

퇴행성 관절염 환자를 대상으로 새로운 진통제 평가를 위한 임상시험자료의 GEE 분석

임희정^{1),2)}, 김윤이¹⁾, 정영복³⁾, 성상철⁴⁾, 안진환⁵⁾, 노권재⁶⁾, 김정만⁷⁾, 박병주^{1),2),8)}

서울대학교 의과대학 예방의학교실¹⁾, 서울대학교 의학연구원 원자력영향역학연구소²⁾,
중앙대학교 의과대학 정형외과학교실³⁾, 서울대학교 의과대학 정형외과학교실⁴⁾, 성균관대학교 의과대학 정형외과학교실⁵⁾,
이화여자대학교 의과대학 정형외과학교실⁶⁾, 가톨릭대학교 의과대학 정형외과학교실⁷⁾, 서울대학교병원 임상시험센터⁸⁾

Analysis of Repeated Measured VAS in a Clinical Trial for Evaluating a New NSAID with GEE Method

Hoi-Jeong Lim^{1),2)}, Yooni Kim¹⁾, Young-Bok Jung³⁾, Sang-Cheol Seong⁴⁾, Jin-Hwan Ahn⁵⁾, Kwon-Jae Roh⁶⁾,
Jung-Man Kim⁷⁾, Byung Joo Park^{1),2),8)}

Department of Preventive Medicine, Seoul National University¹⁾, Institute of Radiation Effect & Epidemiology, Medical Research Center, Seoul National University²⁾, Department of Orthopaedic Surgery, Chung-Ang University³⁾, Department of Orthopaedic Surgery, Seoul National University⁴⁾, Department of Orthopaedic Surgery, Sungkyunkwan University and Samsung Medical Center⁵⁾, Department of Orthopaedic Surgery, Ewha Woman's University Hospital⁶⁾, Department of Orthopaedic Surgery, The Catholic University of Korea⁷⁾, Clinical Trial Center, Clinical Research Institute, Seoul National University Hospital⁸⁾

Objective : To compare the efficacy between SKI306X and Diclofenac by using generalized estimating equations (GEE) methodology in the analysis of correlated bivariate binary outcome data in Osteoarthritis (OA) diseases.

Methods : A randomized, double-blind, active comparator-controlled, non-inferiority clinical trial was conducted at 5 institutions in Korea with the random assignment of 248 patients aged 35 to 75 years old with OA of the knee and clinical evidence of OA. Patients were enrolled in this study if they had at least moderate pain in the affected knee joint and a score larger than 35mm as assessed by VAS (Visual Analog Scale). The main exposure variable was treatment (SKI 306X vs. Diclofenac) and other covariates were age, sex, BMI, baseline VAS, center, operation history (Yes/No), NSAIDS (Y/N), acupuncture (Y/N), herbal medicine (Y/N), past history of musculoskeletal disease (Y/N), and previous therapy related with OA (Y/N). The main study outcome was the change of VAS pain scores from baseline to the 2nd and 4th weeks after treatment. Pain scores were obtained as baseline, 2nd and 4th weeks after treatment. We applied GEE approach with empirical covariance matrix and independent(or exchangeable) working correlation matrix to evaluate the

relation of several risk factors to the change of VAS pain scores with correlated binary bivariate outcomes.

Results : While baseline VAS, age, and acupuncture variables had protective effects for reducing the OA pain, its treatment (Joins/Diclofenac) was not statistically significant through GEE methodology (ITT:aOR=1.37, 95% CI=(0.8200, 2.26), PP:aOR=1.47, 95% CI=(0.73, 2.95)). The goodness-of-fit statistic for GEE (6.55, p=0.68) was computed to assess the adequacy of the fitted final model.

Conclusions : Both ANCOVA and GEE methods yielded non statistical significance in the evaluation of non-inferiority of the efficacy between SKI306X and Diclofenac. While VAS outcome for each visit was applied in GEE, only VAS outcome for the fourth visit was applied in ANCOVA. So the GEE methodology is more accurate for the analysis of correlated outcomes.

J Prev Med Public Health 2004;37(4):381-389

Key Words: VAS pain score, GEE, Goodness-of-fit test for GEE, Osteoarthritis, Clinical Trial

서론

임상연구에서는 한 피험자에 대해서 결과변수를 단지 1회의 방문을 통해서 얻는 것이 아니라 수 회의 방문을 통해 검사를 실시하고 반복측정된 검사결과를 결과변

수로서 사용하는 경우가 많다. 이 경우, 한 피험자가 하나 이상의 결과변수를 갖게 되므로 일변량 이분성 결과변수를 대상으로 하는 분석은 한계를 가지게 된다. 특히, 임상시험과 같이 시험계획서에 개개인의 피험자가 수 회 방문하여 결과변수를 측

정하는 것을 항목화하고 있으며, 그 결과를 바탕으로 신약의 유효성을 판단하는 경우 더욱더 일변량 이분성 결과변수에 대한 분석으로 판단하는 데에는 무리가 따른다고 할 수 있다. 특히, 반복측정이 빈번하게 일어나는 임상시험의 경우에는 반복측정된 다변량 이분성 결과변수 (multivariate binary outcome variable)에 대한 통계분석방법인 GEE 분석방법의 적용이

접수: 2003년 8월 8일, 채택: 2004년 9월 20일

본 연구는 2002년도 한국과학재단 유망여성 과학자 경쟁력 강화 연구지원 #R03-2002-000-00034-0을 받아 수행되었음.

책임저자: 박병주 (110-799 서울시 중로구 연건동 28번지, 전화: 02-740-8325, 팩스: 02-747-4830, E-mail: bjpark@smu.ac.kr)

필요하다. 실제로 임상시험에서 한 피험자 당 반복측정된 자료를 대상으로 이들 간의 상관관계를 고려해 주어야 하는 경우인 퇴행성 관절염 환자를 대상으로 한 새로운 진통제 평가를 위한 임상시험 자료에서 Liang과 Zeger 등이 주창한 GEE (Generalized Estimating Equation) 분석방법을 본 연구에서 적용하고자 한다 [1,2].

GEE 방법은 모수의 추정을 위해서 단기간 관측시간마다 종속변수의 주변분포만을 정의하고 종속변수들 간의 상관관계를 나타내는 가상관행렬(working correlation matrix)을 가정하여 반복측정된 결과변수 간의 상관성을 고려하여 모수를 추정하는 분석방법이다 [1-3]. 의학분야에서는 주로 범주형 자료가 대상이 되며 ‘예’ 또는 ‘아니오’ 혹은 ‘어떤 질병이 있다’ 또는 ‘없다’라는 형태로 이루어진 각각 독립인 일변량 결과변수를 처리하기 위해서 로지스틱 회귀분석 방법이 이용되어왔다. 그러나 각 피험자에 대하여 이분성 결과변수가 수 회 반복측정 되었다면 개개인의 방문 간에는 상관성이 존재할 것이고, 이를 통계적 분석에서 반드시 고려해주어야 한다. 따라서, 이러한 상관관계를 모델에 포함시킬 수 없는 로지스틱 회귀분석보다는 상관관계를 분석과정에서 고려해주는 GEE 분석방법을 적용할 수 있다 [4].

