

신용정보에 따른 입원 및 수술 발생 상대위험도 적용방안 연구: 신용정보원 데이터 이용

A Study on How to Apply the Relative Risk of Hospitalization and Surgery According to Credit Information: Using Korea Credit Information Services Data

전 희 주* (Chun, Heuiju)

인 태 교** (Leen, Taekyo)

황 용 순*** (Hwang, Yongsoon)

〈 목 차 〉

- I. 서 론
- II. 이론적 배경 및 연구방법
- III. 실증분석
- IV. 결론 및 시사점

〈국문초록〉

본 연구는 신용정보원의 신용정보와 보험정보를 이용하여 기존에 개발된 보험상품의 건강상태에 따른 질병발생 상대위험도에 보험가입고객의 신용정보를 추가하는 방법을 제안하고자 한다. 보험계약의 유지자를 대상으로 독립적으로 개인별 신용상태(연체, 대출, 신용 등)에 영향을 받는 입원 및 수술 발생 등 2개의 통계적 예측모형을 개발하여 개별 집단의 평균 보험사고 발생률과 전체 집단의 보험사고

* 제1저자, 동덕여자대학교 정보통계학과 부교수, hjchun@dongduk.ac.kr

** 교신저자, (주)코리안리 상품개발팀 팀장, tkleen@koreanre.co.kr

*** 공동저자, (주)코리안리 상품개발팀 파트장, hwangys@koreanre.co.kr

투고일: 2023.05.30. 심사시작일: 2023.06.01. 최종수정일: 2023.07.11. 게재확정일: 2023.07.24.

발생률의 차이로 정의되는 상대위험도를 계산하였다. 데이터 활용의 제한으로 인하여 기존의 건강상태에 따른 입원/수술 발생과 새롭게 산출하는 신용상태에 따른 입원/수술 발생이 독립이라는 조건하에, 신용정보에 따른 보험사고 발생 상대위험도를 건강상태에 따른 보험사고 발생 상대위험도에 곱함으로써 신용정보를 보험사고 상대위험도에 포함하고자 하였다. 보험사들은 독립적으로 신용상태에 따른 보험사고 발생 상대위험도와 건강상태에 따른 보험사고 발생 상대위험도를 보험상품 가격의 할인 등의 긍정적인 면으로 이용할 수 있을 것이다.

국문색인어 : 로지스틱회귀모형, 신용, 신용정보원 맞춤형DB, 상대위험도

〈ABSTRACT〉

This study proposes a method to add the credit information of insurance customers to the relative risk of disease occurrence according to health status of previously developed insurance products using credit information and insurance information of credit information sources. We developed two statistical prediction models for the occurrence of hospitalization and surgery that are independently affected by individual credit status (delinquency, loan, credit, etc.) targeting the holder of the insurance contract, and calculated the average insurance accident incidence rate of the individual group and the insurance accident rate of the entire group relative risk, defined as the difference in incidence. Due to limitations in data utilization, the relative risk of insurance accidents according to credit information is calculated under the condition that hospitalization/surgery according to existing health conditions and hospitalization/surgery according to newly calculated credit status are independent. By multiplying the relative risk of occurrence, credit information intended to be included in the relative risk of insurance accidents. Insurance companies will be able to independently use the relative risk of an insured event according to credit status and the relative risk of an insured event according to health status as a positive aspect, such as discounting insurance product prices.

Key Words : Credit, Logistic regression model, Korea credit information service customized DB, Relative risk

I. 서론

최근 보험업계는 피보험자의 건강, 행동습관, 신용 등의 개별 정보를 보험 소비자의 개별 보험료 결정에 활용하고 있다. A생명보험사는 국민건강보험공단에 등록된 건강검진 결과를 이용하여 피보험자의 건강나이를 산출하여 보험료에 적용하고 있으며, B생명보험사는 보험소비자가 회사 자체 앱을 사용하도록 하여 소비자의 건강관리 활동 결과를 등록하고 건강관리 등급이 책정되면 보험료 할인을 제공하고 있다.

국내에서 보험계약의 승인심사 또는 보험요율 산정 시에 보험사가 개인 신용정보를 활용한 보험신용평점을 사용하는 면에서는 긍정과 부정의 논란이 있다. 긍정적인 측면에서는 보험신용평점이 개인이 가지고 있는 위험을 보다 정확히 예측할 수 있고 역선택을 줄여 결국 개별 보험소비자들에게 혜택을 줄 수 있다는 반면에 부정적인 측면에서는 개인의 신용이 보험 위험에 영향을 주지 않으며 오히려 저소득층, 사회적 약자에 불리할 수 있고 심지어 특수한 개인들이 부담하는 보험료를 올릴 수 있다고 주장한다(정재욱·여운, 2011).

해외에서는 개인신용정보의 습득과 이용이 훨씬 자유로워 건강정보 외에도 보험소비자의 신용정보를 활용하여 가입심사 뿐 아니라 상품개발에도 반영하는 추세이다(Wu and Guszcz, 2003). 미국의 많은 보험사들은 주택보험·자동차보험에서 가입심사나 보험요율 산정 시 신용정보를 다른 정보와 함께 이용하고 있다. 리스크 솔루션 컨설팅 기업 FICO에 따르면 미국의 개인보험 보험회사의 95%가 신용기반 보험보조지표를 사용하고 있다. 또한 세계적인 생명보험 재보험사인 RGA와 미국의 신용평가기관인 TransUnion은 일반신용정보와 사망률이 밀접한 상관관계가 있고, 일반신용정보를 보험료 산정에 활용하는 것이 허용된다는 점을 이용하여 사망률 추정을 고도화하는 프로젝트도 진행하였다.

한편 국내 보험사의 경우 신용정보원에서 제공받는 보험신용정보를 가입심사와 보험금 지급심사에 활용하고 있으나, 직접적인 상품개발 과정(보험료 및 책임준비금 산출 등)에는 사용하지 않고 있으며, 대출정보·연체정보 등 일반

신용정보의 이용은 쉽지 않은 상황이다.

신용정보원은 대출정보·연체정보 등 일반신용정보를 보험신용정보와 구분하여 수집·보관하고 있으나 현재 보험회사는 일반신용정보와 보험신용정보를 서로 연계하여 활용하지는 않고 있다. 특히, 일반신용정보를 활용한 보험요율 산출에는 일반신용정보와 보험사고 발생 등이 결합된 통계자료가 필요하지만 이에 대한 객관성 있는 자료를 확보하기 어려운 상황이었다. 그러나 2020년 8월 신용정보법 개정에 따라 가명정보의 활용(이용, 제공, 결합 등)이 가능해져 일반신용정보와 보험신용정보를 결합하여 분석할 수 있는 환경이 마련되었다.

이에 본 연구는 미국의 일반신용정보 활용사례를 참고삼아 신용정보원에서 제공하는 일반신용정보와 보험신용정보가 연계된 DB를 이용하여 개인의 신용상태에 따른 보험사고 발생을 예측하는 모형에 의한 상대위험도를 구하여 기존의 건강상태에 따른 보험사고 발생모형과 결합시키는 보험상품 개발 적용방법을 논의하고자 한다.

