creative commons

COMMONS DEED

저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:

저작자표시. 귀하는 원저작자를 표시하여야 합니다.

비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.

변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 이용허락규약(Legal Code)을 이해하기 쉽게 요약한 것입니다.

Disclaimer

dCollection

의학석사 학위논문

국민건강보험공단 표본데이터를 이용한 대장암 예후와 관련된 처방 약물 체계적 평가

2020 년 10 월

서울대학교 대학원
의학과 예방의학 전공
우 형 택

국민건강보험공단 표본데이터를 이용한 대장암 예후와 관련된 처방 약물 체계적 평가

지도 교수 신 애 선

이 논문을 의학석사 학위논문으로 제출함

2020 년 10 월

서울대학교 대학원

의학과 예방의학 전공

우 형 택

우형택의 의학석사 학위논문을 인준함

2021 년 1 월

위 원 장 정 승 용 (인)

부위원장 신 애 선 (인)

위 원 이 중 엽 (인)

초 록

국민건강보험공단 표본데이터를 이용한 대장암 예후와 관련된 처방 약물 체계적 평가

우형택

예방의학교실

서울대학교 의과대학 대학원

배경

대장암 환자에서 흔히 처방되는 당뇨, 고지질혈증 등 비암성(non-cancer) 만성질환의 치료 약물과 대장암 환자의 예후를 살펴본 연구들은 특정 가설 기반하에 한 가지 또는 몇 가지의 약물과 대장암 예후와의 관계를 본 가설주도(hypothesis-driven) 연구로 선택적 보고(selective reporting)가 발생할 경우 편향된 결과가 나올 수 있다. 이에 본 연구에서는 선택적 보고의 문제를 원천적으로 막기 위하여 대장암 환자에서 처방되는 모든 약물과 대장암 예후와의 관계를 체계적으로 살펴본 Medication wide association study (MWAS)를 수행하였다.

방법

2002년에서 2013년까지 추적관찰된 국민건강보험공단 표본코호트DB를 이용하여 대장암으로 진단된 환자에서 진단 이후 처방된 약물과 환자의 예후와의 관련성을 분석하였다. 총 2,178명의 대상자를 무작위 추출을 통해 1:1로 나눠 약물발견셋(drug discovery set) 및 약물검증셋(drug validation set)으로 이용하였다. 다중검정에 따른 위양성 증가는 위발견율(false discovery rate)을 이용하여 통제하였고 약물 사용이 시간에 따라 변함으로 인해 발생하는 불멸시간편향(immortal time bias)은 랜드마크 시점을 1,3년으로 늘려가며 분석을 진행하여 통제하였다. 약물발견셋에서 p값은 0.05 미만이었지만 선정되지 않은 약물과 약물발견셋에서 선정된 약물 간의 약물검증셋에서의 내적타당도 평가(internal validation)에서 재현성의 차이가 있는지 추가 분석을 진행하였다.

결과

약물발견셋에서 총 773개의 약물에 대해 분석을 진행하였고 대장암 예후와 관련이 있는 것으로 나타난 약물은 랜드마크 시점 1년에서 117개, 3년에서 29개였다. 랜드마크 시점과

관계없이 모두 관련성을 보인 약물은 22개였다. 약물발견셋에서 랜드마크 시점과 관계없이 한번이라도 유의했던 124개의 약물을 약물검증셋에서 검증하였을 때 랜드마크 시점 1년에서 유의한 약물은 95개, 3년에서 유의한 약물은 61개였고 약물발견셋과 약물검증셋의 모든 랜드마크 시점에서 유의했던 약물은 H2-Receptor Antagonists(ATC code: A02BA), Other Drugs For Peptic Ulcer And Gastro-Oesophageal Reflux Disease (GORD)(ATC code: A02BX), Mosapride(ATC code: A03FA09), Prednicarbate(ATC code: D07AC18), Methylprednisolone(ATC code: H02AB04), Fluconazole(ATC code: J02AC01), Eperisone(ATC code: M03BX09), Enzymes (Used In Musculo-Skeletal System Inflammatory Conditions)(ATC code: M09AB), Mometasone(ATC code: R01AD09), Pseudoephedrine(ATC code: R01BA02), Acetylcysteine(ATC code: R05CB01), Erdosteine(ATC code: R05CB15), Levodropropizine(ATC code: R05DB27), Chlorphenamine(ATC code: R06AB04), Levocetirizine(ATC code: R06AE09), Other Antihistamines For Systemic Use(ATC code: R06AX) 등 총 17개였다. 약물검증셋에서 진행한 내적타당도 평가를 통한 재현성 차이에 대한 추가 검증에서 약물발견셋에서 MWAS 방법론을 통해 발견된 약물이 그렇지 않은 약물에 비해 약물검증셋에서 전반적으로 더 높은 재현성을 보여주었다.

결론

본 연구를 통해 총 17개의 대장암 환자 예후와 관계된 실마리정보(signal)를 찾아냈으며 MWAS 방법론이 새로운 가설을 찾아내는데 유용할 뿐만 아니라 연구의 재현성을 높이는데 사용될 수 있음을 확인하였다.

주요어 : 대장암, 처방 약물, 다중검정, 랜드마크 분석, 재현성, 선택적 보고
학 번 : 2015-23212

목 차

I. 서 론 1
1. 연구배경 및 필요성 1
2. 통계적 관련성(statistical association)과 인과성(causation) 2
3. 선택적 보고(selective reporting) 2
4. Medication wide association study (MWAS) 3

II. 연 구 방 법 5
1. 자료원 5
2. 대상자 선정 5
3. 연구 설계 및 통계분석 9
3.1. 단계적 연구 설계 9
3.2. 분석대상약물 8
3.3. 통계분석 24
3.3.1. 다중검정의 문제 해결 24
3.3.2. 랜드마크 분석(landmark analysis) 24
3.3.3. 생존분석 및 재현성 검정 25
4 이전 연구결과(previous study)와의 비교 25

III. 결 과 27
1. 연구대상자의 기본적 특징 27
2. Discovery set에서 MWAS 수행결과 30
3. Validation set에서 MWAS 수행결과 34
4. 재현성에 대한 분석 41
5. MWAS 결과와 기존에 대장암 예후와 관련이 있었던 약물과의 비교 43

IV. 고 찰 및 결 론 44
1. 대장암 예후와 관련된 약물의 해석 44
2. 재현성에 대한 분석결과 해석 44
3. 기존 연구에서 대장암 환자의 예후와 관련이 있었던 약물과의 비교결과 해석 46
4. 연구의 장점 및 한계 47
5. 결론 48

참고문헌 49

Abstract 53

표 목차

Table 1. List of colorectal cancer treatments in the National Health Insurance Service 7

Table 2. List of ATC code and name for medication 10

Table 3. General characteristics of the study subjects included in the discovery set and subjects included in the validation set 28

Table 4. Drug associated with colorectal cancer prognosis in discovery set at 1 year landmark point 31

Table 5. Drug associated with colorectal cancer prognosis in discovery set at 3 year landmark point 33

Table 6. Hazard ratio in the validation set based on significance in the discovery set at landmark point 35

Table 7. Number of significant drugs in relation to the selection at the landmark point in discovery set 42

Table 8. Comparison between the result of MWAS findings and candidate colorectal cancer prognosis associated drugs in meta-analyses 44

그림 목차

Figure 1. Schematic flow chart for the association analysis of prescribed medicines with colorectal cancer prognosis 26

I. 서론

1. 연구배경 및 필요성

2017년 중앙암등록본부 암등록통계자료에 따르면 대장암은 우리나라 암 발생률 순위에서 남, 여 모두 3위(연령표준화발생률 남: 39.9명/100,000명, 여: 23.0명/100,000명)를 기록하였고, 이와 더불어 암 조기발견과 치료 성적의 향상으로 대장암 생존율은 꾸준히 증가하여 2017년 암 유병자 1,867,405명 중 13.4%는 대장암 환자였다. [1] 음주, 흡연, 비만, 신체활동 저하와 같은 생활습관 요인은 대장암뿐만 아니라 비암성(non-cancer)만성질환에도 공통적으로 작용하는 위험요인으로서 고혈압, 당뇨, 고지혈증 등은 대장암 환자에서 흔히 동반되는 비암성 만성질환이다. [2, 3] 약물 시판 전에 흔하지 않은 유해작용(adverse effects)을 임상시험을 통해 밝혀내는 것은 제한적이기 때문에 약물 시판 후 약물감시(pharmacovigilance)를 통해 약물의 의도되지 않은 효과를 밝혀낼 필요가 있다. [4] 이와 관련하여 대장암 환자에서 흔히 처방되는 비암성 만성질환의 치료 약물과 대장암 예후를 살펴본 연구들이 활발하게 진행되고 있으며 대표적인 연구는 제2형 당뇨병 치료제인 메포민(metformin), 고지질혈증 치료제로 사용되고 있는 스타틴(statin), 순환기계 질환 예방에 효과가 있다고 알려진 아스피린(aspirin) 등과 대장암 환자의 예후를 본 연구들이다. [5-7] 하지만 이러한 연구들은 대부분 가설주도(hypothesis-driven)하에 한 가지 또는 몇 가지의 약물과 대장암 예후와의 관계를 본 연구로서 선택적 보고(selective reporting)가 발생할 경우 편향(bias)된 결과가 도출될 수 있다. [8, 9] 최근 빅데이터를 활용한 연구가 가능해진 이후 가설에 기반하지 않고 데이터주도(data-driven)로 한 번에 여러 가지 처방 약물과 결과변수와의 관계를 체계적으로 살펴본 연구들이 발표되고 있다. [10-13] 가설에 기반하지 않고 수행되는 데이터주도 분석 방법도 기본적으로 관찰연구가 본질적으로 가지고 있는 편향과 교란요인(confounder)에서 자유롭지 않다. 그러나 데이터주도로 체계적으로 수행되는 분석 방법이 가지는 장점은 이론적으로 다중검정(multiple testing)을 분석 상에서 보정하고 이를 투명하게 보고함으로써 선택적 분석에 따른 편향된 결과 보고를 원천적으로 막을 수 있다는 것이다. [14] 따라서 본 연구는 이러한 장점을 가진 데이터주도 분석 방법을 국민건강보험공단 표본데이터에 적용하여 대장암 생존자에서 처방되고 있는 약물들과 대장암 예후와의 관계를 체계적으로 평가하고자 한다. 데이터주도 분석을 위해서 MWAS (Medication wide association study)를 수행하는데 GWAS (Genome wide association study)가 유전형(genotype)과 결과변수와의 관련성을 평가하는 연구라면 MWAS는 약물(medication)과 결과변수와의 관련성을 평가하는 연구이다. [11]

