantation database

Title	Subtitle	Data	Calculated data	
Pre-transplant evaluation	Yearly census Demographics (R/D))	Transplantation date		
		Birth date (R/D)		
		Gender (R/D)		
	R-D relationship Degree of HLA matching	Donor type by social relationship		
		HLA-A mismatching No.	Matching or mismatching degree	
		HLA-B mismatching No.		
		HLA-DR mismatching No.		
	Lymphocyte crossmatching			
	ABO blood type matching	ABO blood type (R/D)		
		Rh blood type (R/D)		
	Transplantation	Type of transplantation		
		Previous transplantation numbers		
	Underlying disease of recipients	Original liver disease	Original liver disease	
			Status of liver cirrhosis	Child-Pugh score MELD score UNOS status
			Hepatitis B virus status (R/D)	
		Hepatitis C virus status (R/D)		
		Small-for-size	Hepatocellular carcinoma (HCC)	
CMV status (R/D)				
EBV status (R/D)				
Weight and height (R/D)			Body mass index (R/D) Body surface area(R/D) Standard liver volume(R/D) Graft weight-to-recipient weight ratio (GRWR) Resectional liver volume ratio (ERLV)	
Graft (liver) weight		Graft (liver) weight		
		Graft (liver) volume		
Operation	Donor safety Operation time	Remained liver volume		Remained liver volume ratio
		Operation time		
		Cold ischemic time		
	Operation technique	Warm ischemic time		
		Hepatic vein anastomosis		
		Portal vein anastomosis		
		Hepatic artery anastomosis		
	Immediate post-transplant graft function	Bile duct reconstruction		
		Bile production etc.		

Table 3. (Continued)

Title	Subtitle	Data	Calculated data
Post-transplant course	Immunosuppressive regimen	Induction therapy	
		Main immunosuppressive agent	
		Steroid free or type	
	Anti-viral prophylaxis	Triple regimen and type	
		Change of regimen	
		Anti-HBV therapy	
	Complications	Anti-HCV therapy	
		Primary non-function	
		Vascular complication	
	Recurred disease	Biliary complication	
		Non-surgical complication	
		Hepatitis B recurrence	
	Acute rejection	Hepatitis C recurrence	
Hepatocellular carcinoma recurrence			
Diagnosis date		Numbers of rejection	
Chronic rejection	Diagnosis method	Post-rejection survival months	
	Anti-rejection therapy		
	Response rate		
	Diagnosis date	Post-rejection survival months	
	Pathologic findings		
Survival analysis	Patient survival rate	Anti-rejection therapy	
		Mortality date	Survival months
		Cause of mortality	Survival status
	Graft survival rate	Graft status at mortality date	
		Graft fail date	
Follow-up loss	Cause of graft fail		
Living Donor information	Pathology findings	Last follow-up date	
		Steatosis	
	Anatomical variation	Hepatic vein	
		Portal vein	
		Hepatic artery	
		Bile duct	
	Post-operation complication	Post-operation liver dysfunction	
		Surgical complication	
		Non-surgical complication	
	Post-hepatectomy regeneration	Post-operation evaluation date	
Degree of regeneration			
Deceased donor information	Donor condition	Age	
		Sex	
		Cause of brain death	
		History of cardiac arrest	
		Duration of hypotension	
		History of Cardiogenic agent	
	Organ harvest	Multi-organ harvest	
		Cold ischemic time	
	Organ preservation	Allocation distance	
		Type of preservation solution	
Pathology findings	Steatosis		

R = recipient, D = donor.

제시하고자 한다.

### 1) 신장이식 데이터베이스

1990년대 이후 신장이식의 성적이 월등히 향상되면서 이식신 및 환자 생존여부, 급성거부반응의 빈도뿐만 아니라 이식신 기능 정도, 이식신의 조직병리소견, 면역학적인 표식자들이 이식신 성적을 나타내는 종료점(end point)으로 이용되고 있는 추세이다.(5) 특히 이식신 기능 정도와 이식신 생존과의 상관관계에 대하여서는 아직은 논란의 여지가 있는 상태로(6,7) 양자간의 엄밀한 분석이 요구되는 실정이다. 이식신 기능에 대한 대표값으로 혈중 크레아티닌 농도보다는 사구체투과율(glomerular filtration rate, GFR)을 선호하며 이식 후 일정시점의 이식신 기능뿐만 아니라 이식 후 기간 별 이식신 기능의 하락정도(slope)도 유의한 결과로 인정되고 있다.(8) 따라서 이식신기능의 하락추세를 자가신장에서 발생하는 만성 신질환(chronic kidney disease)의 진행과정으로 이해하려고 시도되고 있다.(9,10) 이러한 임상연구를 위하여서는 이식신의 생존 여부는 물론 GFR로 대표되는 이식신기능 정도, 이식신 생검의 병리소견, 만성질환의 동반 여부(빈혈, 지질대사장애, 골질환, 당뇨병, 심혈관계질환 등) 등을 추가로 데이터로 수집하여야 한다.

Table 2는 신이식에 관련된 논문이나 책자에서 자주 사용되는 결과에 따른 데이터베이스의 데이터이다.

