

지리적 가중회귀모형을 이용한 지역별 심정지 발생요인에 관한 연구

박 일 수
(위덕대학교)

김 은 주
(인제대학교)

홍 성 옥
(질병관리본부)

강 성 흥*
(인제대학교)

본 연구에서는 질병관리본부에서 수집한 2009년 심정지 조사 자료 및 지역사회건강 조사 자료에 생태학적 방법론 및 지리적 가중회귀모형(GWR)을 적용하여 지역단위의 심정지 발생 위험요인을 규명함은 물론, 지역별 심정지 예방관리 사업에 GWR 모형 적용의 유용성도 함께 검토하고자 하였다. 그 결과, 지역별 심정지 발생의 주요 요인으로 중등도 신체활동 실천율, 비만율, 고혈압 평생 의사 진단 경험률 및 당뇨병 평생 의사 진단 경험률이 도출되었고, 지역별로 일부 차이가 존재하나 대체로 이들이 높을수록 심정지 발생률이 높게 나타났다. 모형개발에 사용된 자료의 한계로 인해 본 연구에서 도출된 GWR 모형의 설명력은 낮으나, 모형 부합도를 비교한 결과 전통적인 OLS 모형 보다는 우수한 모형인 것으로 나타났다. 개인별 자료 확보 대비 지역사회 단위 자료 확보의 용이성 및 GWR 모형이 가지는 특성 등을 고려해볼 때, 향후 이와 관련된 후속연구가 지속적으로 진행된다면 각 지역별 심정지 발생 예방관리를 위한 우선 사업 순위 선정에 본 연구에서 접근한 방법이 매우 유용할 것으로 판단된다.

주요용어: 심정지 발생률, 생태학적 방법론, OLS모형, 지리적 가중회귀모형

본 연구는 전진호 외(2012)의 '심정지 발생 및 생존 관련 요인의 자료수집 및 관리체계 구축' 연구보고서(질병관리본부 발주)에 수록된 내용의 일부를 수정 및 보완한 것이다.

*교신저자: 강성흥, 인제대학교(hcmkang@inje.ac.kr)

■ 투고일: 2013.6.11 ■ 수정일: 2013.8.26 ■ 게재확정일: 2013.9.4

I. 서론

심정지(Cardiac Death) 또는 급성심정지(Sudden Cardiac Death)는 초기 대응 및 치료가 예후에 중요한 영향을 미치는 응급질환으로 최근 심혈관 질환의 증가와 인구의 고령화로 인해 그 발생환자 수가 점차 증가하고 있다. 우리나라의 연도별 심정지 표준화 발생률을 살펴보면, 인구 10만명당 2007년 39.7명, 2008년 41.4명, 2009년 44.4명으로 국내 심정지 발생은 증가하고 있으나 생존율은 높지 않다. 2009년 119 구급대에 의해 이송된 21,530명의 심정지 환자 중 퇴원 시 생존자는 757명(3.5%)이며, 특히 뇌기능이 회복되어 정상적인 생활이 가능한 퇴원은 219명(1.0%)에 불과하다. 이에 효과적으로 대응하기 위해서는 심정지 발생 관련 위험요인을 규명하고 이를 기반으로 한 예방사업 수행이 필요하다(최정아, 2012).

여러 선행 연구들을 통해 밝혀진 심정지 발생 관련 위험요인들을 살펴보면, 크게 인구사회경제학적 요인, 질환 요인, 생활행태 요인으로 분류할 수 있다. 인구사회경제학적 위험요인으로는 연령, 소득, 교육수준, 독신가구 등이 주요한 것으로 알려져 있다(Strogatz et al., 1988; Reinier et al., 2006; Ong et al., 2011; Sasson et al., 2012). 질환위험요인으로는 허혈성 심장질환, 죽상경화증, 심근병증, 심장판막증, 선천성 심장질환, 당뇨병, 고혈압 등이 있다(Myerburg, 2001; Engdahl et al., 2002; Jouven et al., 2005; Shamseddin & Parfrey, 2011). 생활행태 위험요인으로는 흡연, 음주, 격렬한 운동, 커피섭취 등으로 나타났다(Wannamethee & Shaper, 1992; Albert et al., 1999; Goldenberg et al., 2003; Selb Semerl & Selb, 2004; Wisten & Messmer, 2005; Dahabreh & Paulus, 2011). 그러나 현재까지 진행된 심정지의 위험요인에 대한 연구들은 단편적인 연구로서 심정지의 주요 위험요인을 동시에 고려하여 상대적인 위험의 정도를 파악하지 못한 제한점이 있다. 또한, 심정지 발생환자에 대한 관리는 지역사회주민을 대상으로 하는 지역사회단계 관리로 이루어지는 것이 보다 효과적이다(전진호, 2012). 미국 아리조나주에서는 심정지 관리에서 지역사회의 역할을 확대 시행함에 따라 생존율이 2005년 3.7%에서 2009년 9.8%로 증가하였고(Bobrow, 2010), 스웨덴에서는 1992년 심정지 환자의 입원율이 15.3%에서 2005년 21.7%로 증가하였는데, 그 이유가 지역사회단계에서의 심정지 환자관리가 향상되었기 때문이라고 보고하고 있다(Hollenberg, 2008). 현재 우리나라에서도 심정지 환자에 대한 본격적인 연구를 위해

2009년부터 질병관리본부에서 심정지 환자 의무기록조사와 등록관리사업을 통한 자료 수집 및 기초연구가 이루어지고 있다. 이 자료를 활용한 신상도 외(2011) 연구 결과에서는 지역별 심정지 표준화 발생률(2009년 기준, 인구 10만명당)이 서울(36.0명), 부산(37.6명), 대구(38.3명) 순으로 낮게 나타난 반면, 제주(75.1명), 강원(61.2명), 충북(54.0명) 지역의 경우 타 지역과 비교했을 때 두드러지게 높은 발생률이 나타남을 근거로 효과적인 심정지 관리를 위해서는 지역 특성에 맞는 지역별 심정지 예방 사업 수립이 중요함을 강조하였다.

위의 제시한 내용들을 종합해보면, 지역별 맞춤형 심정지 예방 사업을 수행하기 위해서는 우선 각 지역별 심정지 발생 위험요인을 규명할 필요가 있다. 그러나 지역주민 개개인의 심정지 위험요인에 대한 자료를 수집하는데 현실적인 어려움이 있음에 따라 개인단위가 아닌 지역사회 단위의 자료로 질병의 위험요인과 질병과의 관련성을 보는 생태학적 접근방법론(Ecological Study)의 활용이 필요하다. 더불어 특정질환의 발생 위험요인을 규명하고자 할 때, 흔히 최소제곱법을 통해 모수를 추정하는 OLS분석(Ordinary Least Squares)과 같은 일반적인 회귀분석 방법이 사용된다. 그러나 지역사회 단위로 수집된 자료일 경우, 모든 지역이 동일한 특성을 가짐을 전제로 하는 일반적인 회귀분석 보다는 각 지역별 특성이 질병 발생 위험요인 규명을 위한 통계적 모형에 고려되는 것이 보다 효과적이다.

