예방의학 전공의 과정에 필요한 컴퓨터 지식

순천향의과대학 예방의학교실 안 재 억
(Id; k2ahn@kol.net)

서 론

4년전부터 학회의 요청에 의해 전공의 1년차에게 예방의학 전공의 과정에 필요한 컴퓨터 지식에 대하여 강의를 하였다. 그러나 이 분야는 너무나 빠르게 지식이 변모하여 본 저자가 지도할 자격이 있는지 의심스럽다. 4년전에 썼던 원고는 먼 옛날 이야기 같고 작년에 써 놓았던 원고 역시 진부하게 느껴진다. 작년 7월 이후 1년간 가장 변했던 일은 역시 각 대학이나 병원마다 인터넷 전용선이 설치된 것이다. 컴퓨터를 켜는 순간 LAN선에 접속되어 중앙 컴퓨터에 연결되고 인터넷도 자유 자재로 사용하게된다. 1989년 저자가 예방의학 전공의 1년차 시절에는 XT급 컴퓨터에 팔란티어 워드나 보석글 등으로 각 예방의학 교실에서 작업환경측정 결과나 리포트 보고용으로 워드를 쓰고, dBASE 3+로 자료 입력, SAS 프로그램으로 통계 처리하는 것이 모두였던 PC가 이제는 모든 이에게 있어 매우 편리한 도구로 자리 잡고 있다. 이런 현실에서 이미 많은 분들이 알고 있는 일반적인 프로그램은 제외하고 예방의학이라는 학문과 관련된 것들만 본 저자가 알고 범위에서 소개코자 한다. 또한 새롭게 발전되고 있는 "의료정보학"을 의대생에게 누가 가리킬 것인가도 앞으로 새로운 과제라 생각된다. 이미 미국의 여러 대학에 의료정보학 교실이 설립되어 있다. 그러나 아직 국내 대학에 의료정보학 교실이 독립적으로 존재하는 학교는 없는 걸로 안다. 산업의학 전문의 과정이 생겨 예방의학에서 산업의학의 많은 부분이 제외된 상황에서 예방의학 전공의 과정에 의료정보학을 받아들여야 하지 않을까 생각된다.

국내의 대부분의 병원이 OCS나 EMR 등의 병원 전산화를 하여 이런 곳에 구축되어지는 수많은 의료 데이터베이스를 정보로 분석하는 의사도 미래에는 필요하리라 생각된다. 2,000년대에 전문의가 되는 현 전공의 1년차 선생님들은 미래 상황에 맞게, 예방의학 학문과 같이 컴퓨터 사용을 계속 병행해야 할 것으로 사료된다.

본 론

예방의학 전공의 과정 중 주로 배워야 할 컴퓨터 프로그램은 데이터베이스와 통계프로그램이 주가 될 것이다. 그중 통계학 프로그램은 통계학 공부와 함께 필히 터득하여야 할 과제이다. 다음의 프로그램을 소개하니 기본적으로 공부하시기 바란다.

Ⅰ. 통계 프로그램

1. SPSS for Windows

윈도우 버전의 사용은 DOS 버전과는 비교가 안될 정도로 편리하다. 그 중에서도 SPSS 윈도우 버전이 현재로서는 가장 훌륭해 보인다. SPSS 버전은 6.0까지는 16비트용으로 윈도우 3.1이나 윈도우95에서 사용가능하고, 6.12 버전까지는 윈도우 3.1에서 사용하는 32비트용 프로그램이다. 윈도우 95에서는 6.13 이상 버전이 되어야 정상적으로 동작한다. 6.13 버전 이후로 7.0 버전이 나왔는데 7.0 버전부터는 Output 윈도우에 나오는 화면이 6.1 버전과 많이 틀리다. 현재 7.5 버전까지 있고, 7.5 버전의 약간의 버그 때문에 7.52버전이 판매되고 있다. 프로그램은 미국에서 판매되는 프로그램에는 특별한 제한이 없으나 아시안 버전은 복제 방지를 위하여 키락이나 날짜 프로텍터가 있다. 프로그램은 상당히 고가여서 쉽게 구입할 엄두가 안난다. 그러나 9월초에 실시한 SPSS 사용자 세미나에서 20 Variables와 200 Case 까지 쓸 수 있는 한글 SPSS 7.52 베타 버전 CD를 무료로 배포하였다. 한글 SPSS 7.52 버전은 모든 내용이 한글화되어 있어 통계학 공부하실 때 필히 참고하기 바란다. SPSS 윈도우 버전은 모든 통계 명령어를 화면상에서 보면서 마우스 클릭만으로 해결할 수 있다, 그러나 방대한 자료를 여러 형태로 가공 처리할 경우에는 오히려 프로그램 명령어 방식인 SAS 보다 쓰기가 불편하다. 윈도우즈용 SPSS도 이런 점을 보안하기 위하여 Syntax 명령어 창이 있다. 서울의대 SAS 교재에 있는 예제 파일을 SAV 파일로 변환한 자료가 하이텔 예방의학 통계동호히 자료실에 있으니 공부하시기 바란다. 현재 한글판 참고서적은 "SPSSWIN을 이용한 통계분석"이라는 책이 시중 서점에서 판매되고 있으며, 판매용이 아닌 설명서로는 SPSS 한국대리점 격인 고려정보에서 나온 "초급 및 고급 데이터분석", "시계열분석", "범주형분석" 등 여러 가지 설명서를 판매하고 있다. 96, 97년도 "사용자 사례논문집"에도 좋은 내용이 많으니 참조바란다. 97년 10월경 한글 사용 매뉴얼이 나온다한다. SPSS 개발시 의사가 관여되었다한다. 그래서 그런지 매뉴얼 및 예제 파일에 좋은 의학 예제가 많이 보인다.