본 연구는 ‘퇴행성 관절염 환자를 대상으로 신약인 조인스정의 디클로페낙 서방정에 대한 비열등성 임상시험’의 수행으로 얻어진 자료를 이용하였다. 이 임상시험은 신약인 조인스 200 mg 정 1 일 3 회 28일 주기 투여방법이 디클로페낙 100 mg 서방정과 비교하여 유효성 및 안전성의 측면에서 대조약인 디클로페낙보다 통계적으로 유의하게 나쁘지 않음을 평가하는 것이었다. 결과변수를 퇴행성 관절염 환자에서의 통증의 완화정도로 하였고, 통증에 대한 평가도구로서 100 mm VAS (visual analogue scale)를 사용하였다 [5]. 임상시험은 기저상태에서 VAS가 35 mm 이상인 환자들을 대상으로 약물 투여 2주 후와 4주 후에 다시 VAS를 측정하였고, 4주 후의 VAS값을 결과변수로 사용하였다. 4주 후의 VAS값이 치료법간의 유의한 차이

가 관찰되지 않는지를 평가하기 위하여 이 VAS값을 종속변수로 하고 치료법, 기저상태에서의 VAS값, 성별 및 연령을 공변량으로 간주한 공분산분석을 실시하였으며, 결과변수인 VAS값이 연속변수이므로 임상시험에 대한 결과분석에서는 방문시점에 따른 영향을 평가하기 위해서 repeated measures ANOVA의 방법을 이용하여 분석하였다 [6].

조인스정은 환자 대부분이 노년층인 퇴행성 관절염의 치료 약제로 진통 효과는 우수하면서 부작용을 감소시킬 필요에 의해 개발되었다. 퇴행성 관절염의 치료약물로 흔하게 사용되는 비스테로이드성 진통소염제는 소화장애, 위장관 장애 및 신기능의 감소 등과 같은 전신적인 부작용을 유발하는데 환자의 연령이 증가할수록 부작용 발생의 빈도가 높아지므로 노년층 환자의 치료시 장기간에 걸친 전신투여는 많은 문제점을 내포하고 있었기 때문이다. 조인스정과 관련된 부작용을 신체계통별로 살펴보면 위장관계 이상이 가장 높은 빈도를 보였으며 이 중 속쓰림, 부종 등이 발생하였으며 조인스정과 관련된성 여부가 불명으로 판정된 경우의 부작용은 소화불량, 변비, 혈변 등의 위장관계 이상과 고혈압, 현훈 등이 있었다. 부작용은 시험약물과 관련성이 있는 부작용 중 조인스정 투여군에서 25건, 디클로페낙 투여군에서 46건으로, 디클로페낙 투여군에서 시험약물과 관련성이 있다고 판단된 부작용이 더 많이 발생하였다. 신체 계통별로 살펴보면 이상반응을 살펴보면 대부분의 항목에서 두 군간 유사한 분포를 보였으나 위장관계 증상 중 속쓰림의 경우 조인스정은 4건(3.2%), 디클로페낙 투여군에서는 18건(14.5%)으로 디클로페낙 투여군에서의 발생분율이 통계적으로 유의하게 높아 ($p=0.01$) 조인스정이 더 안전하다는 결과가 보고된 바 있다 [6].

본 연구의 목적은 이전 임상시험의 결과에서 비열등성이 확인되었으므로 상관성이 존재하는 반복측정이 빈번히 일어나는 이분성 범주형 임상시험자료에서 조인스정은 디클로페낙과 비교하여 퇴행성 관절염 환자에서 무릎 관절의 통증을 완화시

킴에 있어서 효과의 차이가 없는지를 알아보기 위해 GEE 방법을 적용해 보고 기존의 임상시험에서 사용되었던 분석방법인 공분산분석법과 repeated measures ANOVA의 방법으로 얻어진 결과와 비교 검토하고자 하는 것이었다.

대상 및 방법

1. 연구대상

퇴행성 관절염(Osteoarthritis, OA) 환자를 대상으로 조인스(SKI306X)정의 유효성과 안전성을 확인하기 위해 디클로페낙 서방정을 대조약으로 하여 무작위배정, 양측 눈가림에 의한 비열등성 임상시험을 실시하였다 [6]. 본 연구계획서는 다기관 공동 임상시험에 참여한 임상연구자들의 소속 기관 IRB로부터 승인을 받았으며 피험자들로부터 사전 동의를 받았다. 본 연구의 대상자는 35세 이상, 75세 미만의 성인 남녀이며 임상적, 방사선학적 소견에 의해 무릎의 퇴행성 관절염 진단이 가능한 환자로서 100 mm VAS에서 무릎관절부위의 통증이 35 mm 이상인 환자이었다. 피험자 선정기준에 적합한 환자를 대상으로 연구 내용을 설명한 후 동의서를 받아 피험자로 확정하였다. 임신 또는 수유 중이거나 임신 가능성이 있는 여성, 퇴행성 관절염 이외에 기타 중양이 동반된 환자, 사고 등의 외상 후 2차적으로 유발된 퇴행성 관절염 환자 등은 제외하였다.

단측검정을 사용하여 계획당시 제 1종 오류와 제 2종 오류를 각각 5%, 20%로 설정하였으며 두 약물의 치료적 동등성을 인정할 수 있는 범위를 결정하기 위하여 임상시험에 대한 기존의 연구결과에 대한 문헌고찰 및 제2상 임상시험 결과를 이용하여 최종적으로 결과변수인 100 mm VAS의 두 군간의 차이가 5.0 mm 이내일 때 두 약물의 비열등성을 인정하기로 연구계획 단계에서 정하였으며 이 때의 표준편차는 15.0 mm으로 선정하였다. 이러한 과정을 거쳐 산출된 피험자수는 각 군당 112명으로 중도탈락률 10%로 가정하여 각 군당 124명, 총 248명의 피험자를 확보하였다.

2. 자료분석

본 임상시험은 총 5회 방문으로 이루어졌는데 첫 방문은 피험 대상자의 적합성 스크리닝 검사를 위해 시험개시 1-2주 이내에 시행되어졌으며 4주간의 투약기에는 세 번의 방문이 이루어졌다. 그리고 투약 완료 1주 후에 추적관찰 방문이 행해졌다. 이 중 투약기에서 이루어진 개개 피험자별 세 번의 방문(투약 0, 2, 4주 후)을 통해 얻어진 VAS값으로 통증감소여부를 측정하여 주 결과변수로 하고 연속변수인 VAS값, 35 mm를 기준으로 그 이상을 통증이 있는 것으로 간주하고, 그 이하를 통증이 없는 것으로 간주하였다. VAS값 35 mm는 임상전문가들의 패널에 의해 결정되어진 값이며 두 번째 방문인 기저상태에서 35 mm 이상을 선택한 환자만을 연구대상으로 선정하였다.

본 연구는 ITT (Intention-To-Treat) 분석법과 PP (Per-Protocol) 분석법을 모두 실시하였다. ITT 분석법은 선정/제외 기준에 적합하였고 임상시험 참여에 자발적으로 동의한 모든 피험자 중 임상시험계획서의 위반이나 중도 탈락과는 상관없이 무작위배정을 받은 모든 환자를 대상으로 시행되었으며 결과분석의 일반적 원칙으로 하였다. PP 분석법은 무작위배정된 피험자 가운데 연구수행도중 중도 탈락된 피험자와 계획서 위반자를 제외하고 임상시험에 참여하여 연구가 수행되는 동안 연구계획서에 따라 순응도 70%이상인 상태로 충실히 참여한 피험자만을 분석대상으로 삼았다[6].