이에 본 연구는 지역사회단위의 심정지 환자 관리 사업을 위한 기반연구로서 생태학적 접근방법론을 통해 수집된 지역사회단위 자료를 활용하고, 심정지 발생 위험요인 규명에 각 지역별 특성을 반영시키기 위해 지리적 가중회귀모형(GWR; Geographically Weighted Regression, 이하 GWR)이라는 통계적 모형을 활용하고자 한다. 또한, 이를 통해 지역사회의 심정지 발생 위험요인을 규명함과 동시에 지역별 심정지 예방관리 사업에 있어 GWR모형 적용의 유용성도 함께 검토하고자 한다.

II. 연구방법

1. 지리적 가중회귀모형

보통 지역단위의 자료 분석에 회귀분석이 흔히 사용되는데, 일반최소제곱법(OLS)으로 추정되는 전통적 회귀분석은 관찰들 간의 독립성과 오차의 동분산성(homoscedasticity)을 가정하므로 공간적 변이의 측면을 고려하지 못한다. 즉, 전통적 회귀분석에서는 표집 위치에 상관없이 관찰들은 서로 독립적이어야 하고, 종속변수의 관찰값과 추정값의 차이인 오차(errors)가 상호 독립적이며 분산이 일정한 것으로 가정한다(Berry & Feldman, 1985; Fox, 1991; 김두섭·강남준, 2008). 그러나 공간적으로 근접한 위치에 표집된 사례일수록 유사한 값을 가지는 경향이 있기 때문에 현실적으로 이러한 가정이 충족되기는 어렵다. 따라서 공간적 이질성 또는 의존성이란 특성들이 존재하는 현상을 OLS 회귀모형으로 분석하게 되면 모수 추정치의 효율성이 떨어지게 된다(조동기, 2009). 이러한 이유로 공간적 변이를 고려하는 공간통계분석기법이 발전되어 왔는데, 그 중에서 Fotheringham et al.(2002)에 의해 개발된 지리적 가중회귀분석이 대표적인 예라 할 수 있다. 지리적 가중회귀분석의 경우, 회귀모형을 사용한다는 점에서 기존 모형들의 장점이 적용됨과 동시에 지역별로 각각 다른 계수의 추정이 가능하여 기존 모형들로는 파악하기 어려운 공간적으로 이질적인 패턴을 확인할 수 있다. 또한, 지리적 가중회귀분석 결과를 GIS(Geographic Information System) 맵을 통해 제시함으로써 보다 시각적으로 데이터들의 공간적 상호작용을 파악할 수 있다(Fotheringham et al., 2002; 강승모, 2011).

다음은 전역적 통계방법(OLS)과 지역적 통계방법(GWR)을 비교한 내용이다.

표 1. OLS와 GWR 비교

전역적 통계 방법 (OLS)	지역적 통계 방법 (GWR)
전체 연구지역의 자료 요약	전역적 통계량의 국지적 분해
하나의 통계량	각 연구대상 지역별 통계량
지도화 할 수 없음	지도화 할 수 있음
GIS에 비친화적	GIS에 친화적
비공간적 혹은 공간적으로 제약	공간적
공간적 동질성을 강조	공간적 이질성을 강조
규칙성 혹은 법칙 탐색	예외 혹은 지역적 특징 탐색

지리적 가중회귀분석의 핵심은 회귀계수를 상수가 아니라 위치의 함수로 간주하는 것이다. 즉, 회귀계수가 모든 지역에서 동일한 것이 아니라 공간적 위치에 따라 달라지도록 모형화 하는 것이다. 알려진 어떤 위치 μ 에 대해 종속변수 y 와 m 개의 독립변수 x 가 있고, 회귀계수와 오차항이 각각 β , ϵ 인 경우 회귀식은 아래 식 (1)과 같이 표현된다(조동기, 2009).

$$y_i(\mu) = \beta_{0i}(\mu) + \beta_{1i}(\mu)x_{1i} + \beta_{2i}(\mu)x_{2i} + \dots + \beta_{mi}(\mu)x_{mi} + \epsilon_i(\mu) \dots\dots (1)$$

y_i = 종속변수 ($i = 1, 2, \dots, n$), n 은 관측치

x_{mi} = m 번째 독립변수

β_{mi} = m 번째 회귀계수(파라미터)

ϵ_i = 오차항

μ = 공간좌표

식 (1)이 일반적인 회귀모형과 다른 점은 각 항에 (μ)가 붙는 것인데, 이는 모수의 추정이 주어진 위치 μ 에 대해 이루어지고 그 위치에 대해서만 의미를 가진다는 것을 나타낸다.

지점에 따른 회귀계수의 추정식은 가중최소자승법에 따라 이뤄지며, 이에 대한 식은 다음과 같다.

$$\hat{\beta} = [X^T W(u) X]^{-1} X^T W(u) Y \dots\dots\dots (2)$$

이 식에서 지리적 가중치 행렬 $W(u)$ 의 각 요소는 가중함수인 커널(kernel)에 따라 계산된다.

커널은 다양한 형태를 취할 수 있지만 일반적으로 가우스 형태인 식(3)이 사용된다.

$$w_i(u) = e^{-0.5(d_i(u)/h)^2} \dots\dots\dots (3)$$

$w_i(u)$ = 공간좌표에서의 관측치 i 에 대한 가중치

$d_i(u)$ = 관측치 i 와 공간좌표 μ 간의 거리

h = 대역폭

커널은 가중치를 만드는 대역폭(Bandwidth)이 고정되어있는 고정방식(Fixed spatial kernel)과 관측치의 수에 따라 다른 대역폭을 적용하는 가변방식(Adaptive spatial kernel)이 있다. 관측치가 조사지역에 규칙적으로 분포할 경우 고정방식을 사용하고, 조사지역 주위에 관측치의 분포가 다양할 경우는 가변방식을 사용한다. 그러나 이에 대한 판단이 확실치 않을 경우, 대개 가변방식을 채택한다. 또한, 적정 대역폭을 설정하기 위해 AIC(Akaike Information Criterion) 또는 CV(Cross Validation)가 사용되는데, 관찰값과 추정값의 차이 및 모형의 복잡성을 고려하는 AIC방식이 선호된다(Charlton & Fotheringham, 2009).

