

공분산구조분석을 이용한 자체충족률 모형 검증

장혜정^{1)†}, 김창엽²⁾

한국보건산업진흥원 산업정보단¹⁾, 서울대학교 의과대학 의료관리학교실²⁾

<Abstract>

Formulating Regional Relevance Index through Covariance Structure Modeling

Hyejung Chang¹⁾, Chang Yup Kim²⁾

Korea Health Industry Development Institute¹⁾, Seoul National University College of Medicine²⁾

Hypotheses in health services research are becoming increasingly more complex and specific. As a result, health services research studies often include multiple independent, intervening, and dependent variables in a single hypothesis. Nevertheless, the statistical models adopted by health services researchers have failed to keep pace with the increasing complexity and specificity of hypotheses and research designs. This article introduces a statistical model well suited for complex and specific hypotheses tests in health services research studies. The covariance structure modeling(CSM) methodology is especially applied to regional relevance indices(RIs) to assess the impact of health resources and healthcare utilization. Data on secondary statistics and health insurance claims were collected by each catchment area. The model for RI was justified by direct and indirect effects of three latent variables measured by seven observed variables, using ten structural equations. The resulting structural model revealed significant direct effects of the structure of health resources but indirect effects of the quantity on RIs, and explained 82% of correlation matrix of measurement variables. Two variables, the number of beds and the portion of specialists among medical doctors, became to have significant effects on RIs by being analyzed using the CSM methodology, while they were insignificant in the regression model. Recommendations for the CSM methodology on health service research data are provided.

Key Words : *relevance index(RI), covariance structure model, health resources, utilization*

† 교신저자 : 장혜정, 한국보건산업진흥원(02-2194-7489, hjchang@khidi.or.kr)

I. 서 론

1. 연구의 배경 및 필요성

1) 의료자원 수급 모니터링 지표

의료자원의 수요과 공급현황을 모니터링함은 보건의료정책을 수립·집행함에 있어 필수적으로 수반되어야 할 기반작업이며, 이를 통하여 국가단위에서 뿐만 아니라 지역단위에서도 수요와 공급이 균형을 이루고 있는 지를 평가하고 불균형이 발생하였을 때 적절히 개입할 수 있는 시스템을 갖추도록 해야 한다. 그러나 양적, 질적으로 변화하는 수요에 맞게 자원을 공급하기 위해서는 실제로 수요와 공급을 구체적으로 이해하여야 하는데, 그 관계가 매우 복잡하다. 예컨대, 수요추정에 기반이 되는 정보인 의료이용에 영향을 미치는 요인들이 매우 다양하고 상호관계가 복잡하며(배상수, 1985; 권순호 등, 1987; 유승흠 등, 1987), 의료이용과 관련한 대부분의 연구 또한 환자 개인의 행태와 그에 영향을 미치는 요인들(예컨대 성·연령, 교육수준, 소득수준, 배우자 유무 등)을 도출하는 데 그치고 있다(Grossman, 1972; Andersen과 Newman, 1973; Sindelar, 1982; Feldstein, 1993). 따라서, 지역단위의 과학적 근거 부재로, 제한된 자원으로써 수요에 부응하는 자원을 공급하기는 매우 어려운 실정이다. 일반적으로, 보건의료체계를 모니터링할 때 고려되는 주요지표는 적합성(relevance), 과정(progress), 효율성(efficiency, 효과성(effectiveness)), 그리고 영향(impact)의 5개 영역이다(WHO, 1981). 여기에서 본 연구에서 논의될 자체충족률은 자원수급의 적합성(relevance) 지표로서 제반 의료자원의 복합적 집합체로 공급된 의료서비스의 역량이 대상 주민의 의료수요에 얼마나 적합한가에 관한 과제이다.

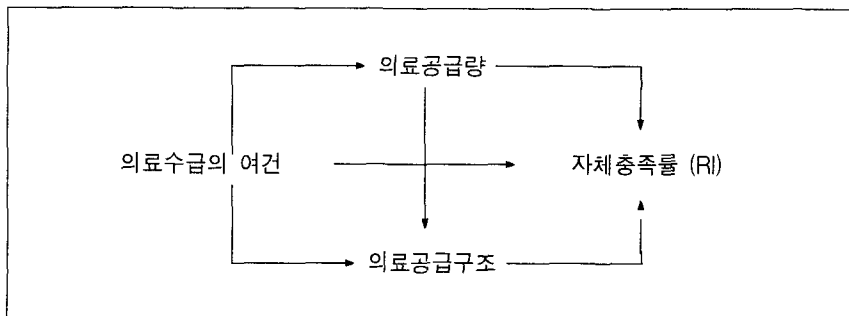
자체충족률은 지역단위의 의료서비스 유출양상을 파악하는 지표로서 제시될 수 있으므로 지역주민중심 친화도(Relevance Index; RI)의 특수한 형태이다. 즉, 지역주민중심 친화도는 한 지역의 주민들이 특정 지역에서 의료서비스를 제공받은 양을 그 지역주민들이 이용한 의료서비스의 총량에 비교한 백분율로써 제시될 수 있으므로, 자체 지역에 대한 친화도를 특히 자체충족률이라 한다. 따라서 자체충족률이 높다는 것은 지역의 의료수요가 지역 내에서 대부분 충족되고 타지역으로의 유출이 상대적으로 적다는 것을 의미한다. j 지역주민중심 자체충족률 RI_j 는

$$RI_j = \frac{j \text{ 지역소재 의료기관을 이용한 } j \text{ 지역 환자의 의료이용량}(U_{jj})}{j \text{ 지역 거주 환자의 총 의료이용량}(U_j)}$$

와 같이 표현된다.

2) 자체충족률 모형

지역단위 보건의료계획은 지역주민들의 의료수요를 자체 지역 내에서 충족하는 방향으로 수립되고 추진된다. 따라서 지역별 의료취약지역을 선정한다든지, 지역보건의료의 우선순위를 설정한다든지, 의료인력 공급량이나 공급구조를 계획하는 경우 등 다양한 측면에서 지역단위 자체충족률 모형은 필요하게 된다(Simcoe York, 2000; Askhms, 2001). 본 연구에서는 지역의 자체충족률에 영향을 미치는 요인들간의 관계를 적절히 설명하는 모형을 검증하고자 시도하였는데, 본 연구의 기본 틀이 된 모형은 [그림 1]과 같다(한달선과 권순호, 1990). 이 모형에 의하면 자체충족률은 의료자원의 공급량 및 공급구조에 의하여 직접적인 영향을 받을 것이며, 지역내 의료기관의 수 및 병상수 등 의료자원의 공급량에 따라서 상호 경쟁관계, 전문화 양상 등의 공급구조가 달라질 가능성이 있으므로 공급량은 자체충족률에 직접적으로 영향을 주기도 하지만 공급구조를 통해 간접적인 영향도 주게 될 것이다. 또한, 의료수요 및 공급에 영향을 미치는 지역사회의 제반 여건, 예컨대 인구구조, 사회경제수준, 지리적 특성, 지역적 특수성 등은 의료이용의 자체충족률에 직접적 영향을 줄뿐만 아니라, 의료공급량이나 공급구조의 결정요인도 되므로 이를 통해 간접적으로 자체충족률에 미치는 효과가 있을 것이다.