본 연구에서는 연속변수인 VAS값 결과변수를 GEE 분석방법을 적용하기 위하여 결과변수를 다음과 같이 정하였다. 결과변수인 DIF 변수는 통증감소여부를 측정할 변수로 기저상태에서 VAS값 35 mm 이상인 통증이 있는 환자만을 대상으로 세 번째 방문 측정에서의 VAS값이 35 mm 미만이면 통증이 감소하였다 (DIF=1)고 간주하였으며, 그대로 VAS값이 35 mm 이상이면 변화없음 (DIF=0)으로 간주하였다. 마찬가지로 기저상태와 네 번째 방문시 측정의 VAS값을 비교하여 개개 피험자에 대하여 이변량 이분 결과변수(bivariate binary outcome variable)를 얻을 수 있었다.

이 경우 2개의 결과변수간에는 상관관계가 존재한다고 할 수 있다. 또한 연령, 성별, 체질량지수인 BMI (Body Mass Index), 연구기관, 치료법(조인스정 vs. 디클로페낙), 기저상태 VAS, 침을 맞은 적이 있는지 여부(Y/N), 한약복용 여부(Y/N), 수술경력 여부(Y/N), 퇴행성 관절염 치료나 통증완화와 관련된 약물복용력(Y/N), 스크리닝 과정에서 확인된 피험자들의 1개월 이내 퇴행성 관절염 치료와 관련된 약물복용 여부(Y/N), 그리고 근육골격계 질환에 해당하는 과거병력(Y/N) 등을 독립변수로 하였다. 여기서 BMI란 체중(Kg)을 신장(m)의 제곱으로 나눈 수치이다.

위의 결과변수를 이용하여 각 투여군 간의 통증완화 정도에 대한 시험약물과 대조약물의 차이성을 평가하는 데 있어 잠재적인 교란변수의 영향을 평가하고자 퇴행성 관절염의 통증 감소에 영향을 미치는 인자로 예측되는 위에서 열거된 변수들을 독립변수들로, 퇴행성 관절염으로 인한 통증감소 여부를 종속변수로 하여 1차 단변수 분석을 GEE 방법에 의한 Wald test를 이용하여 유의성 검정을 시행한 결과 p-값이 0.25 이하인 변수들을 선택하였다. 0.25를 변수 선택 기준으로 택한 이유는 선형회귀 분석 시에는 Bendel과 Afiff에 의한 논문 [7]에 근거하였으며 로지스틱 회귀분석 시에는 Mickey와 Greenland에 의한 논문 [8]에 근거하였다. 각 변수에 대한 단변수 분석의 결과, 통계적으로 유의한 변수인 연령, 성별, 기저상태 VAS, 침술 여부, 한약복용 여부, 근골격계 질환 등의 과거병력을 교란변수로 선정하였다. Full model은 결과변수를 통증감소여부(DIF)로, 독립변수들은 치료법, 연령, 성별, 기저상태 VAS, 침술 여부, 한약복용 여부, 근골격계 질환 등의 과거병력으로 하였다. Backward elimination 방법을 사용하여 이중 GEE 방법에 의한 Wald test 결과 p-값이 0.25보다 큰 변수를 제거한 결과, 치료법, 연령, 기저상태 VAS, 침술 여부로 구성된 최종모형을 얻게 되어 역시 GEE 분석법을 이용한 다변량 분석을 실시하여 대응위험도와 95% 신뢰구간을 추정하였다. 최종모형이 잘 적합되었는가를 보이기 위해 GEE

방법을 위한 모형적합도 검정을 시행하였다. 이 방법은 Horton 등 [9]에 의해 소개되었는데 호스머-렘쇼(Hosmer-Lemeshow)의 적합도 검정의 확장으로 볼 수 있다. 여기서 사용된 방법을 간략히 소개하면 전체 관찰값들을 10개의 집단으로 나누어 아래와 같은 예측확률 이 1/10보다 작거나 같은 개체들을 맨 처음 범주로, 예측확률이 구간 (1/10, 2/10) 사이에 속하는 개체들을 그다음 두 번째 범주 등 열 개의 범주로 범주화하는 것이다.

$$\tilde{P}_g = \left(\frac{e^{\hat{\beta}_0 + \hat{\beta}_1 X_{1i}}}{1 + e^{\hat{\beta}_0 + \hat{\beta}_1 X_{1i}}} \right)$$

여기서 X_{1i} 는 우리가 적합하려는 모형에서 치료법, 연령, 기저상태 VAS, 침술 여부, 한약복용 여부로 구성된 독립변수들을 말하며 $\hat{\beta}_i$ 은 절편과 기울기의 모수 추정치이다. 열 개의 범주로 범주화하기 위해 9개의 그룹 indicator가 필요하며 그것을 정의하면 아래와 같다.

$$I_{ig} = \begin{cases} 1 & \text{if } \tilde{p}_{ig} \text{ is in group } g \\ 0 & \text{otherwise} \end{cases} \quad g=1,2,\dots,9$$

위의 그룹 indicator를 가지고 모형적합도를 검정하기 위해 대립모형(alternative model)을 고려하면,

$$\text{logit}(p_i) = \beta_0 + \beta_1 X_{1i} + \gamma_1 I_{1i} + \gamma_2 I_{2i} \text{ 과 같다.}$$

$\hat{\beta}_i$ 은 절편과 기울기의 모수 추정치이며 I_{ig} 은 indicator들의 모수 추정치이다. 그러므로 모형의 적합도를 검정하기 위한 귀무가설은 아래와 같고 이 귀무가설을 기각하지 않을 때 우리가 제시한 모형이 좋은 모형이라고 할 수 있다.

$$H_0: \gamma_1 = \dots = \gamma_9 = 0$$

대립가설이 $H_1: \gamma = [\gamma_1, \dots, \gamma_9] \neq 0$ 일 때 GEE의 score vector는 아래와 같다.

$$u(\beta, \gamma) = \begin{pmatrix} u_1(\beta, \gamma) \\ u_2(\beta, \gamma) \end{pmatrix} = \sum_{i=1}^n \begin{pmatrix} D'_{1i} V_i^{-1} [y_i - p_i(\beta, \gamma)] \\ D'_{2i} V_i^{-1} [y_i - p_i(\beta, \gamma)] \end{pmatrix}$$

where $D'_{1i} = \frac{\partial [p_i(\beta, \gamma)]}{\partial \beta}$ $D'_{2i} = \frac{\partial [p_i(\beta, \gamma)]}{\partial \beta \gamma}$

그리고 V_i 는 한 피험자에서 다변량 결과변수 Y_{ij} , $i=1, \dots, K$, $j=1, \dots, m$ 의 공분산 행렬이며 아래와 같이 분산함수와 가상관행렬로 구성되어 있다.

$$V_i = \Phi A_i^{1/2} R_i(\alpha) A_i^{1/2} \text{로 구성된 } m_i \times m_i \text{ 대각선 행렬이며 분산함수인 } \text{var}(\mu_{ij}) \text{는}$$

$var(Y)$ 가 μ 에 얼마나 의존적인가를 나타내는 것이다. $R_i(a)$ 는 가상관행렬이다. $p_i(\beta, \gamma)$ 는 indicator들을 포함한 대립모형을 이용하여 산출된 예측확률이다.

β 는 귀무가설 하에서 β 의 추정치일 때 귀무가설을 검정하기 위한 검정통계량은 GEE score statistic이나 GEE Wald statistic을 이용하여 구할 수 있는데 여기서는 아래의 score statistic을 가지고 검정을 하였으며 이 통계량은 자유도가 9인 카이제곱분포를 따른다.