2. 연구자료와 변수설명

본 연구를 수행하기 위하여 질병관리본부의 2009년 심정지 조사 자료와 지역사회건강조사 자료를 수집하였다. 심정지 조사 자료는 119 구급대에 의하여 응급의료센터로 이송된 심정지 환자에 대한 병원 전단계 구급자료 및 병원단계 의무기록조사 결과를 포함하는 자료로서 기본적인 역학조사를 위하여 활용되는 자료이다. 또한, 지역사회건강조사는 우리나라 국민의 건강 수준과 건강 결정요인에 대한 통계를 기초 자치단체(시·군·구) 단위로 얻기 위해 2008년부터 실시해온 단면조사로서, 대상 지역 만 19세 이상 주민을 대상으로 훈련된 조사원이 표본가구를 직접 방문 하여 조사가 이루어진다.

심정지 조사 자료 및 지역사회건강조사 자료를 이용하여 2009년 기준 전국 240개 시군구별(제주도 제외) 심정지 발생에 관한 데이터 셋을 <표 2>와 같이 구성하였다. 분석에 활용된 독립변수는 2009년 지역사회건강조사를 통하여 수집된 자료를 중심으로 조사 자료에 제시된 개인별 가중치를 적용하여 산출한 것으로, 성별 흡연율, 성별 고위험 음주율, 건강검진율, 중등도 신체활동 실천율, 비만율, 지역사회 내 운동프로그램 참가율, 스트레스 인지율, 우울증 경험률, 각 질환별 평생의사진단 경험률 그리고 각 질환별 본인인지 유병률 등 총 22개로 구성하였다. 각 지역별 심정지 발생요인을 규명하기 위해 활용된 종속변수는 2009년 심정지 조사 자료를 이용하여 산출한 지역별 심정지 표준화 발생률로서, 시군구별로 성별 및 각 연령대별¹⁾ 인구수에 성별 조발생률(지역별 심정지 발생자수/전체 인구수)을 곱하여 구하여진 심정지 기대발생자수의 합을 표준화 인구수(시군구별 전체 인구수)로 나누어 산출하였다.

1) 연령대는 30대 이하, 40대, 50대, 60대, 70대 이상으로 구분함.

3. 분석방법

본 연구는 생태학적 연구방법론을 활용하여 자료를 구성하였다. 생태학적 연구방법은 역학적 연구방법론의 하나로서 개인단위가 아닌 지역사회 단위의 자료를 이용하여 질병의 위험요인과 질병과의 관련성을 보는 방법론이다. 이 방법을 사용하는 이유는 개인단위의 위험요인 및 질병발생에 관한 자료가 없고, 지역사회 단위의 자료만 있는 경우에 간접적인 방법으로 질병 위험요인과 질병 발생의 관련성을 분석하기 위함이다(Durrheim & Ogunhanjo, 2004).

이러한 지역사회단위의 자료로만 구성된 경우, 공간적 이질성을 고려하는 분석과정에는 흔히 전역적 통계모형과 지역적 통계모형으로 구분하여 분석할 수 있다. 전역적 통계모형은 공간적 위치에 따른 차이가 없다는 가정에 따라 설정된 모형이며, 지역적 통계모형은 전역적 통계모형을 공간적으로 분해함으로써 공간적 위치에 따른 차이를 고려하는 모형이다. 어떤 변수의 효과가 모든 지역에서 일정하다고 전제하는 전통적 OLS 회귀분석은 전역적 통계모형에 해당되며 지리적 가중치, 커널, 그리고 대역폭이 결합된 GWR은 지역적 통계모형에 해당된다(조동기, 2009). 본 연구에서는 공간적 이질성을 고려하여 지역별 심정지 발생요인을 규명하고자 함에 따라 GWR을 사용하였으며, 분석 수행에는 ArcGIS 10.1을 활용하였다. GWR 모형 설정 시, 지리적 가중치 부여를 위한 커널함수의 유형은 조사대상지역에서의 관측값들이 불규칙적으로 분포함을 가정하여 분포가 조밀한 곳에서는 커널이 작아지고 산발적인 곳에서는 커지도록 해 어느 위치에서나 동일한 관측치 수들이 포함되도록 대역폭을 설정하는 가변방식을 사용하였으며, 적정 대역폭 설정에 있어서는 AIC를 활용하였다.

분석에 활용된 최종 자료에서 제주지역(서귀포시, 제주시)은 제외하였는데, 그 이유는 제주지역의 경우 지리적으로 떨어진 도서지역이기 때문에 분석결과에 대한 신뢰성 자체가 확보되지 못할 뿐 아니라, 이로 인해 전체적인 모형의 부합도 또한 떨어뜨릴 수 있기 때문이다²⁾.

²⁾ 실제 제주도를 포함하여 분석한 결과와 포함하지 않았을 때의 결과에서 모형 부합도(R^2)의 차이가 존재하였음.

표 2. 변수의 정의

구분		설명	
독립변수	남자 현재 흡연율	평생 5갑(100개비)이상 흡연한 사람으로서 현재 흡연하는 사람 (“매일 피움” 또는 “가끔 피움”)의 비율(%)	
	여자 현재 흡연율		
	남자 고위험 음주율	최근 1년 동안 음주한 사람 중 한 번의 술자리에서 7잔 이상을 주 2회 마신다고 응답한 사람의 비율(%)	
	여자 고위험 음주율	최근 1년 동안 음주한 사람 중 한 번의 술자리에서 5잔 이상을 주 2회 마신다고 응답한 사람의 비율(%)	
	건강검진 수진율	최근 2년 동안 건강을 위해 건강검진을 받은 사람의 비율(%)	
	중등도 신체활동 실천율	최근 1주일(7일) 동안 평소보다 몸이 조금 힘들거나 숨이 약간 가쁜 중등도 신체활동을 1회 30분 이상, 주 5일 이상 실천한 사람의 비율(%)	
	비만율	체질량 지수(체중(kg)/신장 ² (m ²)) 25이상인 사람의 비율(%)	
	지역사회 내 운동 프로그램 참가율	최근 1년(365일) 동안 본인이 거주하고 있는 지역의 동사무소, 보건소, 구청, 시청 등의 지방자치단체 운동프로그램에 참여한 적이 있다고 응답한 사람의 비율(%)	
	스트레스 인지율	평소 일상생활 중 스트레스를 “대단히 많이” 또는 “많이” 느끼는 사람의 비율(%)	
	우울증 경험률	최근 1년(365일) 동안 연속적으로 2주(14일) 이상 일상생활에 지장이 있을 정도의 우울감을 경험한 사람의 비율(%)	
	평생 의사 진단 경험률	고혈압	의사에게 해당 질환을 진단 받은 사람의 비율(%)
		당뇨병	
		뇌졸중	
		심근경색	
협심증			
고지혈증			
본인인지 유병률	고혈압	해당 질환을 지금까지 앓은 적이 있는 사람의 비율(%)	
	당뇨병		
	뇌졸중		
	심근경색		
	협심증		
	고지혈증		
종속변수	심정지 표준화 발생률	발생률 = 시군구별 심정지 전체 환자 발생자수/시군구별 전체 인구 ※ 단, 지역별 성별 연령 표준화시킴(인구 10만명당)	

주: 각 변수는 2009년 기준 총 240개의 시군구별 자료로 구성됨(제주도 제외).