2. 통계적 관련성(statistical association)과 인과성(causation)

역학연구에서 특정 요인과 결과 간에 인과성 추론(causal inference)을 중요하게 다루는 이유는 인구집단에서 발생하는 공중보건학적 문제점들을 해결하기 위한 1차예방법(primary prevention) 및 여러 중재법(intervention)들이 인과성에 기반하고 있기 때문이다. [15] 인과성 추론을 위해서는 먼저 관련성(association)이 전제되어야 하는데 이는 인과성을 추론하는데 있어 가장 필수적인 증거가 관련성이기 때문이다. 잘 설계된 역학연구를 통해 요인과 결과 간에 관련성을 밝힐 수 있으며, 역학연구에서는 이러한 관련성 여부를 확률론적 관점에서 통계적 관련성(statistical association)으로 판단한다. [16, 17] 역학연구를 통해 도출된 통계적 관련성이 타당한지에 대해서는 3가지의 가능성을 고려해야 한다. 첫 번째는 통계적 관련성이 우연에 의한 것인지 판단하는 것이다. 즉, 의학연구에서 흔히 사용되는 유의수준 0.05하에서 통계적으로 유의한 결과가 도출되어 연구가설을 채택했을 경우 연구가 오류 없이 완벽하게 수행되었음을 가정해도 연구결과가 우연에 의한 것임을 완전히 배제할 수는 없다. 왜냐하면 유의수준 0.05하에서 연구가설을 채택한다는 것은 귀무가설이 참임을 가정할 때 연구 데이터와 같은 분포가 5%의 이례적 확률로 관찰되어 귀무가설을 포기하고 연구가설을 채택한다는 것을 의미하기 때문이다. 따라서 낮은 확률이지만 우연에 의해 귀무가설이 참임에도 기각하게 되는 1종 오류(type 1 error)가 발생할 수 있다. 두 번째는 연구 설계, 연구대상자 모집, 연구 자료 수집 과정 등에서 체계적 오류(systematic error)인 편향이 발생하여 연구결과에 영향을 미치는 경우이다. 마지막 세 번째는 매개변수(mediating variable)가 아니면서 독립변수(independent variable)와 종속변수(dependent variable)와 모두 관계된 교란요인이 영향을 미치는 경우이다. 이상의 3가지 경우를 고려하여 연구를 통해 도출된 결론의 타당성을 판단할 수 있다. [18, 19]

3. 선택적 보고(selective reporting)

연구 분석과정에서 연구자의 주관이 개입되어 최초의 연구계획과는 달리 연구결과의 통계적 유의성 여부에만 기반하여 연구결과를 발표하는 것을 선택적 보고라고 한다. 선택적 보고는 통계적인 유의성 여부 외에는 특별한 의미가 없는 거짓된 결과를 의미 있는 것으로 도출하는 문제를 일으킨다. 실제로 여러 연구에서 통계적 모형을 통해 산출된 p값의 크기가 0.05 미만인지 여부만을 기준으로 결과를 발표하는 경우의 문제점을 지적하였다. [20-23] 특히 현재 학계에서 활발하게 논의되고 있는 재현성의 위기(reproducibility crisis)와 관련된 보고에서 재현 불가능성을 일으키는 가장 중요한 원인은 선택적 보고이며 이를 극복하기 위해 올바른 통계의 적용이

가장 중요하다고 조사되었다. [24] 통계적 유의성에 기반한 선택적 보고가 어떻게 재현성에 영향을 미치는지 구체적으로 살펴보면 앞서 설명했다시피 비인과적 관계에서도 우연, 편향, 교란에 기인하여 통계적으로 유의한 p값이 도출될 수 있는데 편향 및 교란은 연구 설계 및 분석 단계에서 어느 정도 통제가 가능하다. 즉, 동일한 연구 설계 및 분석을 통해 편향 및 교란의 통제에 있어서는 어느 정도 연구의 표준화가 가능하다. 실제로 관찰연구 질을 향상시키고 궁극적으로는 연구의 표준화를 위한 STrengthening the Reporting of OBservational Studies in Epidemiology (STROBE)와 같은 지침이 개발되어 있고 편향 및 교란과 관계된 연구대상자, 연구 설계, 교란 요인 등을 제시하도록 하고 있다. [25] 하지만 우연이 작용하여 통계적으로 유의한 결과가 도출되었고 이를 선택적으로 보고할 경우 다른 연구에서 재현이 어렵게 되며 특히 연구자의 의도가 개입하여 유의한 결과가 나올 때까지 통계 검정을 다수 시행하여 소위 말하는 p-hacking이 발생할 경우 재현성의 문제는 더욱 심각해진다. [26] 이러한 선택적 보고의 문제를 해결하기 위하여 연구의 사전 등록을 강화하고 있으며 임상시험의 경우 국제 의학저널 편집자 위원회(ICMJE, International Committee of Medical Journal Editors)의 2005년 선언에 따라 ICMJE 소속의 저널 논문 투고를 위해서는 임상시험 사전 등록이 필수화 되었다. [27] 사전 등록을 통해 연구의 투명성 및 재현성의 제고를 달성할 수 있지만 관찰연구의 경우 여전히 연구의 사전 등록이 논문 투고 시 필수 사항이 아닌 경우가 많고 이에 따라 사전 등록의 필요성이 떨어져 사전 등록의 참여도가 낮은 상황이다. [28] 사전 등록과 같은 제도적 강화와 더불어 선택적 보고를 막기 위해 올바른 통계적 검정 수행이 강조되고 있다. 특히 통계적 검정을 수행하는 연구에서 p값의 오용을 막기 위한 성명서가 2016년 미국통계학회에 의해 발표되었는데 선택적 보고에 대해서 언급한 4번째 원칙(principle 4)에서 적절한 통계적 추론 과정은 모든 것을 보고하고(full reporting) 투명하게 보고하는 것(transparency)을 통해 이루어진다고 서술하였다. 또한 의미 있는 과학적 발견은 연구 과정에서 몇 번의 통계 검정이 시행되었고, 보고된 p값은 어떻게 선택되었는지를 모두 밝히는 것을 통해 이루어진다고 덧붙였다. [29]

4. Medication wide association study (MWAS)

2002년 Ozaki 등에 의해 최초로 전장유전체연계분석(Genome-wide association study, GWAS)을 통해 심근경색과 관련된 단일 핵산염기 다형현상(Single Nucleotide Polymorphism, SNP)을 밝혀낸 이후 GWAS 방법론을 이용해 질병을 포함한 다양한 표현형과 관련된 SNP들을 밝혀냈다. [30, 31] GWAS는 생물학적 타당성(biological plausibility)과 이전 연구들에서 제기된 가설을 기반으로 수행되는 가설주도