Ⅱ. 이론적 배경 및 연구방법

1. 이론적 배경

최근 개인의 과거 사회·경제적 행동에 대한 기록을 나타내는 신용(credit) 속성들이 건강과 관련된 질병의 발생 결과와 밀접하게 관련이 있다는 많은 연구 결과들이 나오고 있다(Dean et al., 2019; Dean et al., 2018; Dobkin et al., 2018; Houle et al., 2015). 이러한 연구들은 경제적 상태와 건강 간에 밀접한 관련이 있다는 증거를 제시한다.

개인들의 과거 사회·경제적 행동에 대한 기록을 하나의 점수 또는 평점으로 나타내는 신용평점(credit score)은 개인의 속성, 개인의 위험 행태들, 지불 능력, 파산 위험을 반영한다. 고객의 신용평점은 신용의 긍정적 또는 부정적 상태와 부채를 감당할 능력과 함께 개인의 과거 경제 행위에 대한 요약된

측정을 제공한다. 이러한 신용평점은 건강 상태를 나타내는 질병의 결과들과 밀접한 관련이 있다는 연구들도 계속 제시되고 있다(Dean and Nicholas, 2018; Israel et al., 2014; Li, 2015; Knapp and Dean, 2018).

Dean and Nicholas(2018)은 경제적 상태와 건강상태는 밀접한 관련이 있기 때문에 사회·경제적 요인들뿐만 아니라 건강 또는 질병에 영향을 주는 요인들인 건강의 상태 변수들도 신용평점모형에 포함되어야 한다고 주장한다(Dean and Nicholas, 2018). 또한 이들은 개인의 경제적 상태와 정보를 표현할 수 있는 언어 및 지적인 상태 또한 신용평점과 관계가 있기 때문에 심지어 사회·경제적 요인들은 개인의 심리적 속성과 건강의 상태, 질병의 결과를 예측할 수 있다고 본다(Dean and Nicholas, 2018). Dean and Nicholas(2018)는 하나의 예로 갑작스러운 입원은 실직으로 인한 낮은 신용한도와 낮은 신용평점으로부터 초래될 수 있으며, 이러한 이유로 고객 신용평점이 건강과 질병을 예측하기 위한 측정 수단으로 사용될 수 있음을 보여 주었다.

개인의 신용은 정보의 비대칭을 가져올 수 있는 역선택을 방지하기 위한 보험의 언더라이팅 심사에 사용될 수 있다. 신용기반 보험평점(credit-based insurance score)은 보험고객의 과거의 신용 기록을 이용하여 미래 보험금 신청과 보험청구 비용을 예측하기 위해 사용된다. 이러한 신용기반 보험평점은 일반적으로 개인의 과거 연체 또는 파산 경험, 신용한도 소진율, 신규 신용개설을 위한 문의건수 및 신설 계좌수, 신용기록, 기타 신용거래내역 등을 포함하는 개인의 신용과 관련된 자료를 사용하여 계산된다. 신용평점(credit score)과 신용기반 보험평점 모두 위험(risk)을 예측하는 공통점이 있지만, 신용평점은 부채를 갚을 능력을 예측하는 데 있는 반면에 신용기반 보험평점은 미래 보험금 청구에 대한 가능성을 예측한다. 신용기반 보험평점은 또한 보험금 손실 예측을 위해 사용되며(Golden et al., 2016), 보험사기 방지를 위해 사용된다(Roberts, 2011).

국내에서 신용과 보험사고율과의 관계에 관한 연구로는 정재욱·손승호(2009)와 정재욱·여유희(2011)가 있다. 정재욱·손승호(2009)는 국내에서 신용과 보험사고율에 관한 최초 논문으로 국내 한 손해보험사의 총 자동차보험 계약건수 297,674건수 데이터를 기반으로 보험회사의 개인 보험정보와 KCB

(개인신용평가 전문회사)의 신용정보를 결합하여 자동차보험 사고율 예측모형을 개발하여 개인의 신용등급이 개인 자동차 보험사고와 연관성이 있다는 것을 실증적으로 보여 주었다. 정재욱·손승호(2009)가 이용한 신용 정보는 개설통보, 연체 및 불량정보, 대출실적정보, 카드실적정보, 신상정보이다. 이들은 자동차보험의 경우 개인의 신용등급(1~10등급)은 사고빈도 및 손해율과 매우 밀접한 상관관계가 있음을 제시하였고, 다변량로지스틱 회귀모형의 비교 분석을 통해 신용정보를 추가한 모형이 순수 보험정보만을 고려한 모형보다 모형의 예측력이 더 우수함을 보여주었다. 정재욱·여유희(2011)는 국내의 한 생명보험사의 2007년 1~7월 특정 건강보장상품과 상해보장상품을 가입한 보험계약 가입자 109,689명을 대상으로 보험사 고유정보, KFB(전국은행연합회) 정보, KCB 정보를 활용하여 이들의 보험사고 예측모형을 개발하여 개인신용등급 수준이 낮을수록 보험사고율이 높아진다는 사실을 실증적으로 제시하였다. 이들은 보험사의 고유정보인 인구통계정보와 계약정보, KFB의 대출정보, 보증정보, 개설통보, 연체정보와 KCB의 대출정보, 보증정보, 카드 사용정보, 연체정보를 사용하여 순수 보험사정보를 사용한 모형보다 KFB정보, KCB정보를 모두 포함한 모형이 예측효과가 높음을 보였다.

그러나 정재욱·손승호(2009)의 연구는 국내 특정 한 손해보험사의 총 297,674건 자동차보험 계약건수를 가지고 신용과 자동차보험 사고율 관계로 제한되어 있고, 정재욱·여유희(2011)의 연구는 국내 한 생명보험사의 특정 보험상품의 보험계약 가입자 109,689명에 한정하여 신용과 특정상품 보험사고의 예측모형을 제시하였다. 그러나 본 연구는 모든 보험사의 보험가입고객 중에서 표본 추출된 신용정보원의 대용량 데이터를 이용하여 성별과 연령의 효과를 포함하지 않고 오로지 신용, 연체, 대출 관련 정보를 사용하여 보험사고 담보인 입원과 수술의 발생률 모형을 구한 후 기존의 입원, 수술 등의 보험사고 상품에 결합하는 방법을 제시하고자 한다. 이것은 기존의 입원, 수술 등의 보험사고 발생률에 신용관련 변수들인 신용, 연체, 대출을 결합하여 보험사고 위험을 보다 정확히 예측하고 보험상품의 역선택과 보험가입고객들의 모럴해저드 문제 해결에 도움을 주고자 한다.

2. 연구데이터

본 연구에서 사용되는 데이터는 신용정보원에서 제공하는 맞춤형DB를 사용하였다. 맞춤형DB는 신용정보원이 집중·관리하는 모든 신용정보를 융합하여 효율성을 높인 데이터셋이다. 맞춤형DB는 신용정보원에 등록된 개인 및 법인의 신용정보 중 일부가 표본으로 추출된 후 개인을 재식별 할 수 없도록 비식별 조치를 수행 및 가공되어 개인-기업-보험신용정보 형태로 연계된 DB이다.