GWR 모형에 활용된 독립변수 선택에 있어서는 선행연구를 통해 심정지 발생에 영향을 미치는 요인을 파악한 후, 그들 중 다중공선성이 없는 요인들로 구성하여 모형에 최종 활용하였다. 또한, GWR 분석의 효과성을 검증하기 위해 동일한 독립변수를 활용하여 OLS 분석을 실시하였고, 그 모형의 부합도를 판단하기 위해 R^2 과 AIC 지수³⁾를 사용하였다.

III. 연구 결과

1. 지역의 일반적 특성

분석 자료에 대한 각 주요 변수별 분포를 살펴 본 결과, 평균 남자 현재흡연율은 48.11%, 여자 현재 흡연율은 3.68%, 남자 고위험 음주율은 25.52%, 여자 고위험 음주율은 5.86%, 건강검진 수검률은 59.3%, 중등도 신체활동 실천율은 12.62%, 비만율은 22.70%, 지역사회 내 운동 프로그램 참가율은 5.36%, 스트레스 인지율은 26.74%, 우울증 경험률은 7.23%로 나타났다. 고혈압, 당뇨병, 뇌졸중, 심근경색, 협심증, 고지혈증에 대한 평생 의사 진단 경험률의 전국 평균은 각각 17.34%, 6.38%, 1.78%, 0.95%, 1.28%, 5.89%였으며, 6개 질환에 대한 본인인지 유병률의 전국 평균은 18.01%, 6.54%, 6.30%, 1.84%, 1.03%, 1.49%로 나타났다. 또한, 심정지 표준화 발생률의 전국 평균은 인구 10만명당 47.4로서, 최소 심정지 발생률은 19.3인 경상남도 고성군, 최고 높은 심정지 발생률을 나타낸 지역은 인천광역시 동구로서 71.42였다.

³⁾ AIC지수는 Akaike(1974)가 제안하였으며, 각 모형들을 비교하기 위해 사용되기 때문에 AIC값의 절대량은 의미가 없고 AIC값이 상대적으로 낮으면 더 나은 모형이라 평가함. 또한 일반적으로 두 모형에서 AIC값의 차이가 4보다 작은 경우 두 모형은 사실상 차이가 없는 것으로 간주함(Charlton & Fotheringham, 2009).

표 3. 지역별 일반적 특성

구분	최소값	최대값	평균	표준편차	
남자 현재 흡연율(%)	31.18	60.90	48.11	4.60	
여자 현재 흡연율(%)	0.29	8.34	3.68	1.46	
남자 고위험 음주율(%)	6.26	39.66	25.52	5.93	
여자 고위험 음주율(%)	0.91	13.48	5.86	2.54	
건강검진 수검율(%)	49.35	70.72	59.30	4.07	
중등도 신체활동 실천율(%)	2.24	61.25	12.62	8.54	
비만율(%)	14.65	30.24	22.70	2.78	
지역사회 내 운동 프로그램 참가율(%)	1.19	46.24	5.37	4.25	
스트레스 인지율(%)	7.47	39.35	26.70	5.58	
우울증 경험률(%)	1.11	14.94	7.21	2.66	
평생 의사 진단 경험률	고혈압(%)	10.05	28.27	17.34	4.02
	당뇨병(%)	2.17	9.91	6.38	1.49
	뇌졸중(%)	0.22	4.52	1.78	0.78
	심근경색(%)	0.07	2.36	0.95	0.46
	협심증(%)	0.08	3.45	1.28	0.50
	고지혈증(%)	1.56	10.87	5.89	1.85
본인인지 유병률	고혈압(%)	10.53	28.49	18.01	3.96
	당뇨병(%)	2.17	10.10	6.54	1.51
	뇌졸중(%)	1.80	12.19	6.30	1.98
	심근경색(%)	0.22	4.52	1.84	0.81
	협심증(%)	0.07	2.56	1.03	0.49
	고지혈증(%)	0.11	4.51	1.49	0.61
심정지 표준화 발생률 (인구 10만명당)	19.30(경남 고성)	71.42(인천 동구)	47.42	10.96	

2. 지역사회 단위의 심정지 발생 위험요인 분석

지역사회 단위의 심정지 발생 위험요인을 규명하기 위해 앞서 제시한 바와 같이 GWR 모형을 활용하였다. GWR 분석 시, 독립변수들 간에 다중공선성이 발생할 경우 결과 값이 도출되지 않기 때문에 본 연구에서는 독립변수들 간 다중공선성을 사전에 검토하여 다중공선성이 발생하지 않는 요인들로 최종 독립변수를 구성하여 분석하였다. 또한, 일반적인 회귀모형에 비해 GWR 모형이 보다 유용한지를 평가하기 위하여 동일한 최종 독립변수를 적용하여 도출된 OLS분석 결과를 함께 제시하였다.

2009년 기준 제주도를 제외한 전체 시군구 단위별 심정지 발생 위험요인 분석에 있어 GWR 분석 수행 결과, 주요 변수로는 중등도 신체활동 실천율, 비만율, 고혈압 평생 의사 진단 경험률, 당뇨병 평생 의사 진단 경험률이 선정되었으며, 이들로 구성된 총 240개의 지역별 회귀방정식이 도출되었다. 도출된 GWR 모형의 전반적인 설명력(R^2)은 0.218로, 동일한 독립변수에서 도출된 OLS의 설명력(0.144)보다 높게 나타났으며, 모형 부합도를 나타내는 또 하나의 기준인 AIC값에서는 GWR 모형은 1795.548로 나타나 OLS(1804.662) 모형에 비해 모형의 부합도가 우수한 것으로 나타났다.

표 4. OLS분석결과 및 지리적 가중회귀분석 결과 비교

독립변수	GWR			OLS		VIF
	회귀계수			회귀계수	t	
	최소값	중앙값	최대값			
절편	19.430	28.268	49.467	29.220	5.145	-
중등도 신체활동 실천율	0.182	0.318	0.474	0.293	3.431	1.216
비만율	-0.498	0.161	0.445	0.155	0.621	1.103
고혈압 평생 의사 진단 경험률	-0.446	0.087	0.933	0.099	0.365	2.701
당뇨병 평생 의사 진단 경험률	-0.038	1.677	2.591	1.451	2.056	2.517
모형적합도	$R^2=0.218$ $R^2_{adj.} = 0.173$ AIC = 1795.548			$R^2=0.144$ $R^2_{adj.} = 0.129$ AIC = 1804.662		-

GWR 분석 결과에 따르면, 지역별로 차이는 있지만 대체로 중등도 신체활동 실천율, 비만율, 고혈압 평생 의사 진단 경험률, 당뇨병 평생 의사 진단 경험률이 높을수록 해당 지역의 심정지 발생률은 높게 나타났다. 그러나 중등도 신체활동 실천율을 제외한 비만율, 고혈압 및 당뇨병 평생의사진단 경험률의 경우는 일부지역에서 그 값이 낮을수록 심정지 발생률이 높게 나타나는 결과를 보였다(최소 회귀계수값 음수).