### 2) 간이식 데이터베이스

2004년도 KONOS 의 장기이식현황(1)을 살펴보면 국내에서 시행된 뇌사자에 의한 간이식은 64예에 불과한 반면 생체 간이식은 580예로 국내 간이식의 대부분을 차지하였다. 간이식 데이터베이스도 이를 반영하여 생체 간이식의 업무흐름도에 따라서 작성함이 당연하다. 생체간이식은 수여자에 적합한 공여자를 발굴하는 작업부터 시작되므로 데이터의 등록은 가능한 공여자 검사시점부터 시작되어야 한다(Fig. 9).

생체 간이식에서의 문제는 ‘수여자에게 적절한 크기의 간장이 제공될 수 있는가?’와 ‘공여자는 잔여간장으로 안전할 수 있는가?’ 등이다. 수여자 체중에 대한 이식간의 무게 비율이나 표준 간용적에 대한 절제간 용적비율은 small-for-size graft의 지표로 사용되며, 전체 간용적에 대한 잔여간 용적비율은 donor safety의 지표로 사용된다.(11,12) 아울러 공여자의 해부학적인 특징도 생체 간이식 공여자 여부를 결정하는 인자 중의 하나이다.(13) 수여자의 이식 전 상태를 나타내는 지표로서는 전통적으로 Child-Pugh score 나 UNOS status를 많이 사용하였으나, 2002년 이후부터는 MELD (model for end-stage liver disease) score(14)를 많이 사용하고 있으므로 이를 반영하여야 한다.

이식간 생존율이나 환자 생존을 뿐만 아니라 B형 간염바이러스나 C형 간염바이러스에 의한 만성간염이나 간세포

암의 재발도 이식성적으로 보고된다. 특히 간세포암의 경우에는 전통적인 간이식의 기준인 Milan criteria를 벗어난 간이식 예와 이식 전 시행한 항암요법이 간이식에 미치는 영향에 대한 연구가 최근 들어 활발하게 진행 중에 있다.(15) 간이식 데이터베이스에는 이러한 연구과제에 대한 항목설정이 필요할 것이다.

Table 3은 간이식에 관련된 논문이나 책자에서 자주 사용되는 결과에 따른 데이터베이스의 데이터다.

## 결 론

이식관련 임상연구를 위한 데이터베이스의 개발은 물리적인 설계자나 프로그램 개발자뿐만 아니라 데이터베이스의 목적과 내용을 충분히 이해하고 있는 최종 사용자가 개발단계부터 적극적으로 참여함으로써 성공적인 데이터베이스를 완성할 수가 있다.

## REFERENCES

- 1) 국립장기이식관리센터. 장기이식통계. <http://www.konos.go.kr/>
- 2) 이재호. 데이터베이스의 기본개념. In: 이재호 저. 데이터베이스 시스템 총론. 제2판. 서울, 도서출판 정일; 2002. p. 20-51.
- 3) 김지연, 정종진. 엑세스와 데이터베이스. In: 김지연, 정종진 저. 데이터베이스 기초와 실습. 제1판. 서울, 한빛미디어(주); 2005. p.27-82.
- 4) 김명수, 전경옥, 권기환, 유희철. 자료구축, 관리 및 통계분석. In: 박기일, 조백환, 김유선 저. 신장이식 이론과 실제. 제1판. 서울, 의학문화사; 2004. p.396-404.
- 5) Hariharan S, McBride MA, Cohen EP. Evolution of endpoints for renal transplant outcome. Am J Transplant 2003;3:933-41.
- 6) Kaplan B, Schold J, Meier-Kriesche HU. Poor predictive value of serum creatinine for renal allograft loss. Am J Transplant 2003;3:1560-5.
- 7) Pascual J, Marcen R, Ortuno J. Renal function: defining long-term success. Nephrol Dial Transplant 2004;19(suppl 6):vi3-7.
- 8) Kasiske BL, Gaston RS, Gourishankar S, et al. Long-term deterioration of kidney allograft function. Am J Transplant 2005;5:1405-14.
- 9) Woo YM, Pereira BJB, Gill JS. Chronic kidney disease progression in native and transplant kidneys. Curr Opin Nephrol Hypertens 2004;13:607-11.
- 10) Djamall A, Kendziorski C, Brazy PC, Becker BN. Disease progression and outcomes in chronic kidney disease and renal transplantation. Kidney Int 2003;64:1800-7.
- 11) Tanaka K, Ogura Y. "Small-for-size" and "Small-for-size syndrome" in living donor liver transplantation. Yonsei Med J 2004;45:1089-94.
- 12) Fan ST. Safety of donors in live donor liver transplantation in adults. Arch Surg 2000;135:336-40.

- 13) Nakamura T, Tanaka K, Kiuchi T, et al. Anatomical variations and surgical strategies in right lobe living donor liver transplantation: lessons from 120 cases. *Transplantation* 2002; 73:1896-903.
- 14) Olthoff KM, Brown Jr RS, Delmonico FL, et al. Evolving concepts in liver allocation in the MELD and PELD era. *Liver Transplant* 2004;10(suppl 2):A6-22.
- 15) Kurtovic J, Riordan SM, Williams R. Liver transplantation for hepatocellular carcinoma. *Best Pract Res Clin Gastroenterol* 2005;19:147-60.
-