(hypothsis-driven)의 단일유전자 접근법(single gene approach), 후보유전자 접근법(candidate gene approach) 등과는 달리 가설에 기반하지 않은(hypothesis-free)연구로서 단일유전자 접근법, 후보유전자 접근법 등의 가설주도 연구가 다른 연구에서 재현되지 않는데 대해서 이를 해결하기 위한 방법으로 대두되었다. [32] 실제로 2000년대 초반 시행된 한 연구에 따르면 후보유전자 접근법으로 시행된 166개 연구에서 6개의 연구만이 일관성 있게 재현이 되었고 63개의 연구는 한 번도 재현이 되지 않았으며 재현이 되지 않는 주요 원인으로 선택적 보고에 의한 출판편향(publication bias)을 뽑았다. [33] GWAS가 재현성의 문제를 해결할 수 있는 이유는 가설에 기반한 연구와는 달리 모든 유전자에 대한 분석 결과를 제시하여 출판편향을 줄여주기 때문이다. [31] 이와 같이 가설에 기반하지 않은 연구의 장점 때문에 전장유전체뿐만 아니라 후성유전체(Epigenome-wide association study, EWAS), 진유전체(Exome-wide association study, EWAS), 환경(Environment-wide association study, EWAS)등과 질병을 포함한 다양한 표현형과의 관계를 본 연구들이 발표되고 있으며, 최근에는 컴퓨터 연산 능력의 발달 및 빅데이터 기반의 약물 청구자료에 대한 접근성이 높아짐에 따라 약물과 질병과의 관계를 살펴본(Medication-wide association study, MWAS)연구들도 발표되고 있다. [11, 34-36] MWAS의 장점은 앞서 설명한 GWAS의 장점과 같다. 특정 약물이 아닌 모든 약물과 질병과의 관계를 검정하기 때문에 선택적 보고의 문제가 발생하지 않을 뿐만 아니라 2016년 미국통계학회 성명서의 4번 원칙과 같이 모든 과정을 투명하게 보고하기 때문에 다중검정의 횟수 및 이에 따른 p값에 대한 적절한 보정이 이루어져 기존 가설기반의 연구에 비해 위양성의 가능성이 떨어질 것이라 판단된다. [29] 이와 더불어 부차적으로 새로운 가설의 발견에도 도움을 줄 수 있다. 현재까지 국민건강보험공단 자료 또는 이와 구성이 유사한 국민건강보험심사평가원의 자료를 이용하여 MWAS와 유사한 방법으로 가설 없이 수행된 연구들이 있었지만, 대부분 연구의 초점은 재현성의 확보가 아닌 가설 없이 새로운 가설을 찾아내는 것이 목적이었다. [37-41] 본 연구의 목적은 새로운 가설을 찾아내는 것뿐만 아니라 MWAS방법론을 사용하여 체계적(systematic) 연구 수행을 통해 연구의 재현성을 높일 수 있는 가능성을 찾고 추가적으로 대장암의 예후와 관계된 약물에 대한 새로운 가설 발견 및 건강보험공단 표본데이터가 MWAS방법론 적용에 적절한지를 판단하는 것이다.

II. 연구방법

1. 자료원

본 연구는 2002년부터 2013년까지 추적관찰된 국민건강보험공단 표본코호트DB를 이용하여 수행하였다. 이 자료는 국민건강보험 가입자가 요양기관을 이용한 후 요양기관에서 심사청구를 진행하면서 생성된 자료 중 우리나라 인구의 약 2%인 1,025,340명을 계통적 층화 무작위 추출(systematic stratified random sampling)하여 연구용으로 공개한 자료이다. 실제 임상환경에서 수행된 진료행위에 대한 심사청구 과정에서 생성된 자료이기 때문에 실사용 데이터(real world data)의 성격을 지니고 있으며 2002년에서 2013년까지 12년간의 가입자 자격에 대한 정보, 의료기관 이용 내역, 건강검진결과, 요양기관 정보를 포함하고 있다. 자료는 크게 자격DB, 진료DB, 건강검진DB, 요양기관DB로 이루어지는데 각 DB에 포함된 정보를 간략히 살펴보면 자격DB에는 성별, 연령, 사망년월, 사망원인, 소득분위 등의 인구학적 정보를 진료DB에는 수진자의 요양개시일자, 한국표준질병사인분류코드(Korean Standard Classification of Diseases, KCD)에 기반한 상병명, 수술, 시술, 처방 약물 등의 진료내역 코드 등을 건강검진DB에는 건강검진 수검자의 건강검진 주요 결과 및 문진자료를 포함하고 있으며 마지막 요양기관DB에는 요양기관의 위치, 병상수, 인력 및 의료자원에 관한 정보 등을 포함하고 있다. 각 DB는 수진자 별로 부여된 조인키를 활용하여 연계하여 분석할 수 있으며 본 연구에서는 자격DB, 진료DB 중 명세서(20T), 진료내역(30T), 처방전교부상세(60T)를 활용하여 분석을 시행하였다. 세부 변수 내역은 건강보험공단 표본DB를 소개한 논문에서 확인이 가능하다. [42] 본 연구는 서울대학교병원 연구윤리심의위원회의 승인을 거쳤다. (IRB no. E-2007-005-1137)

2. 대상자 선정

국민건강보험공단 표본코호트DB의 경우 연구 목적으로 자료가 수집된 것이 아니라 심사 청구를 목적으로 수집되었기 때문에 실사용 데이터의 특성을 지닌다. 따라서 질병의 진단명만을 이용하여 대상자를 선정할 경우 대상자 선정이 부정확해지는 문제가 발생한다. 실제로 청구자료 진단명과 의무기록을 비교한 한 연구에 따르면 진단명과 의무기록의 일치도가 약 70%에 불과했다. [43] 이 때문에 정확한 대상자 선정을 위해서는 질병에 대한 조작적 정의가 필요하며 본 연구에서는 대장암 환자를 진료DB상에서 한국표준질병사인분류코드가 결장(C18) 또는 구불결장 이행부(C19) 또는 직장(C20)의 악성종양에 해당하는 진단코드를 가지면서 동시에 대장암 치료를 받은 환자로 정의하였다. 대장암 치료 중 수술, 항암, 방사선 치료에 대한 정보는 선행연구를 참고하였고 추가적으로 내시경적 수술에 대한 코드를 새롭게 포함하였

다. 세부코드는 Table 1과 같다. [44]

Table 1. List of colorectal cancer treatments in the National Health Insurance Service

Treatment	Claim codes
Colonoscopic Operation of Colonic Tumor	
Polypectomy	Q7701, Q7702
Mucosal Resection and Submucosal Resection	Q7703
Endoscopic Submucosal Dissection	QX706
Sigmoidoscopic Operation of Rectosigmoidal Tumor	
Polypectomy	Q7751
Mucosal Resection and Submucosal Resection	Q7752
Operation	
Rt. or Lt. hemicolectomy	QA671, Q2671
Subtotal colectomy	Q1261, Q1262
Total colectomy	QA672, Q2672
Segmental resection	QA673, Q2673
Colectomy with proximal colostomy and distal stump	QA679, Q2679
Transanal Rectal tumor resection	Q2891
Transsacral or parasacral rectal tumor resection	Q2890
Abdominal approach rectal tumor resection	Q2892
Transanal endoscopic microsurgery of rectal tumor resection	Q2893
Anterior resection	QA921, Q2921
Low anterior resection	Q2927, QA922, Q2922
Abdominoperineal resection	QA923, Q2923
Abdominal pull-through operation	QA924, Q2924
Total coloprotectomy with ileostomy	QA925, Q2925
Total coloprotectomy with ileal pouch-anal anastomosis	QA926, Q2926

Table 1. Continued

Treatment	Claim codes
Chemotherapy	
Capecitabine	122701ATB, 122702ATB
5-FU	161430BIJ, 161431BIJ, 161432BIJ
Leucovorin	566132BIJ, 566134BIJ, 622630BIJ, 622631BIJ, 622632BIJ, 521001BIJ, 521002BIJ
Irinotecan	177430BIJ, 177431BIJ, 177432BIJ, 177433BIJ, 177434BIJ, 177435BIJ, 177436BIJ, 177437BIJ
Bevacizumab	554330BIJ, 554331BIJ
Cetuximab	556430BIJ
Oxaliplatin	205830BIJ, 205834BIJ,205803BIJ, 205832BIJ
Radiation therapy	
	HD051, HD054, HD052, HD055, HD053, HD056, HD057, HD058, HD059, HD061, HD071, HD072, HD073, HD080, HD081, HD081, HD082, HD083, HD084, HD085, HD086, HD087, HD088, HD089, HD111, HD112, HZ271

3. 연구 설계 및 통계분석

3.1 단계적 연구 설계

연구가 반복해서 재현(replication)되는지를 확인하기 위한 표준방법(gold standard)은 본 연구에서 사용된 건강보험공단 표본코호트DB와 유사한 코호트 내에서 독립적으로 연구를 시행하여 결과가 동일하게 재현되는지, 즉 외적타당도를 검증(external validation)하는 것이다. 하지만 외적타당도 검증을 위한 외부 자료원이 제한될 경우 동일 자료원을 무작위 추출(random sampling) 방식으로 나눈 후 각각의 데이터셋에서 순차적으로 연구를 진행하여 내적타당도를 평가(internal validation)하는 Two-Stage (TS) 방법을 적용할 수 있다. [45] 본 연구에서는 이를 위해 건강보험공단 표본코호트 DB를 무작위 추출하여 1:1로 나눈 후 각각을 약물발견셋(drug discovery set; 이하 discovery set) 및 약물검증셋(drug validation set; 이하 validation set)으로 설정하고 먼저 약물발견셋에서 약물을 찾고 약물검증셋을 통해 검증하였다. 대상자 선정 기준에 따라 대상자를 선정한 결과 정의에 부합하는 인원은 총 2,178명이었으며 1:1로 나눈 결과 discovery set 및 validation set에 해당하는 인원은 모두 1,089명이었다.