GWR 및 OLS 모형에 따른 추정 심정지 발생률과 실제 관측된 심정지 발생률을 서로 비교하기 위해, 실제 관측된 심정지 발생률과 GWR 및 OLS 모형을 통해 추정된 심정지 발생률의 상관관계를 분석하고 각각의 모형을 통해 도출된 추정 심정지 발생률을 지도 상에 나타내었다.

그 결과, GWR 모형을 통해 추정된 심정지 발생률과 실제 관측된 심정지 발생률과의 상관관계가 OLS 모형을 통한 결과보다 우수하게 나타났으며, GWR로 추정된 심정지 발생률의 분포도 OLS의 결과와 약간의 차이를 보였다(음영의 차이를 통해 확인할 수 있음).

그림 1. GWR 결과에 따른 추정 심정지 발생률 분포

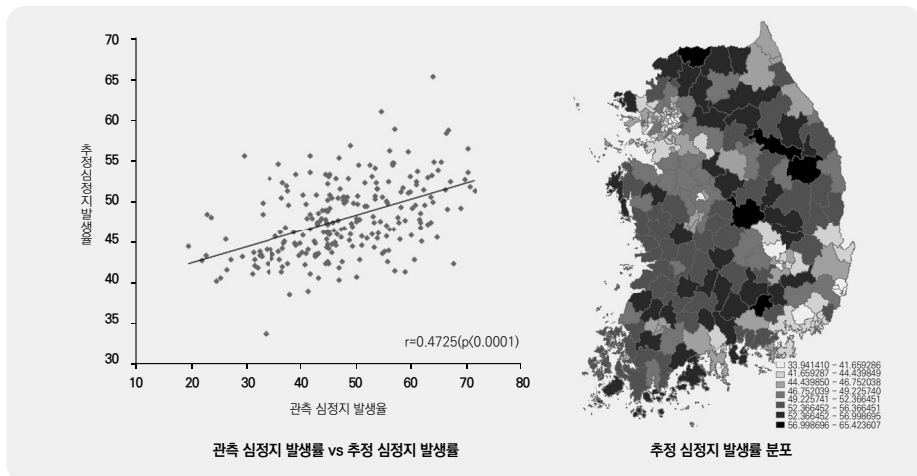
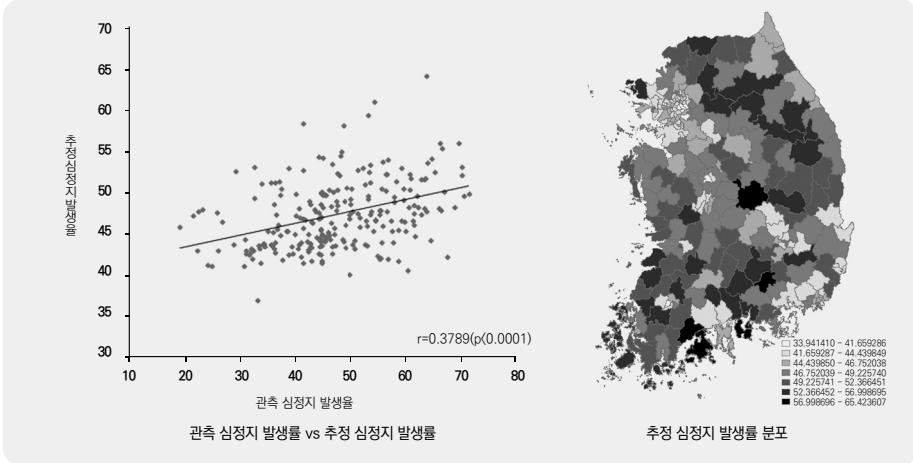


그림 2. OLS 결과에 따른 추정 심정지 발생률 분포



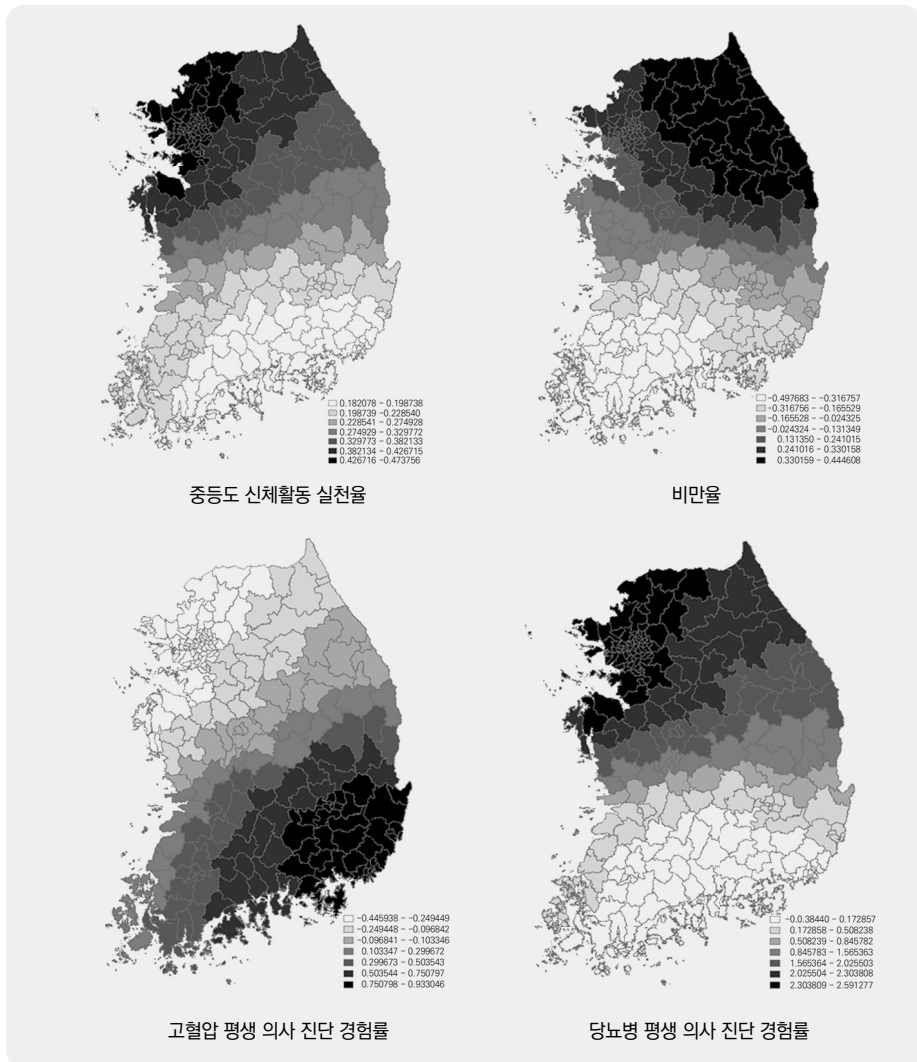
다음의 [그림 3]은 GWR 모형을 통해 산출된 지역별 심정지 발생 위험요인들에 대한 회귀계수를 ArcGIS 10.1을 이용하여 지도에 나타낸 결과물이다. 지역에 해당되는 색농도가 짙을수록 회귀계수 값이 큰 것으로 심정지 발생에 있어 해당 위험요인이 미치는 영향력이 크다는 것을 의미하며 각 지역별로 그 영향력의 차이가 있음을 알 수 있다.