〈표 1〉 신용정보원 맞춤형DB의 테이블 및 데이터 설명

구분	테이블	데이터설명
일반 신용 정보	연체정보 (신용도판단정보)	연체관련 정보 중 금융회사가 등록한 정보: 연체사유, 연체등록월, 연체등록금액, 연체해제월 등
	연체정보 (공공정보)	연체관련 정보 중 법원 등에서 등록한 정보: 연체사유, 연체등록월, 연체등록금액, 연체해제월 등
	개인대출정보 (금융업권)	대출정보 중 금융업권에서 등록한 정보: 대출상품코드, 대출월, 대출금액, 만기월 등
	신용카드개설정보	신용카드 개설 관련 정보: 카드유형코드, 카드개설월 등
	신용카드이용정보	신용카드 한도 및 이용실적 정보: 신용판매한도금액, 현금서비스한도금액, 신용판매이용금액 등
보험 신용 정보	계약관계자정보	계약자 및 피보험자 관련 정보: 계약관계자유형코드, 계약자-피보험자 관계코드 등
	보험계약정보	보험계약 관련 기본 정보: 보험종류코드, 보험계약체결월, 보험계약해지월 등
	보험담보정보	세부 담보 관련 정보: 보험담보그룹코드, 보험담보코드, 보험세부담보금액 등
	청구사고정보	보험청구가 발생한 사고 관련 정보: 사고발생월, 최초사고접수월 등
	청구계약정보	보험청구에 관련된 보험계약 정보: 보험종류코드, 보험계약체결월, 보험시기월 등
	청구지급사유정보	보험청구에 관련된 보험금지급결과 정보: 보험종류코드, 지급사유코드, 지급월 등
	청구지급상세 사유정보	보험청구에 관련된 보험금지급금액 정보: 지급사유코드, 지급월, 상세지급사유코드, 지급금액 등

〈표 1〉에서 보는 바와 같이 일반신용정보는 차주정보, 대출정보, 연체정보

및 카드개설정보 등을 포함하고 있으며, 보험신용정보는 계약관계자정보, 계약정보, 담보정보, 청구정보 등을 포함하고 있어 이를 연계하여 데이터를 사용하였다.

본 연구에서 사용된 맞춤형DB의 데이터 관찰기간은 2018년 6월~2020년 5월로 분석 기준 대상자는 2019년 6월 보험계약을 유지중인 대상자 중에서 만 20세~74세인 1,678,478명을 대상으로 하였다.

3. 연구설계 및 통계적 방법론

1) 연구설계

본 연구에서는 로지스틱회귀모형을 이용하여 개인의 신용상태에 따른 보험사고 즉, 입원발생, 수술발생 등 2개의 발생률을 각각 예측하는 모형을 구하고자 한다.

먼저, 분석대상자는 2019년 6월 보험계약 유지자(만 20~74세)로 정하였고, 이들에 대한 전체 관찰기간(2018년 6월~2020년 5월)을 사전 관찰기관(2018년 6월~2019년 5월)과 사후 관찰기관(2019년 6월~2020년 5월)으로 나누었다. 사전 관찰기간을 통해 분석대상자의 직전 1년 신용상태를 나타내는 속성들을 설명변수로 정의하였고 사후 관찰기간을 통해 분석대상자의 입원과 수술의 보험사고 발생여부를 반응변수로 정의하였다. 반응변수의 정의는 맞춤형DB의 보험사고 청구정보를 이용하여 <표 2>와 같이 정의하였다.

〈표 2〉 보험사고(반응변수) 정의

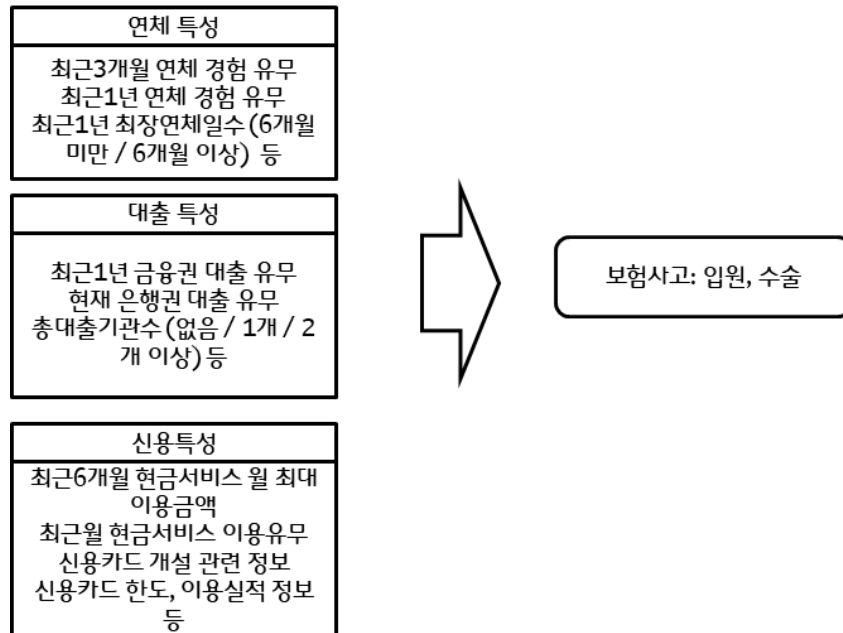
보험사고(반응변수)	지급사유코드
입원	질병입원(정액)
수술	질병수술(정액)

사전 관찰기관 동안의 신용상태 즉, 설명변수의 정의는 맞춤형DB의 연체(overdue), 대출(loan), 신용(credit) 관련 정보 등을 이용하여 <표 3>과 같이 정의하였다.

〈표 3〉 신용상태 주요변수 정의

구분	신용상태(설명변수)
연체	최근 3개월 연체 경험 유무
	최근 1년 연체 경험 유무
	최근 1년 최장연체일수 (6개월 미만, 6개월 이상)
대출	최근 1년 금융권 대출 유무
	현재 은행권 대출 유무
	총 대출기관수 (없음, 1개, 2개 이상)
신용	최근 6개월 현금서비스 월 최대이용금액 (없음, 50만원 이하, 50만원 초과)
	최근 월 현금서비스 이용유무

〈그림 1〉은 본 연구의 연구모형을 도식화한 것으로 연체, 대출, 신용 관련 특성들을 이용하여 반응변수인 보험사고 즉, 입원 및 수술 발생을 예측하는 모형을 찾고, 신용상태에 따른 입원 및 수술 발생 위험도를 객관적으로 측정하고자 한다. 기존의 건강상태에 따른 보험사고 발생 모형은 성별, 연령별 변



〈그림 1〉 연구모형

수를 포함하지만 새롭게 산출하는 신용상태에 따른 보험사고 발생모형은 성별, 연령별 중복을 피하고자 성별, 연령 변수는 제외하고 순수한 신용변수만을 포함한 모형으로 제한하였다.