[그림 1] 의료이용의 자체충족률 모형의 기본 틀

3) 기존 모형검증 방법론의 한계

[그림 1]의 모형이 나타내는 바와 같이, 의료이용의 자체충족률을 설명하고자 하는 가설들은 상호관계가 복잡하게 연결되어 있어서, 설명하고자 하는 변수(즉, 독립변수)와 설명되는 변수(즉, 종속변수)의 역할이 분명하지 않다. 예컨대 '의료공급량' 개념은 자체충족률에는 독립변수의 역할을 하는 요인인 동시에 의료수급의 제반 여건 변수들에 대해서는 종속변수의

역할을 하고 있다. 한편, 지역여건, 공급량 및 공급구조 등 이론적 개념변수(latent variable)를 완벽하게 측정할 수 없다. 다만, 여러 개의 측정 가능한 변수들을 관찰하여 공통된 부분을 개념으로서 조작적으로 정의(operational definition)할 뿐이다. 공통부분으로 정의되는 개념변수와, 이를 간접적으로 측정하는 데 사용된 관찰변수들(observational variables)의 상호작용에 의해서 모형이 성립하므로, 그 복잡성은 심각하다고 할 수 있다.

현재까지 이러한 모형의 분석에는 흔히 여러 단계의 다중회귀분석(multiple regression analysis) 방법을 이용하여 왔다. 그러나 회귀분석모형은 종속변수와 독립변수가 지정되어 하나의 회귀식으로 제시함으로써 종속변수와 독립변수의 관계는 회귀계수에 의해 직접효과밖에 표시되지 않는다. 따라서 두 개 이상의 독립변수가 상관되어 있을 경우, 간접효과 및 제삼변수효과(spurious effect) 등은 해석하지 않는 결과가 된다. 단순히 예측을 위한 연구라면 회귀분석을 이용하겠지만, [그림 1]의 모형과 같이 내용적 유추 내지는 문제의 양상에 대한 설명이 필요한 경우에는 종래의 회귀분석은 큰 도움이 되지 못한다.

이와 같은 한계를 극복하는 분석방법으로 등장한 것이 경로분석(path analysis)이다(Wright, 1934). 변수들간의 관계에서 가지는 모든 효과를 각각의 경로로 표시하고, 가설을 세워 타당성을 검증하여 직접 및 간접적 효과는 물론 제삼변수효과 및 덮인 효과(unanalyzed effect)까지도 분석하는 방법이다. 그러나 이 방법은 회귀분석에서 이용되지 않는 정보를 발굴하여 경로도형으로 표시함으로써 문제의 설명에 도움을 주는 장점이 있지만, 경로도형내 관계가 일방적 방향이어야 한다든지, 이론변수 측정에 오직 하나의 측정변수만 사용된다든지, 변수들이 적어도 구간척도(interval scale) 이상의 수준으로 측정되어야 한다는 등 비현실적인 가정들(assumptions) 때문에 활용상에 제약이 많다(Pedhazur, 1982).

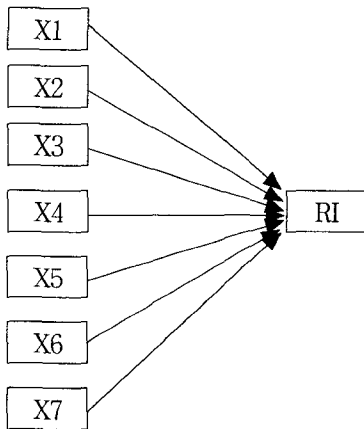
4) 공분산구조분석의 활용

경로분석의 제약점인 가정들을 완화하여 폭넓게 활용되는 분석방법이 공분산구조모형(Covariance Structure Modeling; CSM)¹⁾이다(Joreskog, 1973). 공분산구조분석은 측정모형(measurement model)과 이론구조모형(structural model)의 두 부분으로 구성된다. 측정모형은 개념변수와 관찰변수들간의 연결 관계, 이론구조모형은 이론변수들 간의 관계를 보여주는 모형 부분이다. 따라서 공분산구조분석은 여러 개의 측정모형과 여러 개의 이론구조 모델들의 복합적인 관계를 규명할 때 유용한 분석도구라고 하겠다. 측정모형과 이론구조모형이 합쳐서 공분

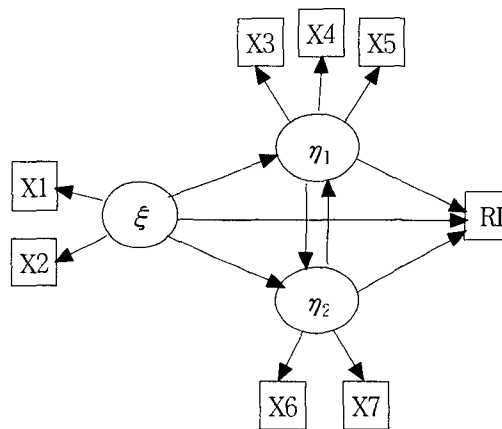
1) 공분산구조분석은 Structural Equation Modelling, LISREL, Path Analysis, Analysis of Latent Variables, Simultaneous Equation Modelling, Causal Modelling 등 다양한 이름으로 알려져 있음. 다른 회귀분석법 등에서는 관찰된 개개의 값 단위로 관찰치와 예측치 간 차이의 함수를 최소화하는 한편, 이 방법론에서는 관찰된 공분산행렬(Covariance Matrix) 또는 상관계수행렬(Correlation Matrix)과 가설화한 모형에 의하여 예측된 행렬과의 차이를 최소화하는 접근방식을 취하므로 공분산의 구조를 분석한다는 의미에서 명명된 것임.