$X_2 = u_2(\beta, 0) \{var[u_2(\beta, 0)]\}^{-1} u_2(\beta, 0) \sim \chi^2_9$
이 통계량의 값이 적을수록 모형의 적합도가 높으며 커질수록 모형의 적합도는 떨어진다. 이 모형의 적합도를 검정하기 위해 Full model에서부터 최종모형까지 score 통계량의 값과 p값을 구해서 모형의 적합도가 높은 모형을 찾았다.

GEE에서 공분산행렬의 종류에는 모형을 근거로 하는 공분산행렬(model based covariance matrix)과 샌드위치 공분산행렬(sandwich covariance matrix)이 있다. 샌드위치 공분산 행렬은 견실한(robust) 혹은 경험적인(empirical) 공분산행렬이라고 부른다. 이 두 공분산행렬의 가장 큰 차이점은 모형을 근거로 하는 공분산행렬은 가상관행렬이 바르게 지정되어졌을 때에만 모형의 모수에 대한 추정치와 이 추정치의 분산에 대한 추정치가 일관되고 [9], 샌드위치 공분산행렬은 가상관행렬이 바르게 지정되지 않았을 때에도 Y에 대한 평균에 대한 모델이 올바르다면 위의 추정치들이 항상 일관된다. 연구자가 가상관행렬을 데이터에 맞게 지정해야 하는데 실제로 적합한 가상관행렬을 고르는데 오류가 발생할 수 있으므로 가상관행렬이 바르게 지정되지 않았다 할지라도 항상 일관된 공분산행렬의 샌드위치 추정치를 사용하는 것이 바람직하다. 그러나 이 추정량은 표본크기가 적을 때나 공변량이 기운 분포(skewed distribution)를 가질 때 바이어스가 개입될 수 있다. 본 연구에서는 GEE 추정치를 얻기 위해 샌드위치 공분산행렬을 사용하였다.

GEE 방법론은 모수를 추정하기 위하여 연구자가 데이터의 성질에 맞는 상관행렬

을 지정해야 한다. 실제의 상관행렬을 알 수 없기 때문에 일정한 형태의 가상관행렬을 가정한 후에 이 상관행렬이 참 상관행렬인 것처럼 가정한 후에 모수를 추정하게 된다 [3]. 가상관행렬은 (1) 독립적(independent), (2) m-종속적(m-dependent), (3) 교환 가능한(exchangeable), (4) 자기회귀성(autoregressive(AR-1)), 그리고 (5) 비구조화된(unstructured) 가상관행렬로 분류되고 있다. 또한 연구자 자신이 가상관행렬을 지정(fixed)할 수 있다. (1)의 경우는 같은 개체로부터 다른 시간에 얻어진 종속변수들 간에 서로 독립관계가 성립함을 가정하는 것을 의미하며, (2)의 경우는 m 번째까지는 상관관계가 성립되지만 그 이후는 독립적임을 가정하는 상관행렬이다. 그리고 (3)의 경우는 각각의 다른 시간에 얻어진 종속변수들 간의 상관정도가 α 로 일정할 것이라는 것을 가정하는 것이며, (4)의 경우는 시간이 지남에 따라 상관의 정도가 희미해짐을 의미한다. 마지막으로 (5)의 경우는 실제의 상관행렬을 도저히 상상할 수 없어서 각기 다른 상관정도를 가진다고 가정한 것이다.

본 연구의 경우 이변량 결과변수이므로 가상관행렬은 매트릭스가 된다. 그러므로 (1)의 경우와 (3)의 형태의 가상관행렬로 축약될 수 있다. 많은 분석가들은 GEE 분석에서 독립적인 가상관행렬을 사용하면서 어떤 가상관행렬도 가정하지 않고 있다. 본 연구에서도 독립적인 가상관행렬을 사용하였고 교환 가능한 가상관행렬도 사용하여 비교하였다. 그 이유는 가상관행렬이 바르게 지정되지 않았을 때조차 모수추정치와 그들의 표준오차들이 항상 일관되는 GEE의 성질 때문이다. 독립적 가상관행렬을 이용해서 얻어진 모형을 근거로 하는 공분산행렬의 표준오차는 로지스틱 회귀분석에서 얻어진 표준오차와 일치한다. SAS를 이용해서 얻어진 GEE의 표준오차는 샌드위치 공분산행렬의 각각의 값들을 제공근하여 얻어진 값들이며 어떤 가상관행렬을 쓰더라도 표준오차의 값들은 항상 같다. 본 연구자료는 GEE 분석법을 위해서 SAS8.1의 PROC GENMOD를, 모형의 적합도를 검정하기 위해서는 SAS

MACRO를 이용하여 분석하였다 [10].

결 과

본 연구의 대상자들은 ITT분석 대상자가 248명이었고, 계획서 위반자와 중도탈락된 경우를 제외하고 프로토콜에서 계획된 치료를 순응도 70% 이상으로 받았던 PP분석의 대상자는 194명이었다. PP분석에서 제외한 피험자 54명에 대한 자세한 정보는 다음과 같다.

이 중 중도탈락한 대상자들은 총 35명이었는데, 조인스투여군에서 20명, 디클로페낙 투여군에서 15명이었다. 이 대상자들중 이상반응의 발생으로 인한 중도탈락한 경우가 각 군당 7명씩 총 14명으로 가장 많았으며, 치료효과 부족으로 인한 탈락이 1명씩 있었다. 그리고 임상시험 계획서상에 병용금지약물이었던 스테로이드를 투여 받은 경우가 조인스 투여군에서 1명이 있었으며 추적관찰이 불가능한 경우가 조인스 투여군에서 7명, 디클로페낙 투여군에서 4명이었다. 피험자의 참여거부로 탈락한 경우가 조인스 투여군에서 4명, 디클로페낙 투여군에서 3명이었다. 그리고 프로토콜에서 계획된 치료를 순응도 70% 미만으로 탈락한 대상자들은 총 6명으로 조인스 투여군에서 2명, 디클로페낙 투여군에서 4명이었다. 계획된 방문일 지연으로 탈락한 대상자들은 총 13명이었는데 조인스 투여군에서 7명, 디클로페낙 투여군에서 6명이었다 (Figure 1).

ITT분석 대상자들 중 남성 18명(7.3%), 여성 230명(92.7%)으로 대부분이 여성이었고, 연령별로는 60-69세(41.5%), 50-59세(38.3%), 70-79세(11.7%), 40-49세(8.5%) 순으로 대부분이 50대와 60대로 구성되어 있었다. 각 약물투여군 내에서도 연령의 균형이 비교적 적절하게 이루어졌으며 조인스 투여군에서 평균 60.2세(표준편차 8.0, 범위 39-77), 디클로페낙 투여군에서 평균 59.8세(표준편차 7.0, 범위 46-75)로 평균연령은 두 군간에 통계적으로 유의한 차이가 없었다. 연령은 퇴행성 관절염에 영향을 주는 중요한 요인으로 알려져 있다. 체질량지수(BMI)는 조인스 투여군에서 평

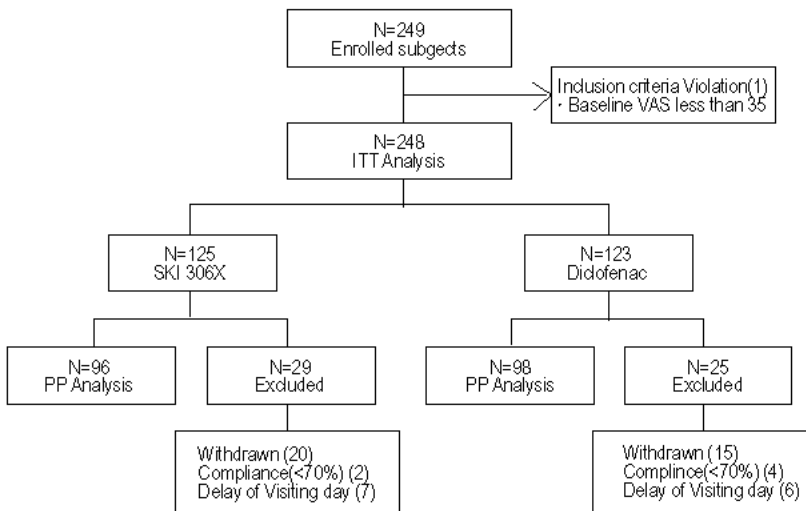


Figure 1. Patient disposition in clinical trial, classified according to ITT analysis and PP analysis excluding withdrawn patients and contact violation.