중등도 신체활동 실천율의 경우, 경기도 지역에 근접한 북부지역일수록 격한 운동이 심정지 환자 발생에 큰 영향을 주는 것으로 나타났고, 이와 반대로 고혈압 평생 의사 진단 경험률의 경우 경상도에 근접한 남쪽 지역일수록 그 영향력이 심정지 발생에 큰 것으로 나타났다.

[그림 4]는 GWR 결과에 따른 각 지역별 결정계수(R^2) 및 표준잔차값을 나타내고 있다. 각 지역별 R^2 값은 본 연구에서 제시한 심정지 발생 위험요인인 중등도 신체활동 실천율, 비만율, 고혈압 평생 의사 진단 경험률 및 당뇨병 평생 의사 진단 경험률이 종합적으로 각 지역별 심정지 발생률을 얼마나 잘 설명하는 지를 나타내는 것이다. R^2 값의 크기가 클수록 본 연구에서 제시한 위험요인이 해당 지역의 심정지 발생 위험요인으로 적합함을 의미하는데, 전체적으로 각 지역별 R^2 값이 0.083에서 0.223의 범위를 보이며, 대체로 강원도 및 경북 일부 지역의 경우 해당 독립변수의 설명력이 높은 반면 전남지역에 가까운 남쪽지역일수록 그 설명력이 낮은 것으로 나타났다. 전체적으로 OLS의 결과보다는 GWR 모형이 우수하기는 하나 GWR 모형 자체의 전체적인 설명력

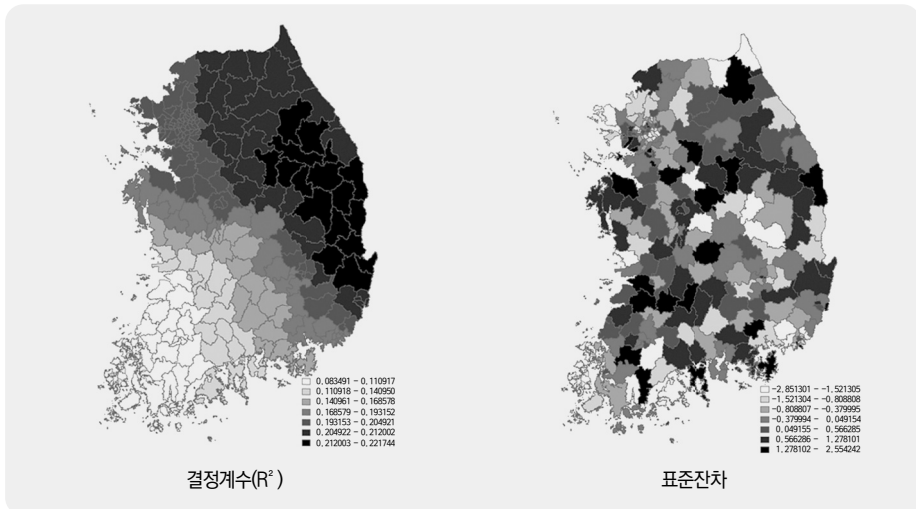
도 다소 낮은 편이며, 각 지역별로도 전남지역에 가까운 남쪽 지역의 경우는 그 설명력이 더욱 더 낮음을 감안할 때, 본 연구에서 제시된 심정지 발생에 영향을 미치는 4가지 요인 이외 추후 연구를 통해 새로운 영향요인을 찾아낼 필요가 있을 것으로 판단된다.

그림 3. GWR 결과에 따른 각 특성별 회귀계수의 지역별 분포



각 지역별 실제 발생률과 발생 위험요인에 따라 추정된 발생률의 차이를 나타내는 표준잔차를 지도화한 것을 살펴보면, 색의 농도가 짙을수록 실제 발생률이 추정된 발생률보다 높은 지역이며, 색의 농도가 옅을수록 실제 발생률보다 추정된 발생률이 더 낮은 지역으로 해석될 수 있다. 또한, 색농도가 중간정도일 경우는 실제 발생률 및 추정된 발생률의 차이가 거의 없는 지역으로 본 연구에서 제시한 모형에 가장 적합한 지역이라 볼 수 있다.

그림 4. GWR 결과에 따른 지역별 결정계수 및 표준잔차 분포



IV. 결론 및 제언

2009년 심정지 발생환자의 퇴원 시 생존율은 3.5% 수준이고 신경학적으로 양호한 생존 퇴원율은 1.0%에 불과하여 우리나라에서 심정지가 발생하면 사망하거나 또는 생존하더라도 정상적인 생활이 어려운 것으로 나타남에 따라 심정지 발생을 예방 관리하는 사업이 매우 필요하다(최정아, 2012). 더불어 심정지 표준화 발생률(2009년 기준)은 인구 10만명당 서울 36.0명, 강원 61.2명, 충북 54.0명, 제주 75.1명으로 지역별로 크게 차이를 보임에 따라 지역사회 특성에 맞는 심정지 사업을 실시할 필요성이 증대되고 있다(신상도 외, 2011).

이를 위해서는 지역별 심정지 발생요인을 규명할 필요가 있으나, 지역주민 개개인의 심정지 위험요인에 대한 자료를 수집하기가 용이하지 않으므로 생태학적 연구방법론을 적용할 필요가 있다. 또한, 회귀분석과 같이 일반적인 통계기법을 적용할 경우 전국 단위의 심정지 발생 위험요인은 규명할 수 있지만 각 지역단위의 심정지 발생 위험요인은 규명할 수 없는 제한점이 있다. 따라서 이러한 제한점을 극복하기 위해 본 연구에서는 GWR 기법을 활용하여 지역사회의 심정지 발생 위험요인을 규명하고자 하였다. GWR의 경우, 관측치들 사이의 거리가 모델에 반영됨으로써 공간상에서 서로 다른 위치에 존재하는 데이터들에 각각 다른 계수를 추정하여 공간적 비정상성(non-stationarity)을 탐색할 수 있다(김익희, 2005; 강승모, 2011).