산의 모형을 이루고 있어 이 모형은 다양한 측정치들을 가진 이론변수들간의 관계를 검증하는데 효과적이다. 이러한 장점 덕분에 기존의 회귀분석이나 경로분석이 가지고 있는 문제점을 해결하는 대안으로서, 사회과학분야에서 뿐만 아니라 의과학 분야에서도 활용도가 높아지고 있다. 보건의료분야 또한 산재한 정보를 체계화하여 보건의료체계의 다양한 난제를 공분산구조분석 방법을 통하여 모형화함으로써 보건의료정책 연구에 광범위하게 활용할 수 있을 것으로 기대된다.

본 연구의 [그림 1]의 개념 틀을 7개 관찰변수에 의하여 측정하는 경우, 기존의 다중회귀분석법에 의하면 [그림 2]의 '모형 1'을 검증하는 반면, 공분산구조분석에 의하면 [그림 3]의 '모형 2'를 검증하게 된다.²⁾ '모형 2'에 등장한 ξ , η_1 , η_2 는 지역여건, 의료공급량, 의료공급구조를 각각 나타내는 개념변수로서, [그림 1]의 기본 틀 이론을 측정할 수 있게끔 한다. 개념변수는 두 개 이상의 관찰변수에 의해 간접적으로 측정되고, 그 공통부분이 개념변수를 나타내는 순수한(error-free) 부분으로 이해된다. 따라서, 개념변수인 ξ , η_1 , η_2 간의 연결관계에 의해 이론구조가 가설화되고, 각 개념변수들을 측정하는 관찰변수들의 연결관계에 의해 측정구조가 가설화된다. 그리하여, 다수의 복합가설을 포함한 모형을 동시에 검증함으로써, 종래의 회귀분석이나 경로분석 방법으로 도출할 수 없었던 문제들을 추가적으로 설명할 수 있게 된다.



[그림 2] 검증 모형 1



[그림 3] 검증 모형 2

2) 도식화한 각 모형에는 오차변수(error terms) 및 오차변수간의 공분산(covariance)은 생략하고 관찰변수와 개념변수간의 관계만 단순화하여 제시하였음.

2. 연구 목적

본 연구는 지역 의료이용의 적합성을 평가하는 지표인 자체충족률(RI)을 과학적으로 설명하는 모형을 제시함으로써, 궁극적으로 지역단위의 보건의료자원 수급 적정화의 정보기반 조성에 기여하고자 한다. 이러한 목적을 달성하기 위한 세부 목표는 다음과 같다.

첫째, 자체충족률(RI)을 설명하는 개념모형을 평가할 수 있도록 유의한 측정요인을 도출한다.

둘째, 유의한 측정요인과 자체충족률과의 관계를 개념변수와 관찰변수간의 직접적 및 간접적 관계로써 가설모형을 설정하고, 이를 공분산구조분석방법에 의해 검증하여 양호한 모형을 찾는다.

셋째, 도출된 자체충족률 모형을 활용하여 지역의 의료이용 현황자료로써 추정치를 고찰한다.

II. 연구방법

1. 연구 자료

지역별 자체충족률을 모형화하기 위하여, 우선 지역의 경제사회적 환경여건, 보건의료자원 공급량 및 공급조직, 보건의료서비스 이용현황을 파악하기 위한 관련 자료를 수집하였다. 중진료권 단위로 자료를 수집하였는데, 138개 중진료권 중 병원급 이상 의료기관을 보유하지 못한 2개 중진료권을 제외하고 136개 중진료권을 대상으로 하였다. 지역별 일반 특성, 보건 의료관련 환경 및 의료자원의 공급현황 자료는 시·군·구 통계연감, 사망원인통계연보, 지방재정연감, 대한병원협회 및 의사협회 통계자료 등을 활용하여 확보하였다. 보건의료서비스 이용 자체충족률은 지역주민이 입원의료서비스를 이용한 총 재원일수 중 자체 지역 의료기관을 이용한 재원일수 비율을 산출하였는데, 지역별 의료이용량 정보는 건강보험공단의 급여 실적 자료를 활용하여 1993년 한 해 동안의 입원의료이용에 근거한 통계를 생성하였다.³⁾

3) 지역별 입원의료이용의 자체충족률은 지역단위의 의료이용 자료로써 산출할 수 있는데, 전국적인 정보를 포괄하는 대표적인 것이 보건복지부의 『환자조사』 자료와 건강보험공단의 『의료보험 급여실적』 자료임. 『환자조사』는 2~3년 단위의 조사로서, 1개월간 퇴원한 환자의 의료이용행태를 파악하므로, 일정한 기간 동안의 의료이용량을 정확히 파악할 수 없을 뿐만 아니라, 소규모 인구 지역의 의료이용 정보는 표본크기로 인해 그 신뢰도가 매우 낮음. 한편, 의료보험 자료는 의료보험을 청구하여 급여된 의료이용만 포괄하므로, 자동차보험, 산업재해보험, 자비자불 등의 환자를 포함하지 못하는 제한점이 있으나, 총이용량의 60%이상을 차지하고 있어서 양적 규모면에서 지역단위 의료이용 유출입 행태를 파악할 수 있으므로 본 연구에서 활용함. 단, 의료서비스를 이용한 환자의 거주지를 파악하여야 하므로, 자격파일과 매핑시켜야하는데 그 규모의 방대함으로 대용량 전산장비와 고비용이 요구되어 매핑된 자료가 1993년 자료밖에 없으므로 이 자료를 활용함.

기초행정구역 단위별로 수집된 자료를 의료전달체계의 기본 단위인 중진료권별로 통합하여 분석의 기반자료를 생성하였다. 각 개념변수별로 수집가능한 다수의 관찰변수들을 선정하여 통계자료를 확보하였으며, 통계항목은 <표 1>과 같다.

<표 1> 자체총족률 모형의 개념변수 측정을 위한 관찰변수

개념	지역 환경여건	의료자원 공급량	의료자원 공급조직
인구		총 보건의료인력수(인구1000당)	
수 노인비율(%)		의사수(인구1000당)	
집 인구밀도		보조인력수(인구1000당)	종합병원 비율
통 조출생률(인구1000당)		병상수(인구1000당)	전문의 비율
계 보건소 등록 결핵환자수(인구1000당)		총 의료기관수(인구1000당)	1차의료기관 비율
정 1인당 지역 예산액		의원수(인구1000당)	민간부분 비율
보 지역구분(6대도시, 시지역, 군지역)		약국수(인구1000당)	
		장비보유대수(인구1000당)	

2. 분석 방법

수집 자료의 상관관계 및 보건의료서비스 이용에 미치는 영향요인을 분석하기 위하여 상관분석(correlation analysis) 및 회귀분석(regression analysis) 모형을 활용하였고, 기존 문헌 및 기초 분석결과에 기초하여 영향요소간의 상관성 및 의료이용에의 영향을 설명하는 가설모형을 정립하였다. 가설모형은 관찰된 변수들 및 잠재적인 개념간의 상호 영향력을 나타내는 총체적 관계를 검증하기 위한 여러 개의 선형구조방정식으로 구성되는 공분산구조분석(covariance structure analysis)을 활용하였다. 통계분석을 위한 전산프로그램으로 SAS (Statistical Analysis System) 통계패키지의 PROC GLM과 PROC CALIS 절차를 이용하였다(SAS, 1996).