3.2 분석대상약물

분석대상약물은 진료DB 중 60T, 즉 원외로 처방되는 약물 중 discovery set과, validation set에서 주성분 코드 앞 4자리를 기준으로 두 데이터셋 모두에서 1회 이상 사용된 약물로 하였고 약제의 ATC (Anatomical Therapeutic Chemical Classification System) 코드가 같은 약물을 같은 약물군으로 묶어 분석하였다. 분석에 포함된 약물군은 총 773개이며 약물군의 목록은 Table 2와 같다.

Table 2. List of ATC code and name for medication

ATC Code	Name	ATC Code	Name	ATC Code	Name
A01A	Stomatological Preparations	A02AX	Antacids, Other Combinations	A02BX08	Sulglicotide
A01AB22	Doxycycline	A02BA	H2-Receptor Antagonists	A02BX11	Troxipide
A01AC02	Dexamethasone	A02BA01	Cimetidine	A02BX13	Alginic Acid
A01AD	Other Agents For Local Oral Treatment	A02BA02	Ranitidine	A02X	Other Drugs For Acid Related Disorders
A01AD02	Benzydamine	A02BA03	Famotidine	A03	Drugs For Functional Gastrointestinal Disorders
A02A	Antacids	A02BA04	Nizatidine	A03A	Drugs For Functional Gastrointestinal Disorders
A02AA	Magnesium Compounds	A02BA06	Roxatidine	A03AA04	Mebeverine
A02AA02	Magnesium Oxide	A02BA08	Lafutidine	A03AA05	Trimebutine
A02AA04	Magnesium Hydroxide	A02BB01	Misoprostol	A03AA09	Difemerine
A02AB	Aluminium Compounds	A02BC	Proton Pump Inhibitors	A03AB	Synthetic Anticholinergics, Quaternary Ammonium Compounds
A02AB01	Aluminium Hydroxide	A02BC01	Omeprazole	A03AB06	Otilonium Bromide
A02AB03	Aluminium Phosphate	A02BC02	Pantoprazole	A03AB17	Tiemonium Iodide
A02AC01	Calcium Carbonate	A02BC03	Lansoprazole	A03AC	Synthetic Antispasmodics, Amides With Tertiart Amines
A02AD01	Ordinary Salt Combinations	A02BC04	Rabeprazole	A03AC05	Tiropramide
A02AD03	Almagate	A02BC05	Esomeprazole	A03AX	Synthetic Anticholinergics, Quaternary Ammonium Compounds
A02AD04	Hydrotalcite	A02BC06	Dexlansoprazole	A03AX04	Pinaverium
A02AD05	Almasilate	A02BX	Other Drugs For Peptic Ulcer And Gastro-Oesophageal Reflux Disease (GORD)	A03AX05	Fenoverine
A02AF	Antacids With Antiflatulents	A02BX02	Sucralfate	A03AX11	Caroverine
A02AG	Antacids With Antispasmodics	A02BX05	Bismuth Subcitrate	A03AX12	Phloroglucinol

Table 2. Continued

ATC Code	Name	ATC Code	Name	ATC Code	Name
A03AX13	Silicones	A04AD	Other Antiemetics	A07AX03	Nifuroxazide
A03AX58	Alverine, Combinations	A04AD12	Aprepitant	A07BC05	Diosmectite
A03BA01	Atropine	A05AA	Bile Acid And Derivatives	A07DA03	Loperamide
A03BB	Belladonna alkaloids, semisynthetic, quaternary ammonium compounds	A05AA02	Ursodeoxycholic Acid	A07DA05	Loperamide Oxide
A03BB01	Butylscopolamine	A05AX	Other Drugs For Bile Therapy	A07EA07	Beclometasone
A03BB05	Cimetropium Bromide	A05BA	Liver Therapy	A07EC01	Sulfasalazine
A03CA02	Clidinium And Psycholeptics	A05BA03	Silymarin	A07EC02	Mesalazine
A03F	Propulsives	A05BA06	Ornithine Oxoglurate	A07FA01	Lactic Acid Producing Organisms
A03FA	Propulsives	A06AB02	Bisacodyl	A07FA02	Saccharomyces Boulardii
A03FA01	Metoclopramide	A06AC01	Ispaghula (Psylla Seeds)	A07XA	Other Antidiarrheals
A03FA03	Domperidone	A06AC08	Polycarbophil Calcium	A09A	Digestives, Incl. Enzymes
A03FA06	Clebopride	A06AD11	Lactulose	A09AA	Enzyme Preparations
A03FA07	Itopride	A06AD12	Lactitol	A09AA02	Multienzymes (Lipase, Protease Etc.)
A03FA09	Mosapride	A06AD15	Macrogol	A10AB04	Insulin Lispro
A04AA	Serotonin (5Ht3) antagonists	A06AD17	Sodium Phosphate	A10AB05	Insulin Aspart
A04AA01	Ondansetron	A06AD19	Magnesium Citrate	A10AC01	Insulin (Human)
A04AA02	Granisetron	A06AD65	Macrogol, Combinations	A10AD01	Insulin (Human)
A04AA03	Tropisetron	A07AA02	Nystatin	A10AE04	Insulin Glargin
A04AA04	Dolasetron	A07AA11	Rifaximin	A10AE05	Insulin Detemir

Table 2. Continued

ATC Code	Name	ATC Code	Name	ATC Code	Name
A10BA02	Metformin	A10BX02	Repaglinide	B01AA03	Warfarin
A10BB01	Glibenclamide	A10BX03	Nateglinide	B01AB	Heparin Group
A10BB08	Gliquidone	A10BX08	Mitiglinide	B01AB11	Sulodexide
A10BB09	Gliclazide	A11CC03	Alfacalcidol	B01AC	Platelet aggregation inhibitors excl. heparin
A10BB12	Glimepiride	A11CC04	Calcitriol	B01AC04	Clopidogrel
A10BD02	Metformin And Sulfonylureas	A11DA01	Thiamine (Vit B1)	B01AC05	Ticlopidine
A10BD03	Metformin And Rosiglitazone	A11EB	Vitamin B-Complex With Vitamin C	B01AC06	Acetylsalicylic acid
A10BD07	Metformin And Sitagliptin	A11GA01	Ascorbic Acid (Vit C)	B01AC10	Indobufen
A10BD08	Metformin And Vildagliptin	A11GB01	Ascorbic Acid (Vit C) And Calcium	B01AC18	Triflusal
A10BD11	Metformin And Linagliptin	A11HA02	Pyridoxine (Vit B6)	B01AC19	Beraprost
A10BF01	Acarbose	A11HA30	Dexpanthenol	B01AC23	Cilostazol
A10BF03	Voglibose	A12AA	Calcium	B01AC30	Combinations
A10BG02	Rosiglitazone	A12AA20	Calcium (Different Salts In Combination)	B01AX	Other Antithrombotic agents
A10BG03	Pioglitazone	A12AX	Calcium, Combination With Vit D and/or other drugs	B02AA02	Tranexamic Acid
A10BH01	Sitagliptin	A12BA01	Potassium Chloride	B02AB04	Camostat
A10BH02	Vildagliptin	A12CC	Magnesium	B02BX02	Carbazochrome
A10BH03	Saxagliptin	A16AA02	Ademetionine	B03A	Iron Preparations
A10BH05	Linagliptin	A16AX	Various Alimentary Tract And Metabolism Products	B03AA07	Ferrous Sulfate
A10BH06	Gemigliptin	A16AX01	Thioctic Acid	B03AB	Iron Trivalent, Oral preparations

Table 2. Continued

ATC Code	Name	ATC Code	Name	ATC Code	Name
B03AB05	Ferric Oxide Polymaltose Complexes	C01DX16	Nicorandil	C04AX20	Buflomedil
B03AB08	Ferric Acetyl Transferrin	C01EB09	Ubidecarenone	C04AX21	Naftidrofuryl
B03AB09	Ferric Proteinsuccinylate	C01EB15	Trimetazidine	C04AX28	Ifenprodil
B03AD04	Ferric Oxide Polymaltose Complexes	C02CA01	Prazosin	C05A	Agents For Treatment Of Hemorrhoids And Anal Fissures For Topical Use
B03BA05	Mecobalamin	C02CA04	Doxazosin	C05AA	Corticosteroids
B03BB01	Folic Acid	C03AA03	Hydrochlorothiazide	C05AD07	Pramocaine
B05BB01	Electrolytes	C03BA08	Metolazone	C05AX03	Other Preparations, Combinations
B05XA03	Sodium Chloride	C03BA11	Indapamide	C05B	Antivaricose Therapy
B05Z	Hemodialytics And Hemofiltrates	C03CA	Sulfonamides, Plain	C05BX01	Calcium Dobesilate
B06AA	Enzymes With Fibrinolytic Properties	C03CA01	Furosemide	C05CA03	Diosmin
C01AA05	Digoxin	C03CA04	Torasemide	C05CA53	Diosmin, Combinations
C01BC03	Propafenone	C03DA01	Spironolactone	C05CX	Other Capillary Stabilizing Agents
C01BC04	Flecainide	C03DB01	Amiloride	C05CX02	Naftazone
C01BD01	Amiodarone	C03EA01	Hydrochlorothiazide And Potassium-Sparing Agents	C05CX03	Hippocastani Semen
C01CA17	Midodrine	C04AC	Nicotinic Acid And Derivatives	C07AA05	Propranolol
C01DA02	Glyceryl Trinitrate	C04AD03	Pentoxifylline	C07AA15	Carteolol
C01DA08	Isosorbide Dinitrate	C04AE02	Nicergoline	C07AB02	Metoprolol
C01DA14	Isosorbide Mononitrate	C04AF01	Kallidinogenase	C07AB03	Atenolol
C01DX12	Molsidomine	C04AX	Other Peripheral vasodilators	C07AB05	Betaxolol