Table 1. Comparison of the baseline characteristics of the study participants by treatment group on intent-to-treat (ITT) population

Variables	Joins (N=125)	Diclofenac (N=123)	Total	unit:person(%)	p-value
Sex					0.97
Male	9 (7.2)	9 (7.3)	18 (7.3)		
Female	116 (92.8)	114 (92.7)	230 (92.7)		
Age					0.72
<49	12 (9.6)	9 (7.3)	21 (8.5)		
50-59	44 (35.2)	51 (41.5)	95 (38.3)		
60-69	53 (42.4)	50 (40.6)	103 (41.5)		
70+	16 (12.8)	13 (10.6)	29 (11.7)		
mean ±SD	60.1 ± 8.1	59.7 ± 6.7	59.9 ± 7.5		0.66
range	38-77	46-75	38-77		
Body mass index(kg/m ²)	25.7 ± 3.2	25.8 ± 2.8	25.8 ± 3.0		0.76
Baseline VAS	66.8 ± 15.4	65.3 ± 16.5	66.1 ± 15.9		0.45
Center*					1.00
CA	25 (20.0)	24 (19.5)	49 (19.8)		
CM	25 (20.0)	25 (20.0)	50 (20.1)		
EH	25 (20.0)	24 (19.5)	49 (19.8)		
SM	25 (20.0)	25 (20.0)	50 (20.1)		
SN	25 (20.0)	25 (20.0)	50 (20.1)		
Operation history(OA)					0.98
Yes	5 (4.0)	5 (4.1)	10 (4.0)		
No	120 (96.0)	118 (95.9)	238 (96.0)		
NSAID †					0.53
Yes	80 (64.0)	74 (60.2)	154 (62.1)		
No	45 (36.0)	49 (39.8)	94 (37.9)		
Acupuncture					0.45
Yes	76 (60.8)	69 (56.1)	145 (58.5)		
No	49 (39.2)	54 (43.9)	103 (41.5)		
Herbal medicine					0.53
Yes	72 (57.6)	66 (53.7)	138 (55.7)		
No	53 (42.4)	57 (46.3)	110 (44.3)		
Past History ‡					0.81
Yes	7 (5.7)	8 (6.4)	15 (6.1)		
No	116 (94.3)	117 (93.6)	233 (93.9)		
Previous therapy §					0.04
Yes	51 (40.8)	35 (28.5)	86 (34.7)		
No	74 (59.2)	88 (71.5)	162 (65.3)		

* : CA, CM, EH, SM, SN stand for five medical centers in Seoul, Korea.
 † : medication(NSAID) history for osteoarthritis(OA) treatment or reducing the pain.
 ‡ : past history of musculoskeletal disease(MS).
 § : medication history related with OA within one month confirmed by screening procedure.

군 25.7이고 디클로페낙 투여군에서 평균 25.8로 두 군간에 통계적으로 유의한 차이는 없었다 (p=0.76). Baseline VAS값은 66 정도로 두 군간에 차이가 거의 없었으며 퇴

행성 관절염과 관련된 수술을 받지 않은 사람들이 대다수였다 (96.0%). 퇴행성 관절염 치료나 통증완화와 관련된 약물복용력이 있는 사람들이 62.1%로 없는 사람들

보다 더 많았다. 과거에 침을 맞은 경험이 있는 사람이나 한약복용 경험이 있는 사람은 조인스군에서 약간 더 많았다. 근골격계 질환의 과거병력을 가지고 있는 사람은 6.1%로 많지 않았으며 스크리닝 과정에서 확인된 피험자들의 1개월 이내의 퇴행성 관절염과 관련된 약물을 복용하지 않은 사람들이 전체적으로 더 많았다. 본 연구는 다섯 개의 연구기관에서 다기관 공동 임상시험의 형태로 이루어졌다. 스크리닝 과정에서 확인된 피험자들의 1개월 이내의 퇴행성 관절염과 관련된 약물 복용 여부 변수를 제외한 나머지 변수에 대하여는 두 군 간의 차이는 없었다. PP 분석에서도 ITT 분석에서와 거의 비슷한 패턴을 보이고 있었다 (Table 1).

ITT 분석결과 기저상태에서 VAS값 35 mm 이상의 통증을 호소한 환자 중에서 세 번째 방문에서 35 mm 미만으로 떨어진 피험자는 조인스군에서 16.8%, 디클로페낙 서방정군에서 16.3%로 별 차이를 보이지 않았으나, 네 번째 방문에서는 35 mm 미만으로 떨어진 피험자는 조인스군에서 38.4%, 디클로페낙 서방정군에서 32.5%로 조인스군에서 더 높았다. 그러나 세 번째 방문에서는 35 mm 미만이었다가 네 번째 방문에서 35 mm 이상으로 옮겨간 피험자는 거의 없었다(1.6%). 거꾸로 세 번째 방문에서는 35 mm 이상이었던가 네 번째 방문에서 35 mm 미만으로 떨어진 피험자는 전체 248명 중 47명(19%)나 되었다. 여기서 투약 후 방문이 거듭될수록 통증은 전체 피험자의 36%가 줄어든 것으로 나타났다. PP 분석결과에서도 마찬가지로 기저상태에서 VAS값 35 mm 이상의 통증을 호소한 환자 중에서 세 번째 방문에서 35 mm 미만으로 떨어진 피험자는 조인스군에서 12.5%, 디클로페낙 서방정군에서 16.3%로 디클로페낙 서방정군에서 더 높았으나, 네 번째 방문에서는 35 mm 미만으로 떨어진 피험자는 조인스군에서 27.1%, 디클로페낙 서방정군에서 24.5%로 조인스군에서 더 높았다 (Table 2).