프로그램이 7.5 버전으로 Upgrade 되면서 Output 윈도우를 고급화시켰다. 7.5 버전부터는 Output 윈도우를 워드프로세서나 프리젠테이션 프로그램 수준으로 기능을 향상시켰다. 특히 인터넷 기능에 발맞추어 HTML 저장 기능이 있어 Output을 HTML로 저장한 후 다시 HWP 프로그램으로 불러들이면 SPSS Output에서 자유 자재로 구현한 칼라화된 도표 등을 HWP 프로그램에서 바로 받아들인다. 물론 MS-Word를 사용한다면 HTML 저장 없이도 Output을 바로 받아들인다. SPSS 7.5 버전이 설치되지 않은 컴퓨터를 위하여 SPSS View 프로그램을 무료로 배포하고 있다.

2. SAS for Windows

SAS 윈도우 버전은 SPSS 윈도우즈처럼 간단하지 않고, 기존 DOS용 SAS와 명령 체계는 거의 흡사하고, 단지 화면 구성만 윈도우 형식이다. 항상 날짜로 프로그램에 프로텍터를 걸어나서 프로그램의 사용 계약 날짜가 지나가면 컴퓨터 상에서 날짜를 되돌려 사용해야한다. SAS로 만들어진 SSD 자료 파일에도 날짜가 기억되어 날짜가 맞지 않으면 프로그램이 구동되지 않는다. 백업본으로 프로그램을 복사하여 쓸 때에는 config.sas 내용을 잘 확인 후 사용하여야 프로그램이 실행된다. 자신의 컴퓨터에서 SAS 윈도우를 SETUP 시킨 후 성공적으로 SAS가 실행되면, 서울의대 유근영 교수님이 쓴 SAS 교재의 내용을 모두 공부해서 터득하여야 한다. 하이텔의 go comlearn으로 가서 SAS 강좌란으로 가면 많은 추가 지식을 얻을 수 있다. 하이텔의 통계연구 동호회에서도 얻을 많은 자료가 많이 있다. 표 1은 윈도우즈용 SAS와 SPSS 프로그램을 간략히

비교한 것이다. 가끔씩 통계 프로그램을 어떤 것을 공부하는 것이 좋겠느냐는 질문을 받을때가 있다. 질문하는 사람의 수준을 보아서 잘 설명해 주는 예방의학 전문의가 되길바란다.

표 1. 윈도우즈용 SPSS와 SAS 프로그램의 비교

	SPSS for Windows	SAS for Windows
추천 대상	① 통계 초보자. ② 비전공자 중 통계분석이 필요한 실무자 (의료인, 직장인 등).	① 시간이 많고 현재 통계 분석이 필요하지 않은 자. (대학생, 대학원 생) ② 통계전공자 ③ 비전공자 중 대단위 자료를 분석할 자.
구입문의	SPSS KOREA (고려정보사) 552-9415	SAS KOREA
통계처리방식	마우스로 변수 클릭	프로그램 작성
자료확장자명	SAV	SSD
외부자료호환	매우 간편히 변환시킴.	다른 파일 변환이 쉽지 않음. DBF->PROC DBF 등 사용
편리한 점	완성한 자료를 분석하는데는 그 어떤 통계프로그램보다 편리함.	자료를 여러 형태로 변화시켜 사용할 때 마우스 방식 보다 편함.
불편함 점	자료를 자주 변형시켜야 할 때 일일이 변수 클릭하여야 하므로 불편함 (보안 : 프로그램 방식을 사용할 수 있음)	간단한 자료에도 모두 프로그램을 사용하여야 하여 불편함.
판매 및 복제방지 방법	판매 또는 대여 방식 키락 사용(아시아 버전만)	1년 단위로 대여 방식 컴퓨터 시간으로
윈도우용 버전정보	SPSS 6.0 (Win 3.1용: 16비트용) SPSS 6.1 6.13(95용) SPSS 7.0 & 7.5 한글 SPSS 7.5	SAS 6.11, 6.12

표 2는 이전에 본인이 게이오 의과대학에 있을 시 받았던 혈중 연 자료로 만들었던 SAS의 프로그래밍이다. SAS 프로그램은 통계 분석보다는 자료의 가공 처리하여야 할 경우 매우 요긴하다. 자료 분석을 위해 처음부터 이렇게 모두 프로그래밍 하기엔 보기는 쉬워도 수많은 에러 메시지를 통과해야 SAS 프로그램을 자유 자재로 쓸 수 있다.