2) 통계적 방법론

예측모형은 보험사고(입원, 수술) 발생 유무를 이항 반응변수로 하는 다중 로지스틱회귀모형을 이용하여 보험사고 발생에 유의한 영향을 주는 요인들을 찾고 그 요인들의 영향정도를 파악하였다(이상연 외, 2023).

$$\ln\left(\frac{\hat{p}_i}{1-\hat{p}_i}\right) = \hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \hat{\beta}_2 x_{i2} + \cdots + \hat{\beta}_k x_{ik} \quad (1)$$

식 (1)을 이용하여 개별가입자(i)의 보험사고 발생률을 의미하는 $\hat{p}_i$ 는 식 (2)와 같이 도출된다.

$$\hat{p}_i = \frac{\exp(\hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \hat{\beta}_2 x_{i2} + \cdots + \hat{\beta}_k x_{ik})}{1 + \exp(\hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \hat{\beta}_2 x_{i2} + \cdots + \hat{\beta}_k x_{ik})} \quad (2)$$

여기서 $x_{i1}, x_{i2}, \dots, x_{ik}$ 는 가입자별 입원과 수술 보험사고 발생에 영향을 줄 것으로 예상되는 <표 3>의 연체, 대출 및 신용 관련 설명변수들이다.

Ⅲ. 실증분석

1. 관찰대상 집단의 일반적인 특성

<표 4>는 분석대상자의 성별, 연령별 분포를 나타낸 표이다. 성별 남녀 비중은 각각 50.2%, 49.8%로 거의 동일하며, 연령대별 비중을 살펴보면 50~59세(21.9, 22.6%), 40~49세(21.9, 21.8%), 30~39세(19.5, 19.0%), 20~29세(18.3, 16.4%) 등으로 순위가 나타나 60세 미만의 연령대가 집단의 81.5,

79.8%를 차지하고 있다.

〈표 4〉 분석대상자 성별, 연령별 분포

연령	남성		여성	
	수	비중(%)	수	비중(%)
20~29	153,744	18.3	137,300	16.4
30~39	163,867	19.5	158,573	19.0
40~49	184,381	21.9	182,433	21.8
50~59	184,374	21.9	189,315	22.6
60~69	124,767	14.8	133,606	16.0
70~74	30,950	3.7	35,168	4.2
계	842,083	50.2	836,395	49.8

〈표 5〉는 분석대상자의 입원 및 수술 발생 분포를 나타낸 표이다. 입원 발생 비중은 3.5%, 수술 발생 비중은 2.6%를 차지하고 있다.

〈표 5〉 분석대상자 입원 및 수술 발생 분포

분석대상자 수	입원		수술	
	발생자 수	비중(%)	발생자 수	비중(%)
1,678,478	59,190	3.5	43,122	2.6

1) 신용관련 변수들에 따른 입원 발생의 관계

〈표 6〉은 최근 3개월 연체경험 유무, 최근 1년 연체경험 유무, 최근 1년 최장연체일수, 최근 1년 금융권 대출 유무, 현재 은행권 대출 유무, 총 대출기관 수, 최근 월 현금서비스 이용유무, 최근 6개월 현금서비스 월 최대이용금액 등의 신용관련 변수들에 따른 입원 발생의 관계를 나타낸 표이다. 여기서 발생자 수는 입원 횟수와 상관없이 사후 관찰기간 동안 1회 이상 입원이 발생하면 1인으로 적용하였다. 최근 1년 연체경험이 있는 집단은 없는 집단 대비 입원 발생 상대비는 1.118배 높게 나타났다. 최근 1년 최장연체일수는 6개월 미만 대비 6개월 이상인 경우가 1.126배로 최장 연체일수가 많은 경우가 입원비

율이 더 높음을 보여준다.

최근 1년 금융권 대출 유무를 보면, 대출이 있는 경우는 없는 경우 대비 입원 발생 상대비는 1.161배 높으며, 총 대출기관수는 대출기관이 많을수록 입원 발생 상대비는 높은 경향을 보이고 있다. 최근 월 현금서비스 이용은 현금 서비스를 받고 있는 경우가 그렇지 않은 경우 대비 입원 발생 상대비는 1.108배 높았다. 최근 6개월 현금서비스 월 최대이용금액은 없는 경우 대비 50만원 초과인 경우는 1.127배 높게 나타났다.

〈표 6〉 신용 상태와 입원 발생과의 관계

구분		대상자수	비율(%)	발생자수	발생률(%)	상대비
최근 3개월 연체 경험	없음	1,602,136	95.5	56,247	3.511	-
	있음	76,342	4.6	2,943	3.855	1.098
최근 1년 연체 경험	없음	1,588,810	94.7	55,676	3.504	-
	있음	89,668	5.3	3,514	3.919	1.118
최근 1년 최장연체일수	6개월 미만	1,594,576	95.0	55,879	3.504	-
	6개월 이상	83,902	5.0	3,311	3.946	1.126
최근 1년 금융권대출	없음	760,659	45.3	24,651	3.241	-
	있음	917,819	54.7	34,539	3.763	1.161
현재 은행권 대출	없음	1,183,917	70.5	41,576	3.512	-
	있음	494,561	29.5	17,614	3.562	1.014
총 대출기관수	없음	837,678	49.9	27,590	3.294	-
	1개	434,361	25.9	15,983	3.680	1.117
	2개 이상	406,439	24.2	15,617	3.842	1.167
최근 월 현금서비스	없음	1,595,910	95.1	55,980	3.508	-
	있음	82,568	4.9	3,210	3.888	1.108
최근 6개월 현금서비스 월 최대이용금액	없음	1,525,158	90.9	53,242	3.491	-
	50만 이하	45,707	2.7	1,716	3.754	1.075
	50만 초과	107,613	6.4	4,232	3.933	1.127

2) 신용관련 변수들에 따른 수술 발생의 관계

〈표 7〉은 최근 3개월 연체경험 유무, 최근 1년 연체경험 유무, 최근 1년 최장연체일수, 최근 1년 금융권 대출 유무, 현재 은행권 대출 유무, 총 대출기관

수, 최근 월 현금서비스 이용유무, 최근 6개월 현금서비스 월 최대이용금액 등의 신용관련 변수들에 따른 수술 발생의 관계를 나타낸 표이다. 여기서 발생자 수는 수술 횟수와 상관없이 사후 관찰기간 동안 1회 이상 수술이 발생하면 1인으로 적용하였다. 최근 3개월 연체경험과 보험사고와의 관계를 보면, 최근 3개월 연체경험이 있는 집단은 연체경험이 없는 집단 대비 수술 발생 상대비는 1.241배 높게 나타났으며, 최근 1년 연체경험이 있는 집단은 없는 집단 대비 수술 발생 상대비는 1.254배 높게 나타났다. 최근 1년 최장연체일수는 6개월 미만 대비 6개월 이상인 경우가 1.287배로 최장 연체일수가 많은 경우가 수술비율이 더 높음을 보여준다.