VAS값의 세 번째 방문과 기저상태에서의 차이와 네 번째 방문과 기저상태에서의 차이, 즉 두 개의 연속결과변수에 대해

Table 2. Change in VAS of knee osteoarthritis according to each visit and treatment

	Visit 2	Visit 3		Visit 4		Total
		<35	≥35	<35	≥35	
ITT ≥35	Diclofenac	20(16.3%)	103(83.7%)	40(32.5%)	83(67.5%)	123(100%)
	Joins	21(16.8%)	104(83.2%)	48(38.4%)	77(61.6%)	125(100%)
PP ≥35	Diclofenac	16(16.3%)	82(83.7%)	24(24.5%)	74(75.5%)	98(100%)
	Joins	12(12.5%)	84(87.5%)	26(27.1%)	70(72.9%)	96(100%)

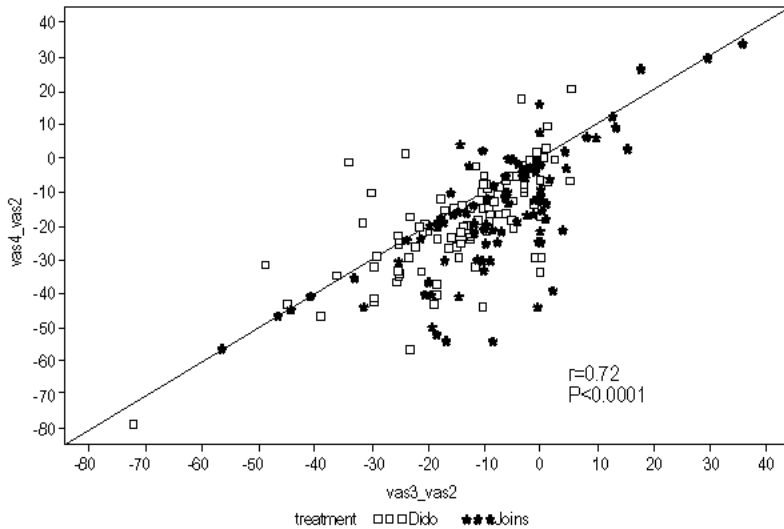


Figure 2. Plot for correlation between VAS3_VAS2 and VAS4_VAS2, where VAS3_VAS2 is the difference of VAS scores between the third visit and baseline, and VAS4_VAS2 is the difference of VAS scores between the fourth visit and baseline (r: correlation coefficient, p-value for $H_0: \rho = 0$).

여 상관계수는 ITT 분석에서는 0.72, PP 분석에서는 0.73으로 두 변수 간에 서로 유의한 선형관계가 있었다 ($p < 0.0001$). 네 번째 방문에서의 VAS값과 두 번째 방문에서의 VAS값의 차이가 세 번째 방문에서의 VAS값과 두 번째 방문에서의 VAS값의 차이보다 큰 것으로 나타났다. 투약 4주까지 통증이 계속해서 감소하였다 (Figure 2).

ITT 분석과 PP 분석을 나누어서 GEE 분석법을 적용하여 얻어진 위험인자들의 대응위험도와 그것의 95% 신뢰구간을 유의수준 0.05에서 살펴보았다. ITT 단변수 분석결과 통계적으로 유의한 변수들을 살펴보면 Baseline VAS값이 1 mm 증가하면 통증이 감소할 대응위험도는 0.96 (95% CI=0.94-0.98)으로 유의했으며 연령에 있어서도 연령이 1세 증가할수록 통증이 감소할 대응위험도는 0.96 (95% CI=0.92-0.99)으로 유의했다. 특히 침술의 과거력이 있는 피험자가 통증이 감소할 대응위험도는 0.43 (95% CI=0.23-0.80)으로 유의했으며 한약복용력이 있는 피험자 역시 통증이 감소할 대응위험도는 0.43 (95% CI=0.23-0.80)으로 유의한 결과를 가져왔

다. PP 단변수 분석결과 역시 유의한 변수들은 ITT 분석결과와 같았으며 GEE 방법에서는 독립 가상관행렬과 교환 가능한 가상관행렬을 사용하였는데 같은 대응위험도와 같은 95% 신뢰구간을 얻었다 (Table 3).

Full model은 결과변수를 통증감소여부(DIF)로, 독립변수들은 치료법, 연령, 성별, 기저상태 VAS, 침술 여부, 한약복용 여부, 근골격계 질환 등의 과거병력으로 하였다. Backward Elimination 방법을 사용하여 이 중 GEE 방법에 의한 Wald test 결과 p값이 0.25보다 큰 변수를 하나씩 제거한 결과, 치료법, 연령, 기저상태 VAS, 침술 여부로 구성된 최종모형을 얻게 되었다. Full model에서 최종모형에 이르는 네 개의 모형들의 적합도를 검정하기 위해 Horton등에 의해 유도되어진 GEE 모형적합도의 통계량과 p값을 얻은 결과 최종모형의 통계량의 값이 가장 적고 p값은 가장 커서 $p=6.55, p=0.68$ 모형의 적합도가 가장 좋았다. 이변량 통증감소여부를 종속변수로서 이 최종모형으로 GEE 방법을 이용하여 다변량 분석을 실시한 결과, 기저상태의

VAS값, 연령, 침술만을 보정한 후에 조인스로 치료한 그룹이 디클로페낙으로 치료한 그룹보다 ITT 분석에서는 1.37배, PP 분석에서는 1.47배 통증이 감소하였으나 통계적으로 유의한 차이를 나타내지 않았다 (ITT와 GEE 분석결과:대응위험도=1.37, 95% CI, 0.82-2.26, PP와 GEE 분석결과:대응위험도=1.47, 95% CI, 0.73-2.95)(Table 4, 5). 이 연구에서의 통증의 감소는 VAS값 35mm이상에서 35mm미만으로 통증의 느낌이 떨어졌다는 것을 의미한다.

고찰

이전의 임상시험은 퇴행성 관절염 환자를 대상으로 디클로페낙의 100 mm VAS값으로 측정된 무릎 관절 부위의 통증 완화 효과에 비해 조인스의 효과가 나쁘지 않음으로 보이기 위한 목적으로 수행된 비열등성 임상시험이었다. 그러나 본 연구에서는 선정기준이 적합한 모든 대상자를 포함하여 실시되는 ITT분석과 계획된 대로 임상시험을 완료한 피험자를 대상으로 한 PP분석 모두 baseline VAS값, 연령, 침술의 과거력을 보정하였을 때, 무릎 관절 부위의 통증 완화 효과는 두군 간에 차이를 보이지 않는 것으로 관찰되었다. 즉, 조인스정은 디클로페낙과 비교하여 퇴행성 관절염 환자에서 무릎 관절의 통증을 완화 시킴에 있어서 효과의 차이를 보이지 않았다.

ITT와 PP 분석을 모두 실시한 이유는 무작위 배정 후에 각종 이유로 중도탈락된 대상자들을 제외하고 분석을 실시할 경우, 배제되는 대상으로 인해 임상시험의 결과가 비뚤어질 수 있기 때문이었다. 특히, 중도탈락의 이유가 배정된 치료와 관련이 있는 경우 더욱 그러한데, 예를 들어 이상반응의 발생으로 인한 중도 탈락의 경우가 시험약에 비해 대조약에서 높은 경우 또는 그 반대인 경우, 중도탈락된 대상자들을 모두 제외하게 된다면, 임상시험에서의 결과는 비뚤어지게 되는 것이다. 따라서, 일단 무작위 배정된 환자들의 경우에 모두 분석에 포함시키는 것이 필요하였다. 본 연구에서는 총 249명의 환자가 무

Table 3. The crude OR and its 95% confidence interval of the baseline characteristics to the OA through GEE analysis

Variables	ITT		PP	
	cOR	95% CI	cOR	95% CI
Baseline VAS	0.96	(0.94,0.98)*	0.92	(0.90,0.95)*
Sex	0.52	(0.23,1.17)	0.48	(0.15,1.47)
Age	0.96	(0.92,0.99)*	0.94	(0.90,0.98)*
BMI	0.97	(0.89,1.05)	0.95	(0.85,1.06)
Center* CA/SN	1.31	(0.67,2.58)	2.15	(0.82,5.63)
CM/SN	0.85	(0.41,1.77)	1.07	(0.38,3.05)
EH/SN	1.08	(0.53,2.21)	1.45	(0.53,3.92)
SM/SN	0.59	(0.28,1.27)	0.52	(0.16,1.67)
Treatment	0.96	(0.52,1.77)	1.18	(0.56,1.92)
Operation History	1.23	(0.45,3.37)	1.62	(0.42,6.31)
NSAID †	0.80	(0.49,1.30)	0.72	(0.39,1.34)
Acupuncture	0.43	(0.23,0.80)*	0.42	(0.26,0.68)*
Herbal medicine	0.43	(0.23,0.80)*	0.51	(0.32,0.83)*
Previous therapy †	0.71	(0.36,1.41)	0.80	(0.48,1.33)

Crude odds ratio and its 95% confidence interval
 * : CA, CM, EH, SM, SN stand for five medical centers in Seoul, Korea.
 † : medication(NSAID) history for osteoarthritis(OA) treatment or reducing the pain.
 † : medication history related with OA within one month confirmed by screening procedure.