Table 2. Continued

ATC Code	Name	ATC Code	Name	ATC Code	Name
C07AB07	Bisoprolol	C08DA01	Verapamil	C09CA10	Fimasartan
C07AB11	S-Atenolol	C08DB01	Diltiazem	C09DA01	Losartan And Diuretics
C07AB12	Nebivolol	C09AA01	Captopril	C09DA02	Eprosartan And Diuretics
C07AG02	Carvedilol	C09AA02	Enalapril	C09DA03	Valsartan And Diuretics
C07BB07	Bisoprolol And Thiazides	C09AA03	Lisinopril	C09DA04	Irbesartan And Diuretics
C07CB03	Atenolol And Other Diuretics	C09AA04	Perindopril	C09DA06	Candesartan And Diuretics
C08CA	Dihydropyridine	C09AA05	Ramipril	C09DA07	Telmisartan And Diuretics
C08CA01	Amlodipine	C09AA08	Cilazapril	C09DA08	Olmesartan Medoxomil And Diuretics
C08CA02	Felodipine	C09AA09	Fosinopril	C09DB01	Valsartan And Amlodipine
C08CA03	Isradipine	C09AA13	Moexipril	C09DB02	Olmesartan Medoxomil And Amlodipine
C08CA05	Nifedipine	C09AA16	Imidapril	C09DB04	Telmisartan And Amlodipine
C08CA06	Nimodipine	C09BB05	Ramipril And Felodipine	C09DB06	Losartan And Amlodipine
C08CA07	Nisoldipine	C09CA01	Losartan	C09DX03	Olmesartan Medoxomil, Amlodipine And Hydrochlorothiazide
C08CA09	Lacidipine	C09CA02	Eprosartan	C10AA01	Simvastatin
C08CA11	Manidipine	C09CA03	Valsartan	C10AA02	Lovastatin
C08CA12	Barnidipine	C09CA04	Irbesartan	C10AA03	Pravastatin
C08CA13	Lercanidipine	C09CA06	Candesartan	C10AA04	Fluvastatin
C08CA14	Cilnidipine	C09CA07	Telmisartan	C10AA05	Atorvastatin
C08CA15	Benidipine	C09CA08	Olmesartan Medoxomil	C10AA07	Rosuvastatin

Table 2. Continued

ATC Code	Name	ATC Code	Name	ATC Code	Name
C10AA08	Pitavastatin	D01AE16	Amorolfine	D07AB08	Desonide
C10AB02	Bezafibrate	D01AE23	Butenafine	D07AB10	Alclometasone
C10AB04	Gemfibrozil	D01BA01	Griseofulvin	D07AB19	Dexamethasone
C10AB05	Fenofibrate	D01BA02	Terbinafine	D07AC01	Betamethasone
C10AD02	Nicotinic Acid	D02AE01	Carbamide	D07AC03	Desoximetasone
C10AX06	Omega-3-Triglycerides Incl. Other Esters And Acids	D02AX	Other Emollients And Protectives	D07AC06	Diflucortolone
C10BA02	Simvastatin And Ezetimibe	D03AX	Other Cicatrizants	D07AC08	Fluocinonide
C10BA03	Pravastatin And Fenofibrate	D03AX12	Trolamine	D07AC10	Diflorasone
C10BX03	Atorvastatin And Amlodipine	D05AX02	Calcipotriol	D07AC11	Amcinonide
D01A	Antifungals For Topical Use	D05AX04	Tacalcitol	D07AC14	Methylprednisolone aceponate
D01AC	Imidazole And Triazole Derivative	D05AX52	Calcipotriol, Combinations	D07AC18	Prednicarbate
D01AC03	Econazole	D05BB02	Acitretin	D07AC19	Difluprednate
D01AC05	Isoconazole	D06AX	Other Antibiotics For Topical Use	D07AD01	Clobetasol
D01AC08	Ketoconazole	D06AX09	Mupirocin	D07CB01	Triamcinolone And Antibiotics
D01AC14	Sertaconazole	D06BA01	Silver Sulfadiazine	D07CC01	Betamethasone And Antibiotics
D01AC16	Flutrimazole	D07AA02	Hydrocortisone	D08AE02	Policresulen
D01AC20	Imidazoles/Triazoles In Combination With Corticosteroids	D07AA03	Prednisolone	D11AH02	Pimecrolimus
D01AE14	Ciclopirox	D07AB01	Clobetasone	D11AX02	Gamolenic Acid
D01AE15	Terbinafine	D07AB02	Hydrocortisone Butyrate	D11AX12	Pyrithione Zinc

Table 2. Continued

ATC Code	Name	ATC Code	Name	ATC Code	Name
G01AA51	Nystatin, Combinations	G04BD06	Propiverine	H02AB08	Triamcinolone
G01AF02	Clotrimazole	G04BD07	Tolterodine	H02AB09	Hydrocortisone
G03CA03	Estradiol	G04BD08	Solifenacin	H02AB13	Deflazacort
G03CA04	Estriol	G04BD09	Trospium	H02BX	Corticosteroids For Systemic Use, Combinations
G03CA07	Estrone	G04BD11	Fesoterodine	H03AA01	Levothyroxine Sodium
G03CA09	Promestriene	G04BX	Other Urologicals	H03AA02	Liothyronine Sodium
G03CA53	Estradiol, Combinations	G04CA	Alpha-Adrenoreceptor Antagonists	H03BA02	Propylthiouracil
G03CA57	Conjugated Estrogens	G04CA01	Alfuzosin	H03BB02	Thiamazole
G03CX01	Tibolone	G04CA02	Tamsulosin	H03CA	Iodine Therapy
G03DA02	Medroxyprogeste	G04CA03	Terazosin	J01AA02	Doxycycline
G03DB06	Chlormadinone	G04CA04	Silodosin	J01AA07	Tetracycline
G03DC02	Norethisterone	G04CB01	Finasteride	J01AA08	Minocycline
G03FA17	Drospirenone And Estrogen	G04CB02	Dutasteride	J01AA20	Combinations Of Tetracyclines
G03HA01	Cyproterone	G04CX01	Prunus Africanae Cortex	J01CA	Penicillins With Extended Spectrum
G03HB01	Cyproterone And Estrogen	G04CX02	Sabalis Serrulatae Ftuctus	J01CA01	Ampicillin
G03X	Other Sex Hormones And Modulators Of The Genital System	H01BA02	Desmopressin	J01CA02	Pivampicillin
G03XC01	Raloxifene	H02AB02	Dexamethasone	J01CA04	Amoxicillin
G04BD02	Flavoxate	H02AB04	Methylprednisolone	J01CA06	Bacampicillin
G04BD04	Oxybutynin	H02AB06	Prednisolone	J01CR02	Amoxicillin And Beta-Lactamase Inhibitor

Table 2. Continued

ATC Code	Name	ATC Code	Name	ATC Code	Name
J01CR04	Sultamicillin	J01EE01	Sulfamethoxazol And Trimethoprim	J01MA14	Moxifloxacin
J01DB01	Cefalexin	J01EE02	Sulfadiazine And Trimethoprim	J01MA15	Gemifloxacin
J01DB05	Cefadroxil	J01FA01	Erythromycin	J01MA22	Tosufloxacin
J01DB07	Cefatrizine	J01FA02	Spiramycin	J01MB04	Pipemidic Acid
J01DB09	Cefradine	J01FA03	Midecamycin	J01RA04	Spiramycin And Metronidazole
J01DB11	Cefroxadine	J01FA06	Roxithromycin	J01XC01	Fusidic Acid
J01DC02	Cefuroxime	J01FA09	Clarithromycin	J01XD01	Metronidazole
J01DC04	Cefaclor	J01FA10	Azithromycin	J01XX01	Fosfomycin
J01DC07	Cefotiam	J01FF01	Clindamycin	J02AC01	Fluconazole
J01DC10	Cefprozil	J01FF02	Lincomycin	J02AC02	Itraconazole
J01DD05	Cefmenoxime	J01GB01	Tobramycin	J04AB02	Rifampicin
J01DD08	Cefixime	J01GB03	Gentamicin	J04AC01	Isoniazid
J01DD10	Cefetamet	J01MA	Fluoroquinolone	J04AK01	Pyrazinamide
J01DD13	Cefpodoxime	J01MA01	Ofloxacin	J04AK02	Ethambutol
J01DD14	Ceftibuten	J01MA02	Ciprofloxacin	J04BA02	Dapsone
J01DD15	Cefdinir	J01MA06	Norfloxacin	J05AB01	Aciclovir
J01DD16	Cefditoren	J01MA07	Lomefloxacin	J05AB09	Famciclovir
J01DD17	Cefcapene	J01MA08	Fleroxacin	J05AB11	Valaciclovir
J01DD18	Cefteram	J01MA12	Levofloxacin	J05AF05	Lamivudine