표 2. SAS 예제 (혈중 연 폭로에 따른 혈중 지질치의 차이)

```
OPTIONS  LS = 80   PS=60  ;
LIBNAME  where  "c:\keio"  ;      자료가 있는 위치를 알려준다
DATA new  ;  SET where.keio ;     new라는 데이터는 c:\keio에 있는 keio.ssd
 ARRAY X[54]    -> DBF에서 결번인것이 0로 SAS로 넘어와 모든 필드의 변수를 .으로
     age   height      weight     bmi
     ob4   ob3   ob2      sp4   sp3   sp2       dp4 dp3 dp2
     go4   go3   go2      gp4   gp3   gp2       gt4 gt3 gt2
     hb4   hb3   hb2      wb4   wb3   wb2
     tc4   tc3   tc2      hdl4 hdl3 hdl2      tg4 tg3   tg2
     pbb49 pbb43 pbb39 pbb33 pbb29 pbb23
```

```
    ala49 ala43 ala39 ala33 ala29
    pbb4 pbb3 pbb2    ala4 ala3 ala29    ;
DO I=1 TO 54;  IF X[I] = 0 THEN X[I] = .  ;  END ; DROP  I ;

IF          age >= 50  THEN Nage = 50 ;   연령을 10대군으로 categorical
IF   40<=  age < 50 THEN Nage = 40 ;
IF   30<=  age < 40 THEN Nage = 30 ;
IF   30>  age  THEN Nage = 20 ;
pbb4 = (pbb49 + pbb43)/2    ;     새로운 변수를 계산 산출
ala4 = (ala49 + ala43)/2    ;

IF   pbb23 = . & pbb29 =. & pbb33=. & pbb39=.  &     결손 자료의 처리
     pbb43=.   & pbb49=.  THEN npbb = 1  ;
ELSE IF  pbb23 < 30 & pbb29<30 & pbb33<30 & pbb39<30
       & pbb43<30  THEN npbb = 2  ;
ELSE IF  pbb23 >= 30 &  pbb29>=30 & pbb33>=30 & pbb39>=30
      &  pbb43 >= 30      THEN npbb = 4  ;
ELSE    npbb =  3  ;

PROC SORT  ; BY   npbb   ;    RUN ;
PROC PRINT  ;
  VAR     pbb49 pbb43 pbb39 pbb33 pbb29 pbb23   ; BY npbb ;
PROC  FREQ  ; TABLE Nage*npbb ;  RUN ;
PROC MEANS   MAXDEC=1   ; 평균 계산 및 소수점 1자리 까지만
VAR  age  height     weight  alc   cig  bmi
     ob4  ob3  ob2      sp4  sp3  sp2      dp4 dp3 dp2
     go4  go3  go2      gp4  gp3  gp2      gt4 gt3 gt2
     hb4  hb3  hb2      wb4  wb3  wb2
     tc4  tc3  tc2     hdl4 hdl3 hdl2     tg4 tg3  tg2
     pbb49 pbb43 pbb39 pbb33 pbb29 pbb23
     ala49 ala43 ala39 ala33 ala29
     pbb4 pbb3 pbb2    ala4 ala3 ala29    ;
 BY  npbb   ;  RUN ;
```

```
PROC GLM   ;
  CLASS   npbb  ;
  MODEL   tc4 tc3  hdl4 hdl3   tg4 tg3 = npbb  bmi age  ;
  RUN  ;                          BMI와 AGE를 adjust하여 각군 비교
```

3. GRIM 또는 EGRET

아직 사용하시는 분은 많지 않으나, 로지스틱 분석에 뛰어나다. 서울의대 예방의학교실 유근영 교수 저 "의학보건학을 위한 범주형 자료 분석론" (서울대학교 출판부)에 소개된 프로그램이다. 프로그램이 DOS용 버전이고 한글 매뉴얼 등이 없으니, 혼자 공부하시기는 어려움이 많다. 대학원 등에서 LOGISTIC 공부하실 때 사용 해 보길 바란다.

4. EPIINFO

EpiInfo 프로그램은 역학에 관련된 프로그램으로 DOS용이지만 다른 통계프로그램이 해결하지 못하는 기능들이 있어 사용해 보아야한다. Raw data가 없는 자료의 통계분석에 용이하다. 프로그램과 간단한 설명서는 예방의학 통계 동호회에 있다.