Table 4. Goodness-of-fit statistic for GEE obtained using independent working covariance matrix and p-value for selecting the best model

Model	Model Covariates	χ^2	p
1	treatment, baseline VAS, acupuncture, age	6.55	0.68
2	treatment, baseline VAS, acupuncture, age, herbal	7.88	0.55
3	treatment, baseline VAS, acupuncture, age, herbal, sex	13.59	0.14
4	treatment, baseline VAS, acupuncture, age, herbal, sex, MS	11.73	0.23

Table 5. The adjusted OR and its 95% confidence interval of treatment(Joins/Diclofenac) for OA through GEE analysis

Variables	aOR*	95% CI	
Treatment*	ITT	Diclofenac	1.00
		Joins	1.37 (0.82,2.26)
	PP	Diclofenac	1.00
		Joins	1.47 (0.73,2.95)

* Adjusted odds ratio and its 95% confidence interval after adjusting for baseline VAS, acupuncture, age.

작위 배정을 받았으므로 이들 모두가 ITT 분석의 대상자들이어야 하였으나 본 연구에서는 무작위 배정의 대상이 아닌 피험자 1명이 스크리닝 단계에서 걸러지지 못하고 무작위 배정을 받은 사실이 임상시험 완료 후에 발견되어 이 피험자를 제외하고 248명을 대상으로 하였다. 이 환자는 기저상태 VAS값이 35 mm이하로 이미 스크리닝 단계에서 걸려져 무작위 배정을 받아서는 안 되는 피험자였기 때문에 PP분석은 물론 ITT분석에서도 제외하였다. ITT분석과 PP분석에서 기초적인 피험자 분포, 단변수 분석 결과로 인한 다변량 분석결과에서도 차이를 보이지 않았다. 따라서 본 임상시험에서 특정한 피험자의 탈락으로 인한 약물효과에 대한 비뚤림의

개입은 없는 것으로 판단되었다.

임상시험에서는 무작위배정을 통해 시험군과 대조군의 비교성을 높인다. 본 연구에서는 대상자가 조인스 투여군과 디클로페낙 투여군에서 모두 여자가 많았고, 연령은 10세 간격으로 범주화하여 분포를 살펴보았을 때 양군에서 연령 분포가 동일하였으며, 평균 연령 분포에 있어서도 차이를 보이지 않았다. 유효성 비교에 영향을 줄 것이라고 예측되는 체질량 지수, 기저 통증 평가 수준(기저 VAS값), 수술력, NSAID 사용, 침술의 과거력, 한약 복용력, 과거력, 과거 치료 여부에 대해서 조사하였을 때, 거의 모든 항목에서 두군 간의 통계적으로 유의한 차이를 보이지 않아 약물 투여전 인구학적 특성은 두군 간에 비교성이 인정되었다. 따라서, 임상시험의 수행을 위한 적절한 시험군 및 대조군의 선정이라고 할 수 있었다.

임상시험 결과보고서에서는 ITT 분석에서는 조인스 투여군과 Diclofenac 투여군간 통증완화 정도의 차이는 -1.15 mm로 관찰되었고 95% CI=-4.59, 2.29로 최종적인 비열등성 평가지표인 신뢰구간의 하한치(-5)

보다 큰 값이므로 조인스정의 디클로페낙 서방정에 대한 비열등성이 확인되었다. 그러나 PP 분석에서는 두 약물투여군간 차이는 -2.08 mm로 최종적인 비열등성 평가지표인 두 약물간 통증완화 정도의 차이에 대한 95% CI=-5.94, 1.76이므로 신뢰구간의 하한치가 연구계획단계에서 고려한 - δ (-5)보다 더 낮게 평가되어 비열등성을 인정할 수 없었다.

본 연구에서는 단변수 분석을 통해 얻어진 유의한 변수들을 다시 살펴보면 나이가 젊을수록 통증이 빨리 감소하고 기저 상태에서 VAS값을 낮게 찍은 사람들에서 더 빨리 통증이 완화되는 것을 관찰할 수 있었다. 또한 과거에 침술이나 한약을 복용한 경험이 있을수록 투약 후에 통증의 감소를 지연시켰다는 것을 알 수 있었다. ITT 분석을 위한 데이터를 가지고 보조적인 수단으로, 이변량 통증감소여부를 종속변수로 하여 연령, 기저상태의 VAS값, 침술여부 등을 보정한 상태에서도 약물치료법은 통증감소에 유의한 차이가 관찰되지 않는지를 평가하기 위한 GEE 분석에서도 유의한 차이가 관찰되지 않아 앞서 실시한 ITT 분석과 동일한 결론을 얻었다. 공분산 분석에서도 GEE 분석법과 같이 유의한 차이가 관찰되지 않았다. 그러나 PP 분석을 위한 데이터를 가지고 위의 교란변수들을 보정한 상태에서도 약물치료법은 통증감소에 유의한 차이가 관찰되지 않는지를 평가하기 위한 GEE 분석에서도 앞서 실시한 ITT 분석과 동일한 결론을 얻었다. 이 데이터에 대한 공분산 분석 결과 GEE와 같은 결론을 얻었다. 서로 다른 통계적 방법으로 접근하였으나 같은 결과를 얻었음을 확인할 수 있었다.

여기서 사용된 공분산 분석법과 Repeated Measures ANOVA 분석법의 보고서에서 인용된 결과를 자세히 살펴보면, 네 번째 방문에서의 VAS값이 치료법간의 유의한 차이가 관찰되지 않는지를 평가하기 위하여 네 번째 방문에서의 VAS값을 종속변수로 하고 치료법, 기저상태에서의 VAS값, 성별 및 연령을 공변량으로 간주한 공분산분석을 실시하였고 VAS값의 약물투여 후 방문시점변화의 영향을 보정한 상

태에서의 통증완화 효과의 두 치료법간의 차이를 평가하기 위해 Repeated Measures ANOVA 분석을 실시하였다. 공분산분석의 결과 네 번째 방문에서의 VAS값은 앞의 공변량을 보정한 후에도 각 치료법간의 차이가 없는 것으로 확인되었으며(ITT: $p=0.45$, PP: $p=0.13$) Repeated Measures ANOVA 분석의 결과는 방문시점의 변화에 따른 VAS값의 변화는 치료법에 따라 통계적으로 유의하게 다르나 (Wilks' $\Lambda=0.956$, $p=0.004$) 방문시점 변화의 영향을 보정한 상태에서 볼 때 치료법에 따른 VAS값은 통계적으로 유의한 차이가 없었다 ($F=2.45$, $p=0.119$) [6]. 본 연구에서 사용한 GEE 방법은 공분산분석법과 비교해보면 차이점은 GEE 방법은 각 방문시점별 VAS값을 이분화하여 통증감소여부로서 모두 적용하여 결과를 얻는 반면 공분산분석법은 네 번째 연속형 VAS값만을 적용하여 치료법에 따른 유의성에 대한 결과를 얻는다. 그러므로 GEE 방법은 공분산분석법이나 Repeated Measures ANOVA 분석법보다 더 정확한 분석이 이루어졌다고 할 수 있겠다.