Table 2. Continued

ATC Code	Name	ATC Code	Name	ATC Code	Name
J05AF07	Tenofovir Disoproxil	M01AB11	Acemetacin	M01AE52	Naproxen And Esomeprazole
J05AF10	Entecavir	M01AB14	Proglumetacin	M01AG01	Mefenamic Acid
J05AF11	Telbivudine	M01AB15	Ketorolac	M01AG02	Tolfenamic Acid
J05AH02	Oseltamivir	M01AB16	Aceclofenac	M01AH01	Celecoxib
J05AP01	Ribavirin	M01AB55	Diclofenac, Combinations	M01AH02	Rofecoxib
J05AX02	Lysozyme	M01AC01	Piroxicam	M01AX	Other Antiinflammatory And Antirheumatic Agents, Non-steroids
J05AX05	Inosine Pranobex	M01AC05	Lornoxicam	M01AX01	Nabumetone
L02AA01	Diethylstilbestrol	M01AC06	Meloxicam	M01AX05	Glucosamine
L02AB01	Megestrol	M01AE	Propionic Acid	M01AX17	Nimesulide
L02BB03	Bicalutamide	M01AE01	Ibuprofen	M01AX21	Diacerein
L02BG03	Anastrozole	M01AE02	Naproxen	M01AX22	Morniflumate
L03AX	Other Immunostimulants	M01AE03	Ketoprofen	M01AX25	Chondroitin Sul
L04AA13	Leflunomide	M01AE04	Fenoprofen	M01AX26	Avocado And Soybean Oil, Unsaponifiables
L04AX03	Methotrexate	M01AE11	Tiaprofenic Aci	M02AA10	Ketoprofen
M01AB	Acetic Acid Derivatives And Related Substances	M01AE12	Oxaprozin	M02AA15	Diclofenac
M01AB02	Sulindac	M01AE13	Ibuproxam	M02AA19	Flurbiprofen
M01AB05	Diclofenac	M01AE14	Dexibuprofen	M02AA23	Indometacin
M01AB08	Etodolac	M01AE16	Alminoprofen	M03BA03	Methocarbamol
M01AB09	Lonazolac	M01AE51	Ibuprofen, Combinations	M03BC01	Orphenadrine (Citrate)

Table 2. Continued

ATC Code	Name	ATC Code	Name	ATC Code	Name
M03BX	Other Centrally Acting Agents (Muscle Relaxants)	M09AB	Enzymes (Used In Musculo-Skeletal System Inflammatory Conditions)	N02CC01	Sumatriptan
M03BX01	Baclofen	M09AB03	Bromelains	N02CC02	Naratriptan
M03BX02	Tizanidine	M09AB52	Trypsin, Combinations	N02CX	Other Antimigraine Preparations
M03BX03	Pridinol	M09AX	Other Drugs For Disorders Of The Musculo-Skeletal System	N03AA02	Phenobarbital
M03BX04	Tolperisone	N01BB02	Lidocaine	N03AB02	Phenytoin
M03BX05	Thiocolchicoside	N01BB10	Levobupivacaine	N03AE01	Clonazepam
M03BX08	Cyclobenzaprine	N02AA05	Oxycodone	N03AF01	Carbamazepine
M03BX09	Eperisone	N02AA55	Oxycodone And Naloxone	N03AF02	Oxcarbazepine
M03CA01	Dantrolene	N02AA59	Codeine, Combinations Excl. Psycholeptics	N03AG01	Valproic Acid
M04AA01	Allopurinol	N02AB03	Fentanyl	N03AX09	Lamotrigine
M04AA03	Febuxostat	N02AD01	Pentazocine	N03AX11	Topiramate
M04AC01	Colchicine	N02AE01	Buprenorphine	N03AX12	Gabapentin
M05BA03	Pamidronic Acid	N02AJ13	Tramadol And Paracetamol	N03AX14	Levetiracetam
M05BA04	Alendronic Acid	N02AX02	Tramadol	N03AX16	Pregabalin
M05BA06	Ibandronic Acid	N02BA01	Acetylsalicylic Acid	N04AA04	Procyclidine
M05BA07	Risedronic Acid	N02BA06	Salsalate	N04AC01	Benzatropine
M05BB03	Alendronic Acid	N02BA16	Imidazole Salicylate	N04BA02	Levodopa And Decarboxylase Inhibitor
M05BB07	Risedronic Acid And Colecalciferol	N02BE01	Paracetamol	N04BA03	Levodopa, Decarboxylase Inhibitor And COMT Inhibitor
M05BX	Other Drugs Affecting Bone Structure And Mineralization	N02CA52	Ergotamine, Combinations Excl. Psycholeptics	N04BB01	Amantadine

Table 2. Continued

ATC Code	Name	ATC Code	Name	ATC Code	Name
N04BC04	Ropinirole	N05BA19	Etizolam	N06AX05	Trazodone
N04BD01	Selegiline	N05BA21	Clotiazepam	N06AX11	Mirtazapine
N05AA01	Chlorpromazine	N05BA23	Tofisopam	N06AX12	Bupropion
N05AB03	Perphenazine	N05BB01	Hydroxyzine	N06AX14	Tianeptine
N05AD01	Haloperidol	N05BE01	Buspirone	N06AX16	Venlafaxine
N05AH03	Olanzapine	N05CD01	Flurazepam	N06AX17	Milnacipran
N05AH04	Quetiapine	N05CD03	Flunitrazepam	N06AX21	Duloxetine
N05AN01	Lithium	N05CD05	Triazolam	N06AX25	Hyperici Herba
N05AX08	Risperidone	N05CF02	Zolpidem	N06BX03	Piracetam
N05AX12	Aripiprazole	N06AA02	Imipramine	N06BX07	Oxiracetam
N05BA01	Diazepam	N06AA09	Amitriptyline	N06BX12	Acetylcarnitine
N05BA02	Chlordiazepoxid	N06AA10	Nortriptyline	N06DA02	Donepezil
N05BA05	Potassium Clorazepate	N06AA23	Quinupramine	N06DA03	Rivastigmine
N05BA06	Lorazepam	N06AB03	Fluoxetine	N06DA04	Galantamine
N05BA08	Bromazepam	N06AB04	Citalopram	N06DX01	Memantine
N05BA09	Clobazam	N06AB05	Paroxetine	N06DX02	Ginkgo Folium
N05BA12	Alprazolam	N06AB06	Sertraline	N07AB02	Bethanechol
N05BA14	Pinazepam	N06AB08	Fluvoxamine	N07AX01	Pilocarpine
N05BA18	Ethyl Loflazepa	N06AB10	Escitalopram	N07AX02	Choline Alfoscerate

Table 2. Continued

ATC Code	Name	ATC Code	Name	ATC Code	Name
N07BB03	Acamprosate	R01BA52	Pseudoephedrine, Combinations	R03CC08	Procaterol
N07CA01	Betahistine	R01BA53	Phenylephrine, Combinations	R03CC11	Tulobuterol
N07CA03	Flunarizine	R02AA05	Chlorhexidine	R03CC12	Bambuterol
N07XX	Other Nervous System Drugs	R02AA09	Benzethonium	R03CK	Adrenergics And Other Drugs For Obstructive Airway Diseases
P01AB02	Tinidazole	R02AA15	Povidone-Iodine	R03DA	Xanthines
P01BA02	Hydroxychloroquine	R03AK06	Salmeterol And Fluticasone	R03DA02	Choline Theophyllinate
P02BA01	Praziquantel	R03AK07	Formoterol And Budesonide	R03DA04	Theophylline
P03AB02	Lindane	R03AK08	Formoterol And Beclometasone	R03DA05	Aminophylline
P03AX	Other Ectoparasiticides, Incl. Sccbicides	R03AL02	Salbutamol And Ipratropium Bromide	R03DA06	Etamiphylline
QJ01FA9	Kitamycin	R03BA02	Budesonide	R03DA07	Theobromine
R01AB06	Xylometazoline	R03BA08	Ciclesonide	R03DA08	Bamifylline
R01AC02	Levocabastine	R03BB04	Tiotropium Bromide	R03DA09	Acefylline Piperazine
R01AD	Corticosteroids	R03CB03	Orciprenaline	R03DA11	Doxofylline
R01AD05	Budesonide	R03CB53	Orciprenaline, Combinations	R03DA51	Diprophylline, Combinations
R01AD08	Fluticasone	R03CC	Selective Beta-2 Adrenoreceptor Agonists	R03DC	Leukotriene Receptor Antagonists
R01AD09	Mometasone	R03CC02	Salbutamol	R03DC01	Zafirlukast
R01AD12	Fluticasone Furoate	R03CC03	Terbutaline	R03DC02	Pranlukast
R01AX03	Ipratropium Bromide	R03CC04	Fenoterol	R03DC03	Montelukast
R01BA02	Pseudoephedrine	R03CC05	Hexoprenaline	R03DX07	Roflumilast