자체총족률을 설명하는 간명함(parsimony)과 부합도(goodness of fit) 및 이론적 고려가 잘 조합된 적정모형(plausible model)을 찾는 과정(MacCallum, 1986)은 실제 사용에 있어서 간명하면서도 현실을 적절히 설명할 수 있도록 내용적 해석이 허용하는 범위에서 단순화시키는 것이다. 구체적으로 영향계수가 통계적으로 유의하지 않거나 이론상 의미가 없는 경우, 그 영향력을 제거함으로써 모형을 단순화하며, 이 과정에서 통계적 유의도는 유의수준 0.05에 의하여 결정하였다. 부합도의 검증을 위하여 Chi-square test와 Goodness of fit index를 사용하였고, 영향계수들은 회귀분석에서와 같이 t-test에 의해 검증하였다.

검증된 최종 모형에 의하여 각 관찰변수 및 잠재변수가 자체충족률에 미치는 직접적 및 간접적 영향력을 분석하였으며, 자체충족률의 관찰치와 모형에 의해서 추정된 예측치를 비교하여 잔차분석(residual analysis)을 수행하며 지역의 특성을 검토하였다. 추가적으로, 향후 변화할 지역사회의 환경, 보건의료자원의 공급 변화 등을 가정하여 지역별 자체충족률의 변화를 최종모형에 의하여 예측하였다.

Ⅲ. 연구 결과 및 논의

1. 관찰변수의 선정

지역환경여건, 자원공급량, 공급구조의 각 개념변수별로 관찰변수들과의 상관성을 분석한 결과, 각 개념을 측정하는 관찰변수들은 매우 높은 설명력을 보였으나, 많은 변수들이 다른 개념과의 상관성으로 인하여 개념변수를 설명하는 데 있어서 '무의미한 변수'⁴⁾로 나타났다. 또한, 이들 관찰변수들 간의 다단계 회귀분석에 의하면, 앞서 제시한 [그림 1]의 모형이 검증하기에 적절하다고 판단되었다.

따라서 본 연구의 목적에 맞게 가설 모형을 검증하기 위해서 이들 관찰변수들 중 개념변수에 대하여 높은 설명력을 갖는 관찰변수들을 취하고, 상관성이 높은 잉여 변수들을 제외시켰다. 일반적으로 하나의 개념변수에 적어도 세 개 이상의 측정변수를 사용할 것을 권장한다(Bentler와 Chou, 1987). 한 개의 개념변수를 한 개의 관찰변수만으로 측정하는 경우 이론적 개념을 측정하기에 부족한 경우가 대부분이며, 또한 계산과정에서 미지수 추정에 문제가 발생할 가능성이 매우 높다. 한편, 측정변수가 너무 많은 경우에도 모형 부합에 있어 어려움이 많다. 따라서 적정 수의 관찰변수를 갖는 간명하면서도 부합도가 높은 공분산구조분석모형을 도출함이 매우 중요하다. 본 연구에서 관찰변수를 선정함에 있어서는 '간명하면서도 부합도 검증에서 통과되는 모형은 이론개발에 매우 유용한 모형이다'는 입장에서 최대한 단순한 측정모형을 선정하고자 하였다. 수집된 <표 1>의 관찰변수들로서 개념변수들을 설명하기 위해서는 <표 2>가 가장 간명하면서도 설명력이 높게 나타났다. 향후 모형의 활용도에 따라 다른 측정변수들을 삽입하여 관찰변수가 수정된 모형을 검증하는 작업도 가능할 것이다.

결과적으로 각 개념을 설명하기 위하여 사용된 관찰변수는 <표 2>에 제시되어 있다. 인구학적, 사회경제적 변수등으로 고려되었던 지역의 환경여건은 도시, 시, 군 등 지역구분으로 개괄적으로 측정하였고, 의료자원의 공급량은 병상수, 의료인력수, 장비보유대수로 시설, 인력, 및 장비의 대표지표로서 측정하였으며, 공급구조는 종합병원 병상과 전문의 비율로 측정

4) 모형내의 다른 관찰변수가 이미 제시한 설명력을 추가적으로 유의하게 높이지 못하는 변수를 칭함.

하였다. 자체충족률에의 개념변수의 영향을 나타내는 [그림 1]의 개념적 모형을 <표 2>의 관찰변수들을 사용하면 [그림 3]과 같이 표현된다.

<표 2> 지역별 자체충족률 설명을 위한 관찰변수

개념 변수	지역 환경여건	의료자원 공급량	의료자원 공급구조
	지역구분(6대도시, 시지역, 군지역)	X ₃ : 병상수(인구1000당)	X ₆ : 종합병원 병상비율
관찰 변수	X ₁ : 시지역=1, 기타=0 X ₂ : 군지역=1, 기타=0	X ₄ : 의료인력수(인구1000당) X ₅ : 고보유율 장비보유대수 (인구1000당)	X ₇ : 전문의 비율

2. '모형 1'의 분석

1) 다중회귀분석법에 의한 '모형 1' 분석

[그림 2]의 '모형 1'을 상례의 다중회귀분석법에 의하여 분석한 결과는 <표 3>과 같다. <표 2>에 제시한 7개의 독립변수로써 자체충족률의 변이를 78% 설명할 수 있었으며 (R-square=0.7799), 지역내 병상수(X₃)와 전문의 비율(X₇)은 다른 변수들에 비해 자체충족률의 설명력을 유의하게 증가시키지 못하는 변수로 판명되었다. 이는 독립변수간의 상관성이 있음에도 불구하고, 회귀모형의 종속변수에 대한 직접효과만 설명하는 특성으로 인하여 간접효과나 제삼변수효과가 무시되었기 때문이다.