질병관리본부의 심정지 조사 자료와 지역사회건강조사 자료를 수집하여 240개 지역별로 심정지 발생의 위험요인에 해당되는 파생변수를 만든 후, GWR을 이용하여 심정지 발생 위험요인을 분석하였다. 모형에 활용된 최종 심정지 발생의 주요 위험요인은 중등도 신체활동 실천율, 비만율, 고혈압 평생 의사 진단 경험률 및 당뇨병 평생 의사 진단 경험률이다. 중등도 신체활동 실천율이 높을수록 심정지 발생률이 높게 나타났는데 이는 Dahabreh & Paulus(2011)와 Whang et al.(2006)이 제시한 격렬한 신체활동(RR=4.98) 및 중등도 혹은 고도의 운동을 할 경우(RR=2.38), 심정지 발생이 증가한다는 연구결과와 매우 유사한 결과였다. 또한, 비만율, 고혈압 평생 의사 진단 경험률, 당뇨병 평생 의사 진단 경험률이 높을수록 심정지 발생률이 높게 나타났는데, 이는 낮은 BMI와 같은 저위험도 생활방식은 심정지 위험도를 낮출 수 있다는 Chiuve et al.(2011)의 연구결과와 심정지 발생의 80%는 관상동맥질환과 관련이 있음을 보고한 Myerburg(2001)의 연구결과 및 당뇨병이 심정지의 위험요인일 가능성이 높다는 Jouven et al.(2005)의 연구결과와 일치하는 것으로 나타났다.

중등도 신체활동 실천율의 회귀계수가 높은 지역은 경기, 서울 및 강원 북부지역 일대, 비만율 회귀계수가 높은 지역은 강원 및 경북 일부지역, 고혈압 평생 의사 진단 경험률의 회귀계수가 높은 지역은 대구, 울산, 부산 및 경남 일부지역, 당뇨병 평생 의사 진단 경험률의 회귀계수가 높은 지역은 경기, 서울 및 강원도 일대로 나타났다. 이러한 연구결과를 중심으로 각 심정지 발생 위험요인별 회귀계수가 높은 지역을 대상으로 심정지 예방관리 사업을 실시하는 것이 사업의 효과성을 높일 수 있다고 판단된다. 즉, 올바른 신체활동을 권장하고 관리하는 사업은 경기, 서울 및 강원도 지역, 비만율 관리

사업은 강원 및 경북 일부지역, 고혈압 관리 사업은 대구, 울산, 부산 및 경남 일부지역 그리고 당뇨 관리 사업은 경기, 서울 및 강원도 지역을 우선적으로 실시하는 것이 합리적일 것으로 판단된다.

본 연구에서는 지역사회별로 심정지 관련 예방관리 사업의 우선순위를 정할 수 있는 정보를 제공함으로써 맞춤형 심정지 사업을 수행 할 수 있는 토대를 제공하고자 하였다. 그러나 본 연구에서 심정지 발생의 위험요인으로 소득, 교육수준과 같은 일부 사회경제학적인 요인 등을 고려하지 못한 점과 심정지 발생의 원인을 심인성(질환성 원인), 비심인성(외상, 질식, 익사, 화상, 전기감전 등) 등으로 분류하여 보다 체계적으로 위험요인을 규명하지 못한 제한점이 있음에 따라 이는 추후 연구를 통해 검토되어야 할 필요성이 있다고 본다. 또한, 본 연구에 활용된 GWR 모형이 일반적인 회귀모형에 비해서는 높은 설명력을 가졌으나, 절대적인 설명력이 다소 낮게 나타났다. 이는 GWR 모형 자체에 대한 효용성이라기보다는 본 연구에서 수집된 지역사회 기반자료 및 이를 활용한 생태학적 접근방법론이 가지는 제한점으로 파악된다.

이러한 제한점에도 불구하고 개인별 자료 확보 대비 지역사회 단위 자료 확보의 용이성 및 GWR 모형이 가지는 특성 등을 고려해볼 때, 각 지역별 심정지 발생 예방관리를 위한 우선 사업 순위 선정에 있어 본 연구에서 활용된 연구방법론은 충분히 검토해 볼 가치가 있는 것으로 판단된다.

박일수는 인제대학교에서 데이터정보학 석사 및 보건학 박사학위를 받았으며, 현재 위덕대학교에서 조교수로 재직 중이다. 주요 관심분야는 의료정보, 보건행정이다.

(E-mail: ispark@uu.ac.kr)

김은주는 인제대학교에서 보건행정학 석·박사통합과정 중에 있다. 주요 관심분야는 의료정보, 보건통계이며, 현재 Healthcare Optimization 등을 연구하고 있다.

(E-mail: wish273575@nate.com)

홍성옥은 경원대학교에서 보건학 박사학위를 받았으며, 현재 질병관리본부 만성병조사과에서 재직 중이다. 주요 관심분야는 의료정보, 보건행정이다.

(E-mail: soh822@hanmail.net)

강성홍은 서울대학교에서 보건학 석사, 인제대학교에서 보건학 박사학위를 받았으며, 현재 인제대학교 보건행정학과 교수로 재직 중이다. 주요 관심분야는 의료정보, 보건행정이다.

(E-mail: hcmkang@inje.ac.kr)

참고문헌

- 강승모(2011). 서울시 도시고속도로 사고취약구간 선정에 관한 연구. 서울: 서울시정개발연구원.
- 김두섭, 강남준(2008). 회귀분석: 기초와 응용. 서울: 나남출판사.
- 김익희(2005). GWR을 이용한 지가설명함수의 공간적 이질성 Spatial Heterogeneity 탐색-서울시 서초구·강남구·송파구를 사례로. 석사학위논문, 서울대학교, 서울.
- 신상도, 외 5명(2011). 2010년 심정지 기초 구급 데이터베이스 구축 및 역학지표 산출을 위한 심층 분석. 서울: 서울대학교병원; 질병관리본부 만성질환관리과 연구비 지원.
- 전진호, 외 7명(2012). 심정지 발생 및 생존 관련 요인의 자료 수집 및 관리체계 구축. 김해: 인제대학교; 질병관리본부 연구비 지원.
- 조동기(2009). 지역 단위 조사연구와 공간정보의 활용: 지리정보시스템과 지리적 가중 회귀분석을 중심으로. 한국조사연구학회, 10(3), pp.1-19.
- 최정아(2012). 병원 외 심정지 의무기록조사 주요결과, 2006-2010. 충북 청원군: 질병관리본부 질병예방센터 만성질환관리과.
- Albert, C. M., Manson, J. E., Cook, N. R., Ajani, U. A., Gaziano, J. M., Hennekens, C. H. (1999). Moderate alcohol consumption and the risk of sudden cardiac death among US male physicians. *Circulation*, 100(9), pp.944-950.
- Berry, W. D., Feldman, S. (1985). *Multiple regression in practice*. California: Sage Publications.
- Bobrow, B. J., Spaite, D. W., Berg, R. A., Stolz, U., Sanders, A. B., Kern, K. B., et al. (2010). Chest compression-only CPR by lay rescuers and survival from out-of-hospital cardiac arrest, *The Journal of the American Medical Association*, 304(13), pp.1447-1454.
- Charlton, M., Fotheringham, A. S. (2009). *Geographically Weighted Regression : A Tutorial on using GWR in ArcGIS 9.3*. [Http://ncg.nhim.ie/ncg/GWR/GWR_Tutorial.pdf](http://ncg.nhim.ie/ncg/GWR/GWR_Tutorial.pdf).
- Charlton, M., Fotheringham A. S. (2009). *Geographically Weighted Regression: white paper*. [Http://ncg.nhim.ie/ncg/GWR/GWR_WhitePaper.pdf](http://ncg.nhim.ie/ncg/GWR/GWR_WhitePaper.pdf).