Table 2. Continued

ATC Code	Name	ATC Code	Name	ATC Code	Name
R05	Cough And Cold Preparations	R05DB07	Oxolamine	R06AX17	Ketotifen
R05CA	Expectorants	R05DB13	Butamirate	R06AX19	Azelastine
R05CA03	Guaifenesin	R05DB15	Zipeprol	R06AX22	Ebastine
R05CA12	Hederae Helicis	R05DB21	Cloperastine	R06AX24	Epinastine
R05CB	Mucolytics	R05DB24	Tipepidine	R06AX25	Mizolastine
R05CB01	Acetylcysteine	R05DB27	Levodropropizine	R06AX26	Fexofenadine
R05CB02	Bromhexine	R05FA	Opium Derivatives And Expectorant	R06AX27	Desloratadine
R05CB03	Carbocisteine	R05FA02	Opium Derivatives And Expectorant	R07AX	Other Respirato
R05CB06	Ambroxol	R06AA04	Clemastine	S01AA	Antibiotics
R05CB07	Sobrerol	R06AA07	Diphenylpyraline	S01AA01	Chloramphenicol
R05CB09	Letosteine	R06AB04	Chlorphenamine	S01AA13	Fusidic Acid
R05CB10	Combinations	R06AD07	Mequitazine	S01AA22	Micronomicin
R05CB14	Neltenexine	R06AE	Piperazine	S01AA30	Combinations Of Different Antibiotics
R05CB15	Erdosteine	R06AE06	Oxatomide	S01AE06	Gatifloxacin
R05DA04	Codeine	R06AE07	Cetirizine	S01B	Antiinflammatory Agents
R05DA09	Dextromethorphan	R06AE09	Levocetirizine	S01BA04	Prednisolone
R05DA20	Combinations	R06AX	Other Antihistamines For Systemic Use	S01BA07	Fluorometholone
R05DB01	Benzonatate	R06AX12	Terfenadine	S01BA13	Rimexolone
R05DB02	Benproperine	R06AX13	Loratadine	S01BC04	Flurbiprofen

Table 2. Continued

ATC Code	Name	ATC Code	Name	ATC Code	Name
S01CA	Corticosteroids And Antiinfectives In Combination	S01EE04	Travoprost	S01XA20	Artificial Tears And Other Indiffernet Preparations
S01CA01	Dexamethasone And Antiinfectives	S01EE05	Tafluprost	S02CA03	Hydrocortisone And Antiinfectives
S01CA07	Fluorometholone And Antiinfectives	S01FA04	Cyclopentolate	S02CA05	Fluocinolone Acetonide And Antiinfectives
S01EA05	Brimonidine	S01FA05	Homatropine	S02CA06	Dexamethasone And Antiinfectives
S01EC01	Acetazolamide	S01FA56	Tropicamide, Combinations	V03AE	Drugs For Treatment of Hyperkalemia And Hyperphosphatemia
S01EC04	Brinzolamide	S01GX	Other Antiallergics	V03AE01	Polystyrene Sulfonate
S01EC05	Methazolamide	S01GX01	Cromoglicic Acid	V03AF03	Calcium Folinate
S01ED01	Timolol	S01KA02	Hypromellose	V06DB	Fat/Carbohydrates/Proteins/Minerals/Vitamins, Combinations
S01ED03	Levobunolol	S01L	Ocular Vascular	V06DX	Other Combinations Of Nutrients
S01ED51	Timolol, Combinations	S01XA	Other Ophthalmologicals	V08AA01	Diatrizoic Acid
S01EE01	Latanoprost	S01XA18	Ciclosporin		

3.3 통계분석

3.3.1 다중검정의 문제 해결

GWAS는 특별한 가설이 없이 수많은 SNP과 질병과의 관계를 검증하기 때문에 항상 다중검정의 문제가 발생한다. 예를 들어 유의수준 0.05로 가설검정을 n회 시행할 경우 적어도 1회 이상 위양성이 발생할 확률은 $1-(0.95)^n$으로 가설검정 시행 횟수가 증가할수록 위양성이 발생할 확률은 증가한다. [46, 47] MWAS 방법론도 GWAS와 마찬가지로 다중검정의 문제가 발생하는데 본 연구에서도 대장암 환자의 예후와 관련된 약물을 찾기 위해 총 773개의 약물에 대한 검증을 시행하였기 때문에 적어도 1회 이상 위양성이 발생할 확률은 $1-(0.95)^{773}$으로 그 확률이 거의 100% 달한다. 이러한 다중검정에 대한 해결책으로 가장 널리 사용되는 방법은 본페로니 교정법(bonferroni correction)으로 유의수준을 검정 횟수로 나눠서 다중검정을 진행하더라도 유의수준을 일정하게 유지시켜주는 방법이다. 간단하게 시행할 수 있는 장점이 있지만 지나치게 엄격하여 검정력이 떨어진다는 단점이 있다. 다중검정을 시행하여 진행되는 연구가 증가하면서 최근 다중검정의 해결책으로 위발견율(False Discovery Rate, FDR)이 제시되고 있다. FDR은 전체 검정에서 위양성을 일정하게 유지시키는 본페로니 교정법과는 달리 양성이라 판단했을 때 위양성을 일정하게 유지시켜주는 방법으로 의미 있는 가설을 찾아낼 가능성을 높여준다. [48] 본 연구에서는 FDR을 이용하여 다중검정의 문제를 해결하도록 하며, 위발견율은 5%로 하여 분석을 진행하였다.

3.3.2 랜드마크 분석(landmark analysis)

약물 사용은 시간에 따라 변하는 변수(time varying variable)로 약물 사용군으로 정의되기 위해서는 생존이 전제되어야 하는데 이를 고려하지 않고 분석을 시행할 경우 약물의 영향과는 무관하게 약물 사용군의 예후가 좋아 보이는 불멸시간편향(immortal time bias)이 발생한다. [49, 50] 본 연구에서는 이러한 편향을 보정하기 위하여 랜드마크 분석을 적용하여 연구를 진행하였다. [51, 52] 랜드마크 분석이란 생존분석에서 노출변수(exposure)가 시간에 따라 변하는 경우 연구자가 설정한 랜드마크 시점 이전에 사망과 같은 결과값(event)이 발생한 경우를 제외하고 분석하는 방법이다. 즉, 특정 시점 이전에 결과값이 발생한 것은 노출변수의 효과와는 무관하며, 랜드마크 시점 이전에 조기에 결과값이 나타나 시간에 따라 변하는 노출변수가 정의되지 않았다고 보고 분석을 하는 것이다. 구체적으로 본 연구에서는 랜드마크 시점을 1년, 3년으로 설정하기로 하며 랜드마크 시점 이전에 사망한 대장암 환자를 순차적으로 제외하면서 분석을 진행하였다.

3.3.3 생존분석 및 재현성 검정

대장암 예후와 관계된 약물을 찾기 위하여 콕스비례위험모형(cox proportional hazard model)을 이용해 생존분석을 진행하였다. 대장암 진단 시점부터 사망 시점까지를 time-to-event로 정의하였고 사망하지 않은 대상자는 2013년12월31일까지를 time-to-censored로 정의하였다. 먼저 discovery set에서 위발견율을 5%, 랜드마크 시점을 1,3년으로 하여 약물을 찾고 이렇게 선정된 약물들을 다시 validation set에서 콕스비례위험모형을 통해 유의수준 5%, 랜드마크 시점 1,3년으로 검증하였다. 모든 분석은 교란요인 보정을 위해서 공변량(covariate)으로 성별, 연령, 대장암 병기에 대한 대리변수로서 대장암 대상자 선정에서 사용되었던 대장내시경적 수술, 수술, 약물, 방사선 등 치료 종류, 동반질환으로 당뇨와 고혈압을 포함하여 분석을 진행하였다. 동반질환에서 당뇨와 고혈압은 한국표준질병사인분류코드가 각각 E10~14, I10~15을 포함한 경우로 정의하였다. (Figure1) 추가적으로 discovery set에서 MWAS를 통해 발견되지는 않았으나 p값이 0.05 미만이었던 약물을 대상으로 validation set에서 분석을 진행하여 discovery set에서 MWAS를 통해 선정된 약물과의 재현성 차이를 살펴보기로 한다. 재현성의 판단은 discovery set에서 선택된 약물과 선택되지 않은 약물의 validation set에서의 유의성에 차이가 있는지에 따라 판단하며 카이제곱검정을 통해 통계적 유의성을 검증하도록 한다. 모든 분석은 SAS 9.4 버전을 통해 진행하였다.