II. 데이터베이스 프로그램

아무리 통계 프로그램이 편하고 훌륭하다 하여도 분석할 자신의 연구 자료가 없다면 통계 공부하여 타인의 자료만 분석하게된다. 연구 자료는 실험으로 얻어지는 자료와 역학적 자료로 나누어진다. 실험으로 얻어지는 자료는 타인의 자료를 얻기 힘들지만 역학적 자료는 전산화에 의해 구축된 자료들이 많다. 의료보험연합회에서 대형 컴퓨터에 담겨진 자료, 병원에서 입원 환자 관리를 위해 저장된 자료, 타 검진기관 등에서 건강관리 업무 등으로 입력한 자료 등이다. 서울대 예방의학교실에서는 의료보험 연합회 자료에 기타 변수를 추가하여, 여러 문헌을 발표한 적이 있다. 이렇게 구축된 자료들을 잘 이용하여 연구 활동에 사용하려면 데이터베이스 계열 프로그램의 기본 지식을 갖추어야 한다. MS-Office에 들어있는 엑세스나 MS사의 폭스프로 등의 서적을 구입하여 공부하시기 바란다. 데이터베이스에 대한 기본적 언어만을 공부하는 것도 1-2개월 이상 책을 보아야한다. 한글오피스에 들어있는 어프로치 프로그램도 상당히 훌륭하다. 간단히 데이터입력용 프로그램을 만들기에 아주 편리하다. 엑세스의 경우는 MDB로 자료가 저장되어 SPSS에서 직접 불러들이지 못하므로 본 저자는 주로 어프로치를 사용하는 편이다. 미래에는 가급적 대형 컴퓨터에 저장되는 DB 언어를 잘 파악해야 할 것으로 보인다. 각 대학이나 병원의 전산 환경을 이해하고 서로 접하는 기회를 갖기 바란다. 작년까지만 해도 의사로서 폭스프로 프로그램에 깊은 지식이 있는 노준양 선생님이 자세한 DB 언어에 대해 강의했으나 1년차 선생님에게 맞는 내용인 것 같지 않다. 1년전 강의록의 대부분의 내용이 하이텔 예방의학 동호회 자료실에 보관되어 있으니 참고하시기 바란다.

III. HTML (hypertext markup language)

세계는 급속히 변해가고 있다. 특히 LAN선을 통한 인터넷의 사용이 앞으로 학문을 배우는 방식에 큰 변화를 일으킬 것 같다. 조만간 현재의 LAN 선보다도 100~200배 빠른 케이블 TV선을 이용한 인터넷 사용이 2,000년대에 각 가정마다 보급된다하니, 학문의 접근방법도 새로워져야 할 것으로 보인다. 일반인들도 대부분의 의학 정보를 가정에서 손쉽게 검색하여 볼 수 있기 때문에 예방의학 학문을 하는 전문의에 입장에서 당연히 의료 정보를 구축하는 방법에 대해 기본적으로 알고 있어야 할 것으로 보인다. WWW에 홈페이지를 구축하는 것은 별로 어렵지 않은 작업이다. 시중에 나와 있는 서적 한 권만 구입하여 1주일 정도만 보면 누구나 다 쉽게 구축할 수가 있다. WWW의 언어가 HTML로 지속되는 한 HTML 언어를 공부하고 홈페이지 구축하는 방법을 공부하시기 바란다.

WEB 브라우져로 정보를 검색하는 일은 이제 당연한 생활이 되었고 넷스케이프나 인터넷 익스프로러 4.0 버전을 이용한 화상 통신도 보편화되고 있다. 현대 사회에 들어 PC는 모든 학문과 같이 연관하여 공부할 도구로 생각된다. 학문적 공부도 중요하겠지만 그 학문을 워드로 입력하여 HTML언어로 저장하고 저장된 내용이 인터넷 홈페이지에 구축되는 의료 정보화 과정이 매우 빠르게 발전되고 있다. 이미 몇 개 학회가 자신의 문헌은 CD로 제작하고 웹에서 검색할 수 있도록 하고있다. 예방의학 교과서 내용이나 예방의학 관련 학회지 검색 등을 인터넷 웹에 언제 누가 구축할지는 모르겠지만 조만간 타 학회지가 모두 이런 서비스가 될 때에는 없어서 창피한 학회가 될 가능성도 높다. 예방의학 전공의 과정에 입문한 선생님들의 진취적인 사고 방식에 의해 훌륭

한 미래가 있길 바란다. 조금 더 의료정보에 관심을 갖고 예방의학 학문을 정보화하여 일반인에게 질병을 예방하는 의료 지식을 전달하는 우리 학회가 되길 바란다.