- Chiuve, S. E., Fung, T. T., Rexrode, K. M., Spiegelman, D., Manson, J. E., Stampfer, M. J., et al. (2011). Adherence to a low-risk, healthy lifestyle and risk of sudden cardiac death among women. *The Journal of the American Medical Association*, 306(1), pp.62-69.
- Dahabreh, I. J., Paulus, J. K. (2011). Association of episodic physical and sexual activity with triggering of acute cardiac events: systematic review and meta-analysis. *The Journal of the American Medical Association*, 305(12), pp.1225-1233.
- Durrheim, D. N., Ogunbanjo, G. A. (2004). Making sense of statistics for the family practitioner: What are ecological studies?. *South African Family Practice*, 46(4), p.48.
- Engdahl, J., Holmberg, M., Karlson, B. W., Luepker, R., Herlitz, J. (2002). The epidemiology of out-of-hospital 'sudden' cardiac arrest. *Resuscitation*, 52(3), pp.235-245.
- Fotheringham, A. S., Brunsdon, C., Charlton, M. (2002). *Geographically weighted regression: The analysis of spatially varying relationships*. New York: Wiley.
- Fox, J. (1991). *Regression diagnostics: An Introduction*. California: Sage Publications.
- Goldenberg, I., Jonas, M., Tenenbaum, A., Boyko, V., Matetzky, S., Shotan, A., et al. (2003). Current smoking, smoking cessation, and the risk of sudden cardiac death in patients with coronary artery disease. *Archives of Internal Medicine*, 163(19), pp.2301-2305.
- Hollenberg, J., Herlitz, J., Lindqvist, J., Riva, G., Bohm, K., Rosenqvist, M., et al. (2008). Improved survival after out-of-hospital cardiac arrest is associated with an increase in proportion of emergency crew-witnessed cases and bystander cardiopulmonary resuscitation. *Circulation*, 118(4), pp.389-396.
- Jouven, X., Lemaître, R. N., Rea, T. D., Sotoodehnia, N., Empana, J. P., Siscovick, D. S. (2005). Diabetes, glucose level, and risk of sudden cardiac death. *European Heart Journal*, 26(20), pp.2142-2147.
- Myerburg, R. J. (2001). Sudden cardiac death: exploring the limits of our knowledge.

Journal of Cardiovascular Electrophysiology, 12(3), pp.369-381.

- Ong, M. E., Earnest, A., Shahidah, N., Ng, W. M., Foo, C., Nott, D. J. (2011). Spatial variation and geographic-demographic determinants of out-of-hospital cardiac arrests in the city-state of Singapore. *Annals of Emergency Medicine*, 58(4), pp.343-351.
- Reinier, K., Stecker, E. C., Vickers, C., Gunson, K., Jui, J., Chugh, S. S. (2006). Incidence of sudden cardiac arrest is higher in areas of low socioeconomic status: a prospective two year study in a large United states community. *Resuscitation*, 70(2), pp.186-192.
- Sasson, C., Cudnik, M. T., Nassel, A., Semple, H., Magid, D. J., Sayre, M., et al. (2012). Identifying high-risk geographic areas for cardiac arrest using three methods for cluster analysis. *Academic Emergency Medicine*, 19(2), pp.139-146.
- Selb Semerl, J., Selb, K. (2004). Coffee and alcohol consumption as triggering factors for sudden cardiac death: case-crossover study. *Croatian Medical Journal*, 45(6), pp.775-780.
- Shamseddin, M. K., Parfrey, P. S. (2011). Sudden cardiac death in chronic kidney disease: epidemiology and prevention. *Nature Reviews Nephrology*, 7(3), pp.145-154.
- Strogatz, D. S., Siscovick, D. S., Weiss, N. S., Rennert, G. (1988). Wife's level of education and husband's risk of primary cardiac arrest. *American Journal of Public Health*, 78(11), pp.1491-1493.
- Wannamethee, G., Shaper, A. G. (1992). Alcohol and sudden cardiac death. *British Heart Journal*, 68(5), pp.443-448.
- Whang, W., Manson, J. E., Hu, F. B., Chae, C. U., Rexrode, K. M., Willett, W. C., et al. (2006). Physical exertion, exercise, and sudden cardiac death in women. *The Journal of the American Medical Association*, 295(12), pp.1399-1403.
- Wisten, A., Messner, T. (2005). Young Swedish patients with sudden cardiac death have a lifestyle very similar to a control population. *Scandinavian Cardiovascular Journal*, 39(3), pp.137-142.

A Study on Factors Related with Regional Occurrence of Cardiac Arrest Using Geographically Weighted Regression

Park, Il Su

(Uiduk University)

Hong, Seong Ok

(Korea Centers for
Disease Control & Prevention)

Kim, Eun Ju

(Inje University)

Kang, Sung Hong

(Inje University)

This study sought to use the ecological study design and geographically weighted regression(GWR) to identify regional factors regarding the occurrence of cardiac arrest and to investigate the utility of GWR. Our data came from two main sources: 2009 Out-of-Hospital Cardiac Arrest Surveillance and Community Health Survey. As the result, risk factors on cardiac arrest are moderate-intensity physical activity, obesity rate, diagnosis rates of hypertension and diabetes mellitus which generally increase incidences of cardiac arrest. Our finding showed that GWR increased goodness-of-fit of model in comparison with traditional OLS regression. Considering the usefulness of ecological studies and GWR, methods of this study will help policymakers in terms of planning health programs and identifying priority regions for prevention of cardiac arrest.

Keywords: Incidence Rate of Cardiac Arrest, Ecological Study, OLS, GWR