<표 3> 다중회귀분석법에 의한 추정결과

독립변수	추정치	표준오차	t	P-value	표준화 추정치
상수	38.37	8.68	4.42	0.0001	0.0000
X ₁	-24.73	5.20	-4.76	0.0001	-0.5019
X ₂	-39.81	5.50	-7.24	0.0001	-0.8173
X ₃	-0.34	0.45	-0.75	0.4541	-0.0567
X ₄	4.80	0.87	5.51	0.0001	0.5683
X ₅	-4.93	1.27	-3.88	0.0002	-0.2853
X ₆	0.38	0.06	6.29	0.0001	0.3327
X ₇	-0.37	8.60	-0.04	0.9655	-0.0021

2) 공분산구조분석방법에 의한 '모형 1' 분석

한편, [그림 2]의 '모형 1'을 공분산구조분석을 위한 프로그램에 의하여도 분석할 수 있다. 공분산구조분석은 독립변수 및 종속변수의 지역별 개개 관찰치를 분석하는 회귀분석과는 달리, 변수간의 공분산 또는 상관계수 행렬의 구조를 분석하지만, 궁극적으로 동일한 표본을 분석하고 있으므로 공분산구조분석에서 가장 흔히 사용되는 최소자승법(least square method)으로 모형을 검증하는 경우 동일한 결과를 산출하게 된다. 즉, 공분산구조분석은 7개의 독립변수와 종속변수인 자체충족률 간의 영향계수를 이용하여 하나의 방정식으로 제시하는 결과를 제시하는데, <표 3>의 표준화 추정치 결과와 동일하고, R-square 값인 0.7799도 일치하고 있다.

$$RI = -0.5019 * X_1 - 0.8173 * X_2 - 0.0567 * X_3 + 0.5683 * X_4 - 0.2853 * X_5 + 0.3327 * X_6 - 0.0021 * X_7 + 0.4691 E$$

3. 공분산구조분석에 의한 '모형 2' 분석

1) 최종모형찾기

[그림 3]의 모형을 각 변수들간의 영향력 유무 및 오차변수들간의 상관성을 가정하면서 다단계의 모형검증을 통하여 최종 "모델찾기(specification search)"를 시도하였다. 다시 말해서, 각 관찰변수 및 개념변수들 간의 영향력([그림 3]에서 화살표로 표시됨), 가정된 측정오차들 및 그들 간의 상관성을 추정해야할 모수(parameter)로 설정하여 각 모수들이 0인지 아닌지의 가설을 전체적인 구조방정식 내에서 검증하였다. 다단계의 검증작업을 거쳐 최종적으로 산출된 10개 구조방정식을 모수의 추정치를 사용하여 연립방정식으로 표현하면 <표 4>와 같고, 유의한 오차변수(error terms)들의 분산공분산 행렬(variance-covariance matrix) Φ_e 및 Φ_{de} <표 5>에 제시하였다.

<표 4>와 <표 5>은 [그림 4]로 동시에 도식화된다. 3개의 개념변수인 지역의 일반적 여건, 의료자원 공급량 및 공급구조 ξ , η_1 , η_2 는 각각 둘 또는 세 개의 유의한 관찰변수들에 의하여 측정되고, 개념변수들은 직접 및 간접적으로 자체충족률을 설명하고 있다.

[그림 4]의 최종모형은 원자료의 상관계수 행렬을 82.2% 설명하고 있었으며, 관찰치에 대한 구조방정식 모형의 적합도를 측정하는 지표들 역시 매우 높은 결과를 나타냈다. 즉, 카이자승치와 그의 확률치($X^2 = 1.8905$, $df = 3$, $prob > X^2 = 0.5954$)에 의하여 모형이 잘 맞는 것으로 나타났다. 특히, 표본크기의 변화나 다변량 정규분포의 위반에 별 영향을 받지 않는 기초부합치(Goodness of Fit Index; GFI)와 조정부합치(Adjusted Goodness of Fit Index;

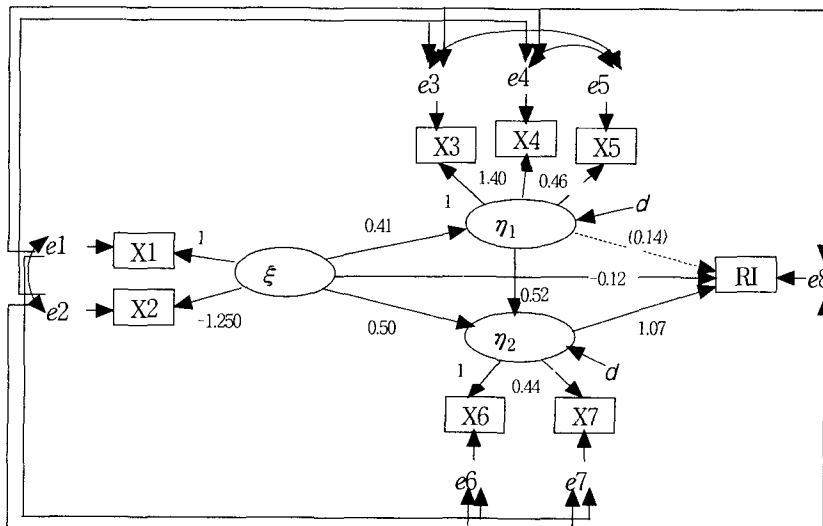
<표 4> 자체충족률 최종모형의 관찰변수 및 개념변수간의 연립방정식

관찰변수(Observational variables) 방정식	개념변수(Latent variables) 방정식
$X_1 = 1.00 \xi + e_1$	
$X_2 = -1.25 \xi + e_2$	
$X_3 = 1.00 \eta_1 + e_3$	$\eta_1 = 0.41 \xi + d_1$
$X_4 = 1.40 \eta_1 + e_4$	
$X_5 = 0.46 \eta_1 + e_5$	$\eta_2 = 0.50 \xi + 0.52 \eta_1 + d_2$
$X_6 = 1.00 \eta_2 + e_6$	
$X_7 = 0.44 \eta_2 + e_7$	
$RI = -0.12 \xi + 0.14 \eta_1 + 1.07 \eta_2 + e_8$	

<표 5> 최종모형 오차변수의 분산공분산 행렬(variance-covariance matrix) Φ_e 및 Φ_d

Φ_e	e1	e2	e3	e4	e5	e6	e7	e8
e1	1.00							
e2	-.92	1.00						
e3	.16	-.18	.41					
e4	.20	-.26	.00	.13				
e5	.00	.00	.38	.42	.88			
e6	.47	-.54	.00	.00	.00	.37		
e7	.44	-.44	.15	.00	.00	.00	.88	
e8	.51	-.68	.00	.00	.00	.00	.00	.18

Φ_d	d1	d2
d1	.57	
d2	.00	.47



[그림 4] 자체충족률 모형의 최종 분석 결과

AGFI 지표도 0.9966과 0.9589로 각각 나타나 Silvia(1988)의 기준에 의하여 양호한 모형으로 판단할 수 있다.