〈표 7〉 신용 상태와 수술 발생과의 관계

구분		대상자수	비율(%)	발생자수	발생률(%)	상대비
최근 3개월 연체 경험	없음	1,602,136	95.5	40,715	2.541	-
	있음	76,342	4.6	2,407	3.153	1.241
최근 1년 연체 경험	없음	1,588,810	94.7	40,271	2.535	-
	있음	89,668	5.3	2,851	3.180	1.254
최근 1년 최장연체일수	6개월 미만	1,594,576	95.0	40,388	2.533	-
	6개월 이상	83,902	5.0	2,734	3.259	1.287
최근 1년 금융권대출	없음	760,659	45.3	17,646	2.320	-
	있음	917,819	54.7	25,476	2.776	1.197
현재 은행권 대출	없음	1,183,917	70.5	30,015	2.535	
	있음	494,561	29.5	13,107	2.650	1.045
총 대출기관수	없음	837,678	49.9	19,770	2.360	-
	1개	434,361	25.9	11,678	2.689	1.139
	2개 이상	406,439	24.2	11,674	2.872	1.217
최근 월 현금서비스	없음	1,595,910	95.1	40,635	2.546	-
	있음	82,568	4.9	2,487	3.012	1.183
최근 6개월 현금서비스 월 최대이용금액	없음	1,525,158	90.9	38,541	2.527	-
	50만 이하	45,707	2.7	1,264	2.765	1.094
	50만 초과	107,613	6.4	3,317	3.082	1.220

최근 1년 금융권 대출 유무를 보면, 대출이 있는 경우는 없는 경우 대비 수

술 발생 상대비는 1.197배 높으며, 총 대출기관수는 대출기관이 많을수록 수술 발생 상대비는 높은 경향을 보이고 있다. 최근 월 현금서비스 이용은 현금 서비스를 받고 있는 경우가 그렇지 않은 경우 대비 수술 발생 상대비는 1.183배 높았다. 최근 6개월 현금서비스 월 최대이용금액은 없는 경우 대비 50만원 미만의 수술 발생 상대비는 1.094배, 50만원 초과인 경우는 1.220배 높게 나타났다.

2. 통계적 모형

보험가입자의 신용상태가 입원, 수술 발생률에 영향을 주는지를 확인하기 위해 최근 3개월 연체경험 유무, 최근 1년 연체경험 유무, 최근 1년 최장연체일수, 최근 1년 금융권 대출 유무, 현재 은행권 대출 유무, 총 대출기관수, 최근월 현금서비스 이용유무, 최근 6개월 현금서비스 월 최대이용금액 등의 신용관련 변수들을 설명변수로 하고 입원, 수술 발생률을 각각 예측하는 최적의 로지스틱회귀모형을 찾고자 한다.

위 설명변수들 중 보험사고 발생에 영향을 주는 요인들을 찾기 위해 다음과 같은 절차를 수행하였다. 먼저 입원, 수술의 반응변수들에 대해 신용상태 등을 나타내는 각 설명변수로 하는 단순로지스틱회귀모형을 적용하였다. 그리고 설명변수를 대상으로 후진제거법(backward elimination method)을 통해 모형 선택기준인 AIC가 가장 작은 설명변수들의 조합으로 최적의 회귀모형을 선택하였다. 모형적합을 확인하기 위해서 이탈도(deviance) 통계량의 p -값 >0.05 인지를 점검하였고 예측치와 실제 관측치의 차이를 이용한 모형 적합도 통계량인 Hosmer & Lemeshow 카이제곱 통계량이 p -값 >0.05 인지를 추가적으로 확인하였다.

1) 입원발생 모형

로지스틱회귀모형을 이용하여 오직 신용관련 변수들에 의한 입원발생 유무를 적합한 결과, <표 8>은 AIC가 가장 작은 최적모형이 포함한 변수들은 은행 대출 유무, 최근 1년 금융권대출, 총 대출기관수, 최근 1년 최장연체일수, 최

근 3개월 연체 등 중요도 순으로 나타났다.

〈표 8〉 입원 발생에 유의미한 영향을 미치는 변수

설명변수	df	Wald- χ^2	p-value
현재 은행권 대출 유무	1	140.5	<.0001
최근 1년 금융권 대출 유무	1	75.8	<.0001
총 대출기관수	2	43.5	<.0001
최근 1년 최장연체일수	1	20.9	<.0001
최근 3개월 연체 경험 유무	1	6.8	0.0092

〈표 9〉의 모형적합 결과를 살펴보면, 입원발생의 오즈(odds)는 은행대출 경우 ‘없음’ 대비 ‘있음’의 경우 0.869(=exp(-0.1406))배, 최근 1년 금융권대출의 경우 ‘없음’ 대비 ‘있음’의 경우 1.178(=exp(0.1637))배, 총 대출기관수의 경우 ‘없음’ 대비 ‘2개 이상’의 경우 1.112(=exp(0.1059))배 높게 나타나는 경향을

〈표 9〉 입원 발생 로지스틱회귀모형

설명변수	df	$\hat{\beta}$	s.e.	p-value
상수	1	-3.4012	0.0065	<.0001
현재 은행권 대출	있음	-0.1406	0.0118	0.0001
	없음	0	0.0000	-
최근 1년 금융권 대출	있음	0.1637	0.0185	<.0001
	없음	0	0.0000	-
총 대출기관 수	1개	0.0393	0.0198	0.0475
	2개 이상	0.1059	0.0206	<.0001
	없음	0	0.0000	-
최근 1년 최장연체일수	6개월 이상	0.1749	0.0378	<.0001
	6개월 미만	0	0.0000	-
최근 3개월 연체 경험	있음	-0.1044	0.0400	0.0090
	없음	0	0.0000	-
통계량		df	χ^2	p-value
이탈도(deviance)		1,678,471	511,702	1.0000
Hosmer & Lemeshow 카이제곱		8	5.4697	0.7064
AIC		511,716		

보였다.

〈표 10〉은 〈표 9〉의 보험가입자의 신용상태에 따른 집단별 입원 예측 발생률과 전체 평균 3.5264 대비 상대비 예측 결과를 크기순으로 나타낸 표이다. 보험사는 개인들의 신용상태에 따른 상대위험도를 기존의 건강상태에 따른 상대위험도에 곱함으로써 보험가입자의 신용에 따른 위험을 포함시킬 수 있을 것이다. 그러나 현재 건강보험데이터와 신용정보원의 데이터의 결합이 가능하지 않기에 신용변수들에 의한 입원발생이 건강상태에 따른 발생과 독립이라는 가정을 하였다는 한계점이 존재한다.