세 번째 방문시의 VAS 값과 기저상태에서의 VAS 값의 차이를 결과변수로 했을 때 두 치료법의 차이는 유의하였고 조인스정(평균 12.2 감소)이 디클로페낙 서방정(평균 7.9 감소) 보다 통증이 빨리 감소하였지만 네 번째 방문시의 VAS 값과 세 번째 방문시의 VAS 값의 차이를 결과변수로 했을 때 두 치료법의 차이는 역시 유의하였지만 이번에는 디클로페낙 서방정(평균 7.9 감소)이 조인스정(평균 4.2 감소)보다 더 많이 감소한 것을 알 수 있다. 전체적인 차이, 즉 네 번째 방문시의 VAS 값과 기저상태에서의 VAS 값의 차이를 결과변수로 했을 때 두 치료법의 차이, 조인스정(평균 16.9 감소)과 디클로페낙 서방정(평균 15.7 감소)의 차이는 통계적으로 유의하지 않았다. 위의 결과는 Repeated Measures ANOVA 분석의 결과와 대응되어지는데 역시 같은 결과를 얻었음을 확인할 수 있다 (Table not shown).

본 임상시험은 국내 임상시험 자료에 대해 GEE분석방법 및 GEE에 대한 모형적합

도를 검증하여 적합도가 높은 모형을 찾는 연구라는 점에서 의의를 가질 수 있다. 임상시험의 경우 피험자들로부터 여러 번의 방문을 통해 이변량 혹은 다변량의 결과변수를 얻는 경우가 많다. 흔히 기존의 임상시험에서는 반복 측정된 자료라고 할지라도 각각의 측정에서 얻어진 결과변수가 독립적이라고 가정하고 로지스틱 회귀분석의 방법을 주로 적용하였다. 그러나, 반복측정이 일어난 자료에서는 측정간에 상관성이 존재할 가능성이 있고, 이를 무시할 경우 결과를 비뚤어지게 해석할 수 있다. 개별 피험자들에서 반복 방문을 통해 얻어진 값들이 모두 독립적인 것으로 가정하고 로지스틱 회귀분석을 하게 되면, 실제로 이 값들이 독립적이지 않기 때문에 통계적인 독립성에 위배됨은 물론이고, 종종 유의수준이 과도하게 커지거나 신뢰구간이 지나치게 좁아지는 일이 발생한다고 알려져 있다 [11]. Zhang 등의 연구에서는 근골격계 질병에 대한 연구에서 이분형 결과 변수를 분석할 때 로지스틱 회귀 분석법과 GEE분석방법을 비교하였는데, 로지스틱 회귀분석에 얻어진 대응위험도와 GEE분석방법에서 얻어진 대응위험도가 일치하는데 비해 로지스틱 회귀분석에서의 통계적 검정력보다 GEE분석에서의 검정력이 더 높아짐을 보였다.

조인스정과 같이 퇴행성 관절염에서 통증에 대한 평가를 결과변수로 하는 경우에서 한 피험자의 반복된 평가 결과 간에는 상관성이 존재할 것이다. 그 때문에 근골격계 질환과 같이 한명의 피험자에서 여러 관절에 대하여 측정을 하는 경우, 또는 본 임상시험과 같이 동일 피험자에서 여러 번의 측정을 반복하는 경우 GEE분석방법이 더 적절하다는 것을 보여주었다 [11,12]. 본 임상시험은 이변량 이분성 결과변수를 분석 대상으로 하는 임상시험 자료에서 GEE분석을 적용하였는데, GEE 분석은 이변량에 국한되지 않고 다변량의 경우에도 응용될 수 있다. 또한 이변량 순서형 결과변수에 대해서도 GEE를 가지고 cumulative logit link와 다항분포를 사용하여 proportional odds model을 적합시켜서 모

수의 추정치와 모수의 분산에 대한 추정치를 산출할 수 있는데, 실제 의학 연구자료에 유용하게 적용될 수 있을 것으로 기대되었다.

결론

본 연구에서 사용한 GEE 방법은 임상시험 결과보고서에서 사용했던 공분산분석법과 대응되어지는데 이 두 분석법 모두 교란변수들을 보정한 후에도 약물치료법은 통증의 감소에 유의한 차이를 나타내지 않았으나 차이점은 GEE 방법은 각 방문시점별 VAS값을 이분화하여 통증감소 여부로서 모두 적용하여 결과를 얻는 반면 공분산분석법은 네 번째 연속형 VAS값만을 적용하여 치료법에 따른 유의성에 대한 결과를 얻었다. 그러므로 GEE 방법은 공분산분석법보다 더 정확한 분석이 이루어졌다고 할 수 있겠다.

감사의 글

조인스(SKI 306X)정의 유효성과 안전성을 확인하기 위한 본 임상시험에 참여하여 데이터 구축에 도움을 주신 관계자 여러분께 감사드립니다.

참고문헌

1. Liang KY, Zeger SL. Longitudinal data analysis using generalized linear models. *Biometrika* 1986; 73: 13-22
2. Zeger SL, Liang KY. Longitudinal data analysis for discrete and continuous outcomes. *Biometrics* 1986; 42: 121-130
3. Shin DY, Park TS. A Study on the Use of Working Correlation Matrices in the GEE Approach to the Analysis of Repeated Binary Data. *Appl Statistic* 1996; 11: 15-27 (Korean)
4. Heimendinger J, Feng Z, Emmons K, Stoddard A, Kinne S, Biener L, Sorensen G, Abrams D, Varnes J, Boutwell B. The working well trial: baseline dietary and smoking behaviors of employees and related worksite characteristics. The working well research group. *Prev Med* 1995; 24: 180-193
5. Gaston JF, Gustafsson M. Rheumatoid arthritis: determination of pain characteristics and comparison of RAI and VAS in its measurement. *Pain* 1990; 41: 35-40

6. Jung YB, Seong SC, Lee MC et al. A four-week, randomized, double-blind trial of the efficacy and safety of SKI306X, a herbal antiarthritic agent versus diclofenac in osteoarthritis of the knee. *Clinical Trial Report of SK Pharm* 2000; 11-74 (Korean)
7. Bendel RB, Afifi AA. Comparison of stopping rules in forward regression. *JASA* 1977; 72: 46-53
8. Mickey J, Greenland S. A study of the impact of confounder-selection criteria on effect estimation. *Am J Epidemiol* 1989; 129: 125-137
9. Horton NJ, Bebchuk JD, Jones CL et al. Goodness-of-fit for GEE: An example with mental health service utilization. *Statist Med* 1999; 18: 213-222
10. SAS/STAT User's Guide, Versoion 8, Cary, NC, USA: SAS Institute Inc.; 1999
11. Zhang Y, Glynn RJ, Fekson DT. Musculoskeletal disease research: should we analyze the joint or the person? *J Rheumatol* 1996; 23: 1130-1134
12. Koo HW, Kwark M, Lee Y, Park BJ. Comparison of Efficiency between Individual Randomization and Cluster Randomization in the Field Trial. *Korean J Prev Med* 2000; 33(1): 51-55