2) 직접 및 간접적 효과 해석

공분산구조분석의 영향계수의 추정치의 크기를 서로 비교할 때는 다음의 두 가지 사항에 주의하여야 한다(Cohen과 Cohen, 1983). 첫째는 비실험 자료들을 기초로 계산된 추정치가 가지는 통계적 정밀함에 대한 의문이다. 통계적 정밀함을 자료가 통계적 가정을 만족시킬 때 비롯되는 것인데, 일반적으로 통계적 가정을 위반하는 경우가 많다. 단순히, 추정치가 0인지 아닌지 또는 방향이 +인지 -인지 정보의 유의도 검증은 자료가 방법론에서의 가정을 조금 위반해도 비교적 강건한(robust) 결정으로 이끌어 주지만, 추정치의 값 자체를 해석하고자 할 때에는 가정의 위반이 가져오는 결과가 훨씬 심각하게 잘못 유도할 수 있는 것이다. 둘째는 측정의 단위 문제이다. 많은 경우 관찰자료들의 측정단위가 임의적(arbitrary)일 경우가 많다. 그러므로 측정치의 크기를 해석하고자 할 때 많은 조심이 필요한 것이다. 따라서 한 모형에서 상이한 변수들의 효과의 크기를 비교하고자 할 경우, 공분산자료에서 측정변수의 분산이 모두 비슷한 크기이거나 측정변수가 모두 표준화된 경우가 바람직하다. 따라서, 본 연구에서는 분석결과로서 산출되는 영향력의 유의성 유무에 초점을 두었고, 추가적으로 총 효과뿐만 아니라 직접 및 간접적 효과의 크기를 비교해 보았는데, 이는 상관계수행렬을 분석 자료로 사용하였기에 가능하다.

잠재변수 의료자원의 공급구조(η_2)는 자체충족률에 직접적인 영향만을 미치는 반면, 의료자원 공급량(η_1)은 직접적으로는 통계적으로 유의성이 낮은 미미한 영향만을 미치고 오히려 공급구조를 통하여 간접적인 영향을 유의하게 미친다. 그 영향력은 간접적인 영향력의 크기가 0.55이고 직접적인 영향력의 크기가 0.14인 결과를 비교하면 쉽게 알 수 있다. 더욱이 잠재변수인 지역의 일반적 여건(ξ)은 직접적 영향보다는 훨씬 더 큰 간접적 영향력을 갖는다. 그 경로도 3개나 되어 공급량(η_1)만을 통과하는 경로, 공급구조(η_2)만을 통과하는 경로, 그리고 공급량(η_1)을 통한 후 공급구조(η_2)를 통과하는 경로에 의하여 영향력을 미치며 직접적으로 자체충족률에 미치는 영향력은 -0.12에 불과하다. 따라서 모형의 각각의 영향계수에 의한 직접적인 효과만을 볼 것이 아니라 간접적인 효과까지 포함한 전체효과⁵⁾를 보는 것이 매우 중요하다. 각 개념변수들이 영향을 미치는 관찰변수 및 다른 개념변수에 미치는 총효과 및 간접적 효과는 <표 6>에 요약하였다. 지역의 자체충족률에는 의료자원의 공급구조가 가장 큰 영향을 미치고 있고(1.07), 지역 일반여건과 의료자원 공급량은 상대적으로 적은 영향

5) 총효과 = 직접효과 + 간접효과로 정의됨.

을 미친다(0.69~0.70). 더욱이, 지역의 의료자원 공급량은 자체충족률에 0.69의 총 효과를 미치는데, 이중 직접효과가 0.14, 공급구조를 통한 간접효과가 0.55이고, 지역의 일반여건도 총 효과 0.70 중 의료자원의 공급구조를 통한 간접효과가 0.54로 대부분을 차지하여, 의료자원의 공급구조가 자체충족률에 미치는 심대한 효과를 입증하고 있었다.

<표 6> 자체충족률 모형의 모든 변수들에 대한 개념변수의 총 효과 및 간접적 효과

종속변수	독립변수의 종속변수에의 총 효과			독립변수의 종속변수에의 간접적 효과		
	ξ	η_1	η_2	ξ	η_1	η_2
X ₁	1.0000	0.0000	0.0000	0.00000	0.00000	0
X ₂	-1.2508	0.0000	0.0000	0.00000	0.00000	0
X ₃	0.4141	1.0000	0.0000	0.41409	0.00000	0
X ₄	0.5792	1.3987	0.0000	0.57918	0.00000	0
X ₅	0.1888	0.4559	0.0000	0.18880	0.00000	0
X ₆	0.7139	0.5159	1.0000	0.71387	0.51585	0
X ₇	0.3158	0.2282	0.4423	0.31577	0.22818	0
η_1	0.4141	0.0000	0.0000	0.00000	0.00000	0
η_2	0.7139	0.5159	0.0000	0.21361	0.00000	0
RI	0.7015	0.6906	1.0729	0.82271	0.55346	0

<표 7> 자체충족률에 대한 개념변수들의 직접 및 간접적 효과 요약

	지역 일반여건		의료자원 공급량		의료자원 공급구조	
직접적 효과	$\xi \rightarrow \text{RI}$	-0.1212	$\eta_1 \rightarrow \text{RI}$	0.1372	$\eta_2 \rightarrow \text{RI}$	1.0729
간접적 효과	$\xi \rightarrow \eta_1 \rightarrow \text{RI}$	0.0578				
	$\xi \rightarrow \eta_2 \rightarrow \text{RI}$	0.5367	$\eta_1 \rightarrow \eta_2 \rightarrow \text{RI}$	0.5534	(해당없음)	
	$\xi \rightarrow \eta_1 \rightarrow \eta_2 \rightarrow \text{RI}$	0.2292				
	(계 : 0.8228)					
총 효과	0.7015	0.6906	1.0729			

4. 자체충족률 모형에 의한 지역별 의료자원수급 현황 분석

136개 중진료권별로 관찰된 자체충족률 현황을 [그림 4] 및 <표 4>에 제시된 지역별 자체

충족률 추정모형에 의하여 추정된 자체충족률과 비교하여 잔차분석(residual analysis)을 시도하였다. 지역에 따라서, 모형에 의하여 관찰치에 매우 근접하게 추정된 지역이 있는 반면, 관찰치와 큰 차이를 갖게 추정된 지역들이 있기 마련이다. 그러나 관찰치와 추정치의 차(difference)인 잔차(residuals)는 평균이 0이고, 정규분포(W: Normal 0.985986; Pr<W=0.7909)를 하고 있었다.