〈표 10〉 집단별 입원 예측 발생률 및 상대위험도 산출

최근 3월 연체	최근 1년 최장연체일	최근 1년 금융권대출	은행 대출	총 대출 기관수	대상자	발생률 (%)	상대 위험도
있음	6개월 미만	없음	있음	없음	-	2.5428	0.7211
있음	6개월 미만	없음	있음	1개	-	2.6420	0.7492
없음	6개월 미만	없음	있음	없음	-	2.8148	0.7982
있음	6개월 미만	없음	있음	2개 이상	-	2.8190	0.7994
있음	6개월 미만	없음	없음	없음	825	2.9154	0.8267
없음	6개월 미만	없음	있음	1개	-	2.9242	0.8292
있음	6개월 미만	있음	있음	없음	-	2.9817	0.8455
있음	6개월 이상	없음	있음	없음	-	3.0142	0.8547
있음	6개월 미만	없음	없음	1개	49	3.0287	0.8588
있음	6개월 미만	있음	있음	1개	192	3.0975	0.8784
없음	6개월 미만	없음	있음	2개 이상	-	3.1195	0.8846
있음	6개월 이상	없음	있음	1개	-	3.1311	0.8879
없음	6개월 미만	없음	없음	없음	721,194	3.2259	0.9148
있음	6개월 미만	없음	없음	2개 이상	39	3.2307	0.9161
없음	6개월 미만	있음	있음	없음	-	3.2990	0.9355
있음	6개월 미만	있음	있음	2개 이상	1,699	3.3039	0.9369
없음	6개월 이상	없음	있음	없음	-	3.3348	0.9457
있음	6개월 이상	없음	있음	2개 이상	-	3.3398	0.9471
없음	6개월 미만	없음	없음	1개	2,167	3.3508	0.9502
있음	6개월 미만	있음	없음	없음	174	3.4164	0.9688
없음	6개월 미만	있음	있음	1개	215,867	3.4266	0.9717
있음	6개월 이상	없음	없음	없음	26,925	3.4534	0.9793

최근 3월 연체	최근 1년 최장연체일	최근 1년 금융권대출	은행 대출	총 대출 기관수	대상자	발생률 (%)	상대 위험도
없음	6개월 이상	없음	있음	1개	-	3.4637	0.9822
있음	6개월 이상	있음	있음	없음	-	3.5315	1.0014
있음	6개월 미만	있음	없음	1개	768	3.5484	1.0062
없음	6개월 미만	없음	없음	2개 이상	910	3.5735	1.0134
있음	6개월 이상	없음	없음	1개	1,477	3.5868	1.0171
없음	6개월 미만	있음	있음	2개 이상	261,357	3.6542	1.0362
있음	6개월 이상	있음	있음	1개	2,208	3.6678	1.0401
없음	6개월 이상	없음	있음	2개 이상	-	3.6937	1.0474
없음	6개월 미만	있음	없음	없음	79,293	3.7781	1.0714
있음	6개월 미만	있음	없음	2개 이상	1,814	3.7837	1.0730
없음	6개월 이상	없음	없음	없음	5,226	3.8189	1.0829
있음	6개월 이상	없음	없음	2개 이상	1,374	3.8246	1.0845
없음	6개월 이상	있음	있음	없음	-	3.9049	1.1073
있음	6개월 이상	있음	있음	2개 이상	11,426	3.9107	1.1090
없음	6개월 미만	있음	없음	1개	199,527	3.9235	1.1126
없음	6개월 이상	없음	없음	1개	286	3.9658	1.1246
있음	6개월 이상	있음	없음	없음	2,503	4.0430	1.1465
없음	6개월 이상	있음	있음	1개	621	4.0550	1.1499
없음	6개월 미만	있음	없음	2개 이상	108,701	4.1827	1.1861
있음	6개월 이상	있음	없음	1개	8,993	4.1981	1.1905
없음	6개월 이상	없음	없음	2개 이상	187	4.2276	1.1988
없음	6개월 이상	있음	있음	2개 이상	1,191	4.3224	1.2257
없음	6개월 이상	있음	없음	없음	1,538	4.4680	1.2670
있음	6개월 이상	있음	없음	2개 이상	15,876	4.4746	1.2689
없음	6개월 이상	있음	없음	1개	2,206	4.6386	1.3154
없음	6개월 이상	있음	없음	2개 이상	1,865	4.9426	1.4016
전체					1,678,478	3.5264	-

〈표 10〉의 결과를 바탕으로 평균 3.5264 대비 상대위험도가 1미만, 즉 전체 입원발생률보다 낮은 발생률의 집단을 그룹화하여 재산출하면 〈표 11〉과 같다. 보험사는 이러한 객관적인 입원발생 상대위험도를 적용하여 입원발생 상대위험도가 전체 평균보다 낮은 보험 가입고객들에 대해 보험료 또는 수수료 할인 등의 다양한 혜택을 제공할 수 있다.

〈표 11〉 입원 발생의 상대위험도에 따른 재분류 결과

구분	상대위험도 < 1	상대위험도 ≥ 1
입원 발생률(%)	3.2771	3.8671
상대비	0.9293	1.0966

2) 수술발생 모형

〈표 12〉는 로지스틱회귀모형을 이용하여 수술 발생유무를 적합한 최적모형에 포함된 변수들이다. 포함된 설명변수는 최근 1년 연체 유무, 최근 1년 최장연체일수, 최근 1년 금융권 대출 유무, 최근 월 현금서비스 이용 유무, 총 대출기관수 순으로 나타났다.

〈표 12〉 수술 발생에 유의미한 영향을 미치는 변수

설명변수	df	Wald- χ^2	p-value
최근 1년 연체 경험 유무	1	61.1	<.0001
최근 1년 최장연체일수	1	37.0	<.0001
최근 1년 금융권 대출 유무	1	15.1	<.0001
최근 월 현금서비스 이용 유무	1	10.1	0.0001
총 대출기관수	2	9.8	0.0076

〈표 13〉의 모형적합 결과를 살펴보면, 수술발생의 오즈(odds)는 최근 1년 연체의 경우 '없음' 집단 대비 '있음' 집단의 경우 0.753(=exp(-0.2838))배, 최근 1년 최장연체일수의 경우 '6개월 미만' 대비 '6개월 이상'의 경우 1.709(=exp(0.5362))배, 최근1년 금융권 대출의 경우 대출을 받지 않는 경우 대비 대출을 받은 경우 1.1879(=exp(0.1722))배, 현금서비스 이용의 경우 '없음' 대비 '있음'의 경우 1.092(=exp(0.0880))배 높게 나타나는 경향을 보였다. 총 대출기관수를 보면 총 대출기관수가 기준범주인 "없음" 대비 총 대출기관수가 "1개"와 "2개" 각 범주와의 차이 유의성을 나타내지 않지만 총 대출기관수 "1개" 대비 총 대출기관수가 "2개"인 오즈는 1.044(=exp(0.0433))배임을 알 수 있다.

〈표 13〉 수술 발생 로지스틱회귀모형

설명변수		df	$\hat{\beta}$	s.e.	p-value
상수		1	-3.7532	0.0077	<.0001
최근 1년 연체 경험	있음	1	-0.2838	0.0937	<.0001
	없음	0	0.0000	0.0000	-
최근 1년 최장연체일수	6개월 이상	1	0.5362	0.0955	<.0001
	6개월 미만	0	0.0000	0.0000	-
최근 1년 금융권 대출	있음	1	0.1722	0.0216	<.0001
	없음	0	0.0000	0.0000	-
최근 월 현금서비스 이용	있음	1	0.0880	0.0224	0.0001
	없음	0	0.0000	0.0000	-
총 대출기관수	1개	1	-0.0180	0.0223	0.4198
	2개 이상	1	0.0253	0.0228	0.2680
	없음	0	0.0000	0.0000	-
통계량			df	χ^2	p-value
이탈도(deviance)			1,678,471	400,378	1.0000
Hosmer & Lemeshow 카이제곱			8	10.723	0.2179
AIC			400,392		

〈표 14〉는 〈표 13〉의 보험가입자의 신용상태 변수들에 의한 집단별 수술 예측 발생률과 전체 평균 2.5691 대비 상대비 예측 결과를 크기순으로 나타낸 표이다. 보험사는 개인들의 신용상태에 따른 상대위험도를 기존의 건강상태에 따른 상대위험도에 곱함으로써 보험가입자의 신용에 따른 위험을 포함시킬 수 있을 것이다. 그러나 이 또한 건강정보와 신용정보의 결합의 한계성으로 신용 변수들에 의한 수술발생이 건강상태에 따른 발생과 독립이라는 가정을 하였다는 한계점이 존재한다.

〈표 14〉 집단별 수술 예측 발생률 및 상대비 산출

최근 1년 연체	최근 1년 최장연체일	최근 1년 금융권대출	총 대출 기관수	현금 서비스	대상자	발생률 (%)	상대 위험도
있음	6개월 미만	없음	1개	없음	50	1.7040	0.6633
있음	6개월 미만	없음	없음	없음	847	1.7344	0.6751
있음	6개월 미만	없음	2개 이상	없음	41	1.7780	0.6921

최근 1년 연체	최근 1년 최장연체일	최근 1년 금융권대출	총 대출 기관수	현금 서비스	대상자	발생률 (%)	상대 위험도
있음	6개월 미만	없음	1개	있음	-	1.8579	0.7232
있음	6개월 미만	없음	없음	있음	-	1.8910	0.7361
있음	6개월 미만	없음	2개 이상	있음	-	1.9385	0.7545
있음	6개월 미만	있음	1개	없음	1,015	2.0178	0.7854
있음	6개월 미만	있음	없음	없음	196	2.0537	0.7994
있음	6개월 미만	있음	2개 이상	없음	3,572	2.1051	0.8194
있음	6개월 미만	있음	1개	있음	5	2.1994	0.8561
있음	6개월 미만	있음	없음	있음	-	2.2385	0.8713
없음	6개월 미만	없음	1개	없음	2,164	2.2507	0.8761
없음	6개월 미만	없음	없음	없음	720,829	2.2907	0.8916
있음	6개월 미만	있음	2개 이상	있음	40	2.2944	0.8931
없음	6개월 미만	없음	2개 이상	없음	907	2.3479	0.9139
없음	6개월 미만	없음	1개	있음	2	2.4527	0.9547
없음	6개월 미만	없음	없음	있음	343	2.4962	0.9716
없음	6개월 미만	없음	2개 이상	있음	1	2.5584	0.9958
없음	6개월 미만	있음	1개	없음	406,115	2.6624	1.0363
없음	6개월 미만	있음	없음	없음	78,642	2.7095	1.0546
없음	6개월 미만	있음	2개 이상	없음	298,089	2.7769	1.0809
있음	6개월 이상	없음	1개	없음	1,763	2.8781	1.1203
없음	6개월 미만	있음	1개	있음	9,219	2.9003	1.1289
있음	6개월 이상	없음	없음	없음	32,147	2.9290	1.1401
없음	6개월 미만	있음	없음	있음	629	2.9515	1.1488
있음	6개월 이상	없음	2개 이상	없음	1,561	3.0016	1.1683
없음	6개월 미만	있음	2개 이상	있음	71,870	3.0247	1.1773
있음	6개월 이상	없음	1개	있음	-	3.1347	1.2201
있음	6개월 이상	없음	없음	있음	4	3.1899	1.2416
있음	6개월 이상	없음	2개 이상	있음	-	3.2688	1.2723
있음	6개월 이상	있음	1개	없음	13,943	3.4007	1.3237
있음	6개월 이상	있음	없음	없음	4,034	3.4604	1.3469
있음	6개월 이상	있음	2개 이상	없음	29,995	3.5457	1.3801
있음	6개월 이상	있음	1개	있음	85	3.7020	1.4410
있음	6개월 이상	있음	없음	있음	7	3.7668	1.4662
없음	6개월 이상	없음	1개	없음	-	3.7870	1.4741
없음	6개월 이상	없음	없음	없음	-	3.8532	1.4998

최근 1년 연체	최근 1년 최장연체일	최근 1년 금융권대출	총 대출 기관수	현금 서비스	대상자	발생률 (%)	상대 위험도
있음	6개월 이상	있음	2개 이상	있음	363	3.8594	1.5022
없음	6개월 이상	없음	2개 이상	없음	-	3.9479	1.5367
없음	6개월 이상	없음	1개	있음	-	4.1212	1.6041
없음	6개월 이상	없음	없음	있음	-	4.1930	1.6321
없음	6개월 이상	없음	2개 이상	있음	-	4.2956	1.6720
없음	6개월 이상	있음	1개	없음	-	4.4670	1.7387
없음	6개월 이상	있음	없음	없음	-	4.5445	1.7689
없음	6개월 이상	있음	2개 이상	없음	-	4.6553	1.8120
없음	6개월 이상	있음	1개	있음	-	4.8580	1.8909
없음	6개월 이상	있음	없음	있음	-	4.9420	1.9236
없음	6개월 이상	있음	2개 이상	있음	-	5.0620	1.9703
전체					1,678,478	2.5691	-

〈표 14〉의 결과를 바탕으로 평균 2.5691 대비 상대비가 1미만, 즉 전체 수술 발생률보다 낮은 발생률의 집단을 그룹화하여 재산출하면 〈표 15〉와 같다. 보험사는 이러한 객관적인 수술발생 상대위험도를 통해 수술발생 상대위험도가 전체 평균보다 낮은 보험 수요자들에 대해 보험료 또는 수수료 할인 등의 혜택을 제공할 수 있다.

〈표 15〉 수술 발생의 상대비에 따른 재분류 결과

구분	상대위험도 < 1	상대위험도 ≥ 1
수술 발생률(%)	2.2887	2.7850
상대비	0.8908	1.0840

IV. 결론 및 시사점

본 연구는 신용정보원에서 제공하는 일반신용정보와 보험신용정보가 연계된 맞춤형DB를 활용하여 개인의 신용상태와 보험사고와의 발생관계를 통계적 모

형을 통해 인과관계를 분석하였다. 보험사고 예측모형 도출을 위해 반응변수가 이항변수인 경우 자주 사용되는 일반화선형모형 중 하나인 로지스틱회귀모형을 선택하였다. 기존의 연구에서 검증된 신용정보를 참고하여 연체, 대출, 신용 관련 설명변수들을 정의하였다. 통계적 모형은 로지스틱회귀모형을 이용하여 신용 상태에 따른 입원 및 수술 발생을 예측하였고, 최종적으로 개발된 통계적 모형에 기반하여 신용상태별 입원 및 수술 발생 위험도를 객관적으로 측정하였다.