추정모형에 의한 자체충족률 추정치가 관찰된 현황보다 매우 낮게 예측된 지역들로는 상주시(46.43), 광주광역시(33.12), 전주시(31.53), 부산광역시(30.79), 그리고 대구광역시(27.73) 순의 지역들이었고, 반대로 관찰치보다 높게 추정된 지역들로는 예산군(-65.21), 문경시(-55.76), 영덕군(-44.62), 장흥군(-40.98), 영천시(-36.87) 등이었다.

잔차의 절대값이 큰 지역들은 흔히 유사한 다른 지역에 비해서 공급자원의 양이나 구조가 다르기 때문에 나타난다. 자체충족률이 저추계(underestimate)된 지역들은 대부분 인구규모가 큰 대도시 지역들로서 의료자원을 충분히 확보하여 지역내 의료수요를 대부분 충족할 뿐만 아니라 근교지역의 의료수요도 일부 충족시키는 거점형 지역들이다. 이들 지역들은 수적 열세로 인하여 모형설정과정에서 지역의 현황이 제대로 반영하지 못했을 가능성도 있다. 이는 3개 지역유형별 평균을 적용하여 추계하여 본 결과인 <표 8>에 의해서도 나타난다. 중소도시 및 군지역의 잔차가 각각 -2.04와 -3.61인 반면, 대도시지역의 잔차는 24.81에 이른다. 따라서 최종모형은 대도시 지역의 자체충족률을 저추계하는 경향이 있다. 반면, 고추계(overestimate)된 지역들은 현재의 의료자원현황에 비추어 자체충족률이 매우 낮음을 시사한다. 자원의 공급량이나 공급구조에 의하면 좀 더 높은 자체충족률을 나타내어야 하나 자원의 시설수준, 기술 또는 경영능력이 취약하여 지역내 의료수요를 충분히 해결하지 못하고 있거나, 지역내 환자보다는 외부 지역으로부터 유입한 환자가 많은 경우 등으로 해석할 수 있다. 이들 극단적인 추정치를 갖는 지역들은 세부현황을 면밀히 분석하여 논의하여야 할 것이며, 일반적으로 열외변수(outliers)로서 간주하게 된다.

<표 8> 자체충족률 추정모형에 의한 추정결과 지역유형별 분석

지역유형	관찰변수							잠재변수			RI 추정치	RI 관찰치	잔차
	x1	x2	x3	x4	x5	x6	x7	지역 여건	공급 량	공급 구조			
6대도시	0	0	6.69	10.21	3.38	39.71	0.64	0	22.51	39.99	58.45	83.26	24.81
시지역	1	0	6.84	8.73	3.32	33.66	0.71	1	20.56	33.97	51.35	49.31	-2.04
군지역	0	1	5.17	7.18	3.23	11.26	0.58	-1.25	16.69	11.52	23.00	19.39	-3.61

136개 중진료권의 잔차들은 약 ± 10 의 범위내에 중위 50%의 지역들이 속하여 있었고, 특히 <표 9>에 제시되어 있는 최종양 7개지역은 본 연구의 최종 자체충족률 모형에 의하여 매우 적합하게 추계되었다.

<표 9> 자체충족률 추정모형에 의하여 설명력이 높은 7개 지역 분석

중진료 권명	관찰변수							잠재변수			RI 추정치	RI 관찰치	잔차
	x1	x2	x3	x4	x5	x6	x7	지역 여건	공급 량	공급 구조			
포천군	0	1	5.98	8.10	2.78	19.51	0.75	-1.25	18.58	19.84	33.24	32.62	-0.62
영광군	0	1	6.86	9.28	3.41	48.80	0.70	-1.25	21.39	48.11	65.51	65.17	-0.34
성주군	0	1	0.67	4.35	1.46	1.00	0.44	-1.25	7.42	1.19	5.53	5.23	-0.29
담양군	0	1	2.38	5.16	2.47	1.00	0.39	-1.25	10.72	1.17	7.79	8.41	0.63
경주시	1	0	5.02	8.75	2.97	35.83	0.71	1	18.61	36.14	52.33	53.34	1.00
익산시	1	0	6.67	10.08	3.33	52.34	0.51	1	22.29	52.57	72.49	73.50	1.01
김제시	1	0	2.62	5.00	1.98	1.00	0.83	1	10.52	1.37	9.43	10.61	1.19

이들 지역들의 의료자원 공급량 및 공급구조를 나타내는 변수들이 최소한 각 지역이 속한 지역유형의 평균에 이른다고 가정하는 경우, <표 10>과 같이 자체충족률의 변화를 예측할 수 있다. 영광군의 경우에만 모든 관찰변수들이 지역유형별 평균치를 초과하여 예측된 자체충족률에 변화가 없었고, 다른 6개 지역은 한 개 이상의 변수가 지역유형별 평균값으로 상향 조정되었다. 예컨대, 병상수 보유량의 경우 성주군과 담양군은 군지역 평균인 인구1천당 5.17로 추가 공급되는 경우를 가정해 본 것이고, 경주시, 익산시, 김제시는 시지역 평균인 6.84를 가정하여 자체충족률을 예측해 본 것이다. 이러한 가정에 입각하여 추계한 결과에 의하면 포천군, 경주시, 익산시는 공급량과 구조에 있어서의 변화가 미미하여 자체충족률에 거의 변화가 없으나, 성주군, 담양군, 김제시에는 큰 변화가 예측되었다. 5~8%로 낮은 자체충족률을 보이던 성주군과 담양군은 다른 공급량 및 공급구조 여건이 유사지역의 수준에 이른다면 자체충족률이 23%까지 제고될 수 있으며, 김제시 역시 같은 가정으로 10%의 자체충족률이 약 50%까지 상승될 가능성이 있다. 따라서, 현재의 의료자원현황은 취약하지만 공급량을 늘린 다든지, 공급구조를 좀 더 전문화한다든지 하여 유사지역과 비슷한 수준의 자원기반을 갖춘다면 모형에 의하여 추정된 자체충족률만큼 높아질 가능성이 있음을 시사하고 있다.