보험소비자들의 신용을 반영한 보험상품 개발을 위한 법적인 측면과 소비자 보호측면에서의 정책적 고려사항은 다음과 같다. 첫 번째 법적 허용여부 측면이다. <신용정보법 제32조 제 6항 제9호 2의 개인신용정보의 제공.활용에 대한 동의>로 신용정보법 상 통계작성 등을 위한 가명 정보의 제공과 활용은 신용정보주체의 정보활용 동의가 면제되므로 가명 정보를 활용하여 통계 및 모형 개발은 가능하다. 또한 <보험업감독규정 제7-73조의 보험요율 산출의 원칙>에 의하면 보험업감독규정에 따라 '객관성 있는 국내외 통계자료 등'으로 신용모형을 활용한다면 보험요율 개발이 가능하다. 둘째로 소비자 보호 측면이다. <보험업감독업무 시행세칙 별표18 보험상품심사기준>에 의하면 기존 고객의 가입거절, 가입한도 제한, 부담보 및 보험료 상승 등에 신용모형 결과의 적용이 보험소비자에게 합당하지 않은 이유로 불리하게 이용되지 않아야 한다고 명시하고 있다. 이와 같이 보험상품 개발에 신용을 반영하는 것은 법적으로는 문제가 없고 개인의 신용위험을 정확히 예측함으로써 인해 역선택을 방지할 수 있는 장점이 있지만, 소비자 보호측면에서 저소득층, 신용불량자와 같은 사회적인 약자에게 보험료를 할증하여 불이익을 준다는 것은 상충된 측면이다. 그래서 신용을 반영한 보험상품 개발은 신용불량자에게 제약을 가하기보다는 신용우량자에게 보험료 또는 수수료 할인 등의 혜택을 주는 방향으로 이루어지는 것이 현재로서는 가장 적합한 방법으로 사료된다.

본 연구의 결과를 통한 기대효과 및 시사점은 다음과 같다.

첫째, 신용상태가 좋은 고객에게 보험료 할인을 제공하는 것은 고객에게 지속적인 신용관리 및 개선의 실질적인 동기부여가 되고, 보험회사는 고객의 신용개선 독려로 효율적인 보험리스크 관리를 해나갈 수 있다. 또한 보험업계 차원에서는 전통적인 보험업에서 소비자의 신용관리로 역할범위를 확장하는 계기

가 될 수 있을 것이다.

둘째, 신용정보원의 공공빅데이터 활용이 활성화 된다면 기존 보험시장의 확대뿐만 아니라 다양한 보험상품 개발로 이어져 보험소비자의 다양한 위험에 대한 보장서비스가 가능한 동시에 소비자의 보험상품 선택권이 넓어질 수 있을 것이다.

셋째, 최근 보험회사들은 보험업무 각 분야에 사물인터넷, 빅데이터, 인공지능 등 인슈어테크 도입을 활발히 추진하고 있으나, 가장 근간이 되는 양질의 보험 관련 데이터 확보에 어려움을 호소하고 있다. 향후 신용빅데이터를 기반으로 한 보험·신용정보 모형에 의해 보험리스크가 더욱 정확히 예측될 수 있을 것이며 신용정보 외에도 건강정보 등의 다양한 공공빅데이터가 보험산업에 도입될 수 있는 계기가 될 수 있을 것으로 기대한다.

참 고 문 헌

- 이상연 · 조은성 · 전성현 · 홍석철, “국민건강보험 빅데이터를 활용한 주요 만성질환 발생률 예측모형의 개발과 활용”, 보험학회지 133(23) (2023): 23-48.
- 정재욱 · 여윤희, “생명보험 언더라이팅 시 개인신용정보의 활용 효과 분석”, 금융연구 25(1) (2011): 25-56.
- 정재욱 · 손승호, “개인신용정보와 개인용자동차보험 사고발생확률 및 손해율 간의 상관관계 연구”, 리스크관리연구 20(2) (2009): 17-58.
- Dean, L. T., Knapp, E. A., Sngun, S., Ransome, Y., Qato, D. M., and Visvanathan, K., “Consumer credit, chronic disease and risk behaviours,” *Journal of Epidemiology Community Health* 73(1) (2019) 73-78.
- Dean, L. T., and Nicholas, L. H., K., “Using Credit Scores to Understand Predictors and Consequences of Disease,” *American Journal of Public Health; Washington* 108(11) (2018): 1503-1505.
- Dean, L. T., Schmitz, K. H., Frick, K. D., et al., “Consumer credit as a novel marker for economic burden and health after cancer in a diverse population of breast cancer survivors in the USA,” *Journal of Cancer Survivorship* 12(3) (2018): 306-315.
- Dobkin, C. Finkelstein, A., Kluender, R., and Notowidigdo, M. J., “The economic consequences of hospital admissions,” *American Economic Review* 108(2) (2018): 308-352.
- Golden, L. L., Brockett, P. L., Ai, J., and Kellison, B., “Empirical Evidence on the Use of Credit Scoring for Predicting Insurance Losses with Psychosocial and Biochemical Explanations,” *North American Actuarial Journal* 20(3) (2016): 233-251.
- Houle, J. N., Collins, J. M. and, Schmeiser, M. D., “Flu and finances: influenza outbreaks and loan defaults in US cities, 2004-2012,” *American Journal of Public Health* 105(9) (2015): 75-80.
- Israel, S., Caspi, A., Belsky, D. W., Harrington, H., Hogan, S., Houts, R., Ramrakha, S., Sanders, S., Richie, P., and Moffit, T. E., “Credit

- scores, cardiovascular disease risk, and human capital,” *Psychological and Cognitive Sciences* 111(48) (2014): 17087-17092.
- Li, Y., Gao, J., Enkavi, A. Z., Zaval, L., Weber, E. U., Johnson, E. J. (2015). “Sound credit scores and financial decisions despite cognitive aging,” *Economic Science* 112(1) (2015): 65-69.
- Knapp, E. A., and Dean, L. T., “Consumer credit scores as a novel tool for identifying health in urban U.S. neighborhoods,” *Annals of Epidemiology* 28(10) (2018): 724-729.
- Roberts, T., The use of credit scorecard design, predictive modelling and text mining to detect fraud in the insurance industry, North-West University Ph.D Dissertation, 2011.
- Wu, C. P., and Guszcza, J. C., “Does Credit Score Really Explain Insurance Losses? Multivariate Analysis from a Data Mining Point of View,” *Casualty Actuarial Society Forum Winter* (2003): 113-138.