<표 10> 지역여건의 변화를 가정한 자체충족률 추정

증진료 권명	관찰변수							잠재변수			RI 추정치	RI 관찰치	잔차
	x1	x2	x3	x4	x5	x6	x7	지역 여건	공급 량	공급 구조			
포천군	0	1	5.98	8.10	3.23*	19.51	0.75	-1.25	18.58	19.84	33.38	32.62	-0.76
성주군	0	1	5.17*	7.18*	3.23*	11.26*	0.58*	-1.25	16.69	11.52	23.00	5.23	-17.77
담양군	0	1	5.17*	7.18*	3.23*	11.26*	0.58*	-1.25	16.69	11.52	23.00	8.41	-14.59
경주시	1	0	6.84*	8.75	3.32*	35.83	0.71	1	20.59	36.14	53.70	53.34	-0.36
익산시	1	0	6.84*	10.08	3.33	52.34	0.71*	1	22.46	52.65	72.70	73.50	0.80
김제시	1	0	6.84*	8.73*	3.32*	33.66*	0.83	1	20.56	34.03	51.41	10.61	-40.80

* 지역유형별 평균값으로 대체함.

5. 분석방법론 활용의 논의

이제까지 본 논문은 지역의 의료이용 자체충족률을 종래의 분석방법이 아닌 공분산구조분석 방법론을 활용함으로써 개념변수 및 측정변수간 상호관계를 직접 및 간접적 효과로써 분석할 수 있었고, 다중회귀분석 결과보다 설명력이 높은 모형을 제시할 수 있었다. 그러나 공분산구조 분석의 활용을 제안함에 있어 마지막으로 ‘우연적 이론개발(capitalization on chance)’(Cliff, 1983)의 논리를 지적하고자 한다. 표본은 모집단의 특성을 반영하는 이외에 표본 특유의 일반화될 수 없는 특성들이 있는데, 이는 한 표본자료를 근거로 모형을 개발하는 경우 그 표본에서만 우연히 발견된 중요하지 않은 내용까지도 이론개발에 첨가될 수 있다. 따라서 일단 한 모형을 검증한 다음에는 일반화를 위한 재검증을 반드시 해야 한다.

IV. 결 론

본 연구는 의료자원수급의 적정화를 위한 기초 연구로서 의료자원의 공급과 의료이용에 영향을 미치는 관련 요소들의 관계를 파악하여 의료이용 자체충족률을 공분산구조분석을 활용하여 설명하고자 하였다. 자체충족률에 영향을 미치는 세 가지의 개념, 즉, 지역의 일반여건, 의료자원 공급량 및 공급구조를 측정하는 관찰변수들을 각각 선정하고, 이들 관찰변수와 개념변수들의 상호관계 및 개념변수간의 영향력에 의하여 자체충족률에 미치는 효과를 파악하였다.

공분산구조분석을 활용함으로써 개념변수들 및 관찰변수들 간의 복잡한 상호영향력을

동시에 검증할 수 있게끔 하였으며, 종래의 다중회귀분석에 의한 결과보다 높은 설명력을 보였다. 자체충족률에는 지역의 일반여건이나 의료자원 공급량보다는 의료자원의 공급구조가 상대적으로 큰 영향력을 미치고 있었으며, 의료자원 공급량은 직접적으로보다는 공급구조를 통하여 간접적인 영향을 더욱 크게 미친다. 지역의 일반적 여건은 자체충족률에 직접적으로 영향을 미칠 뿐만 아니라, 의료자원 공급량과 공급구조를 통하여 유의한 간접적인 영향을 갖는다.

설명하고자 하는 변인과 설명되어지는 변인이 복합적으로 얽혀 있는 본 연구의 가설모형과 같은 복잡한 모형을 검증할 수 있는 방법인 공분산구조분석을 활용하여 보건의료정책 분야에서 유용하게 활용될 자체충족률을 의미있게 설명할 수 있었다. 공분산구조분석 방법은 보건의료관리 및 정책 분야에서 제기되는 다양한 문제들에 대한 복합적인 가설 모형을 검증하는데 유용하게 활용될 수 있을 것으로 기대된다.

참 고 문 헌

- 권순호, 한달선, 이규식. 의료수요분석의 방법과 효과에 대한 고찰. 한림대학 사회의학연구소. 1987
- 배상수. 의료이용에 영향을 미치는 요인에 관한 분석. 예방의학회지 1985; 18(1): 13~24
- 유승흠, 조우현, 박종연, 이명근. 도서지역주민의 의료이용양상과 그 결정요인. 예방의학회지 1987; 20(2): 287~300
- 한달선, 권순호. 입원의료의 진료권별 자체충족도에 관한 연구. 예방의학회지 1990; 23(3): 285~295
- Anderson R, Newman JF. Societal and individual determinants of medical care utilization in the United States, *Milbank Memorial Fund Quarterly* 1973; winter: 51~95
- Bentler PM, Chou C. Practical issues in structural modeling. *Sociological Methods & Research* 1987; 16: 78~117
- Cliff N. Some cautions concerning the applications of causal modeling methods. *Multivariate Behavioral Research* 1983; 18: 115~126
- Cohen J, Cohen P. *Applied multiple regression correlation analysis for the behavioral science*, 2nd ed. Hillsdale, New Jersey: Lawrence Erlbaum Associates, 1983
- Feldstein PJ. *Health care economics*. 4th ed., Delmar Publishers Inc., 1993; 7: 5~93
- Grossman M. On the concept of health capital and the demand for health. *Journal of Political Economy* 1972; 80(2): 223~255

- Joreskog KG. A general method for estimating a linear structural equation system. In Goldberger AS & Duncan OD (Eds.), Structural equation models in the social sciences. New York: Seminar press. 1973
- MacCallum R. Specification searches in covariance structure modeling. Psychological Bulletin 1986; 100: 107~120
- Pedhazur EJ. Multiple regression in behavioral research. N.Y.: Holt, Rinehart and Winston Inc., 1982
- Statistical Analysis System. SAS user's guide: Statistics. Cary, NC: SAS Institute, 1996
- Sindelar JL. Differential use of medical care by sex. Journal of Political Economy 1982; 90(5): 1003~1019
- Silvia S. Effects of sampling error and model misspecification on goodness-of-fit indices for structural equation models. Ph.D. Dissertation, Ohio state University, Columbus, Ohio, 1988
- World Health Organization. Health program evaluation: Guiding principles. Health for All Series No. 6. Geneva, 1981
- Wright S. The method of path coefficients. Annals of Mathematical Statistics 1934; 5: 161~215
- <http://www.askhms.com/Askhms/connation.html>
- <http://www.askhms.com/Askhms/connys.html>
- <http://www.askhms.com/Askhms/Priority.html>
- http://www.dhc.simcoe-york.on.ca/ac_phys.htm