4. 이전 연구결과(previous study)와의 비교

재현성검정에 대한 확장으로 기존에 메타분석에서 대장암 환자에서 생존과 관계 있는 것으로 보고된 약물인 metformin, statin, aspirin이 본 연구에서도 의미 있는 것으로 나타나는지를 비교하였다. [53-55]

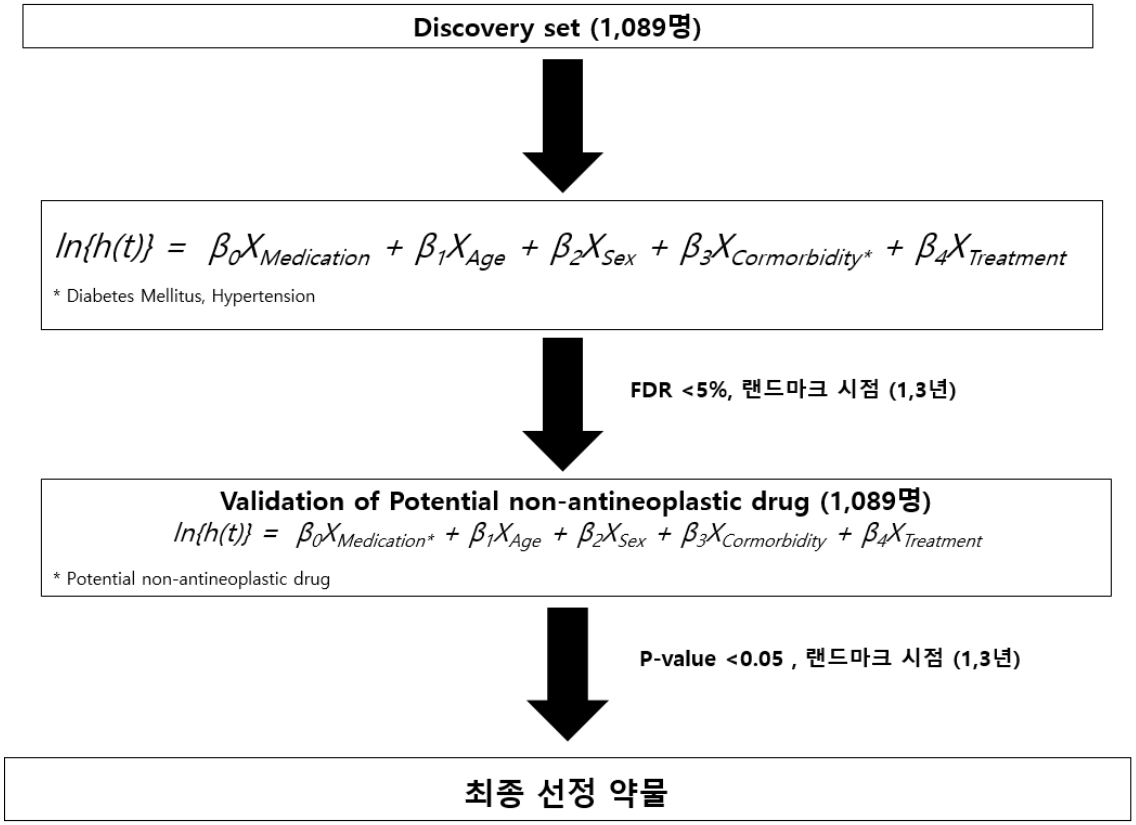

Figure 1. Schematic flow chart for the association analysis of prescribed medicines with colorectal cancer prognosis

III. 결과

1. 연구대상자의 기본적 특징

Table 3은 연구에 포함된 대상자들의 기본적 특징을 제시한 것이다. 추적관찰 기간을 제외하고는 discovery set과 validation set에서 대체적으로 기본적 특징의 분포가 유사하였다. 각각을 살펴보면 먼저 성별 구성비에서는 discovery set, validation set 모두에서 남자 환자의 분율이 각각 63.5%, 63.9%로 여성에 비해 더 많이 포함되었다. 대장암을 최초 진단 받은 나이의 분포를 살펴보면 discovery set, validation set 모두 60-69세까지는 연령이 증가할수록 최초 진단을 받은 환자가 증가하다가 70세 이후부터는 감소하는 동일한 패턴을 보여 주었다. 랜드마크 시점 1,3년에서 사망자의 수는 discovery set과 validation set 모두에서 증가하였고 분율로 비교했을 때에도 유사한 분포를 나타냈다. 대장암 환자가 받은 치료의 분포 역시 유사하였으며 수술만을 받은 환자의 분율이 discovery set, validation set 모두에서 각각 34.3%, 37.7%로 가장 높았다. 동반질환의 분포는 당뇨는 discovery set에서 유병률이 35.8%, validation set에서 36.4%였고 고혈압의 경우 각각 55.6%, 54.0%였다. 대장암 진단을 받은 후 평균 추적관찰기간을 살펴보면 discovery set은 49.9개월(표준편차 38.4개월), validation set은 53.4개월(표준편차 39.8개월)이었으며 표에서는 제시하지 않았으나 이는 통계적으로도 유의한 차이였다. (p값<0.05)

Table 3. General characteristics of the study subjects included in the discovery set and subjects included in the validation set

Variable	Subjects included in the discovery set N= 1,089	Subjects included in the validation set N=1,089
Sex, n (%)		
Male	691 (63.5)	696 (63.9)
Female	398 (36.5)	393 (36.1)
Age at diagnosis of colorectal cancer, n (%)		
15-29	6 (0.6)	8 (0.7)
30-39	31 (2.8)	29 (2.7)
40-49	126 (11.6)	143 (13.1)
50-59	283 (26.0)	255 (23.4)
60-69	340 (31.2)	331 (30.4)
70-79	239 (21.9)	264 (24.2)
>79	64 (5.9)	59 (5.4)
Number of deaths after diagnosis, n (%)		
≤1year	61 (18.2)	69 (22.5)
≤3year	229 (68.2)	195 (63.5)
Total	336 (100)	307 (100)
Received cancer treatment, n (%)		
Colonoscopy only	163 (15.0)	150 (13.8)
Op only	374 (34.3)	411 (37.7)
Op with CTx	84 (7.7)	94 (8.6)
Op with RTx	128 (11.8)	112 (10.3)
Op with Both RTx and CTx	113 (10.4)	101 (9.3)
RTx or CTx without Op	227 (20.8)	221 (20.3)

Table 3. Continued

Variable	Subjects included in the discovery set N= 1,089	Subjects included in the validation set N=1,089
Prevalence of Comorbidity, n (%)		
Diabetes mellitus	390 (35.8)	396 (36.4)
Hypertension	605 (55.6)	588 (54.0)
Follow-up period, months		
Mean±SD	49.9±38.4	53.4±39.8

2. Discovery set에서 MWAS 수행결과

Table 4 와 Table 5 은 FDR 5%, 랜드마크 시점을 1,3 년으로 하여 대장암 예후와 관계가 있는 약물을 제시한 표이다. 분석을 시행한 773 개의 약물 중 랜드마크 시점을 3 년까지로 했을 경우에도 대장암 예후와 관계가 있었던 약물은 29 개였다. 랜드마크 시점을 1 년으로 했을 때 FDR 값은 $6.47x10^{-5}$ 에서 $7.57x10^{-3}$ 사이였으며 117 개의 약물이 대장암 예후와 관계가 있었다. 랜드마크 시점을 3 년까지 늘렸을 때 FDR 값은 $6.61x10^{-5}$ 에서 $1.92x10^{-3}$ 사이였으며 29 개의 약물이 선택되었는데 이 중 랜드마크 시점을 1 년으로 했을 때 포함된 약물이 22 개였고 새롭게 7 개의 약물이 추가되었다. 랜드마크 시점을 1,3 년으로 모두에서 유의했던 약물의 ATC 명칭은 H2-Receptor Antagonists(ATC code: A02BA), Other Drugs For Peptic Ulcer And Gastro-Oesophageal Reflux Disease(ATC code: A02BX), Mosapride(ATC code: A03FA09), Prednicarbate(ATC code: D07AC18), Methylprednisol(ATC code: H02AB04), Fluconazole(ATC code: J02AC01), Megestrol(ATC code: L02AB01), Eperisone(ATC code: M03BX09), Enzymes Used In Musculo-Skeletal System Inflammatory Conditions(ATC code: M09AB), Mometasone(ATC code: R01AD09), Pseudoephedrine(ATC code: R01BA02), Acetylcysteine(ATC code: R05CB01), Erdosteine(ATC code: R05CB15), Levodropropizine(ATC code: R05DB27), Chlorphenamine(ATC code: R06AB04), Levocetirizine(ATC code: R06AE09), Other Antihistamines For Systemic Use(ATC code: R06AX), Other Peripheral vasodilators(ATC code: C04AX), Propionic Acid(ATC code: M01AE), Tramadol And Paracetamol(ATC code: N02AJ13), Opium Derivatives And Expectorant(ATC code: R05FA), Azelastine(ATC code: R06AX19)였다.

Table 4. Drug associated with colorectal cancer prognosis in discovery set at 1 year landmark point

ATC code	Name	P-value	FDR
M09AB	Enzymes (Used In Musculo-Skeletal System Inflammatory Conditions)	$3.04x10^{-25}$	$6.47x10^{-5}$
M01AE	Propionic Acid	$1.51x10^{-23}$	$1.29x10^{-4}$
R06AB04	Chlorphenamine	$8.70x10^{-15}$	$1.94x10^{-4}$
M03BX09	Eperisone	$1.07x10^{-14}$	$2.59x10^{-4}$
A02BX	Other Drugs For Peptic Ulcer And Gastro-Oesophageal Reflux Disease (GORD)	$1.11x10^{-14}$	$3.23x10^{-4}$
R05DB27	Levodropropizine	$2.01x10^{-14}$	$3.88x10^{-4}$
R06AX	Other Antihistamines For Systemic Use	$5.76x10^{-14}$	$4.53x10^{-4}$
R05CB01	Acetylcysteine	$8.77x10^{-14}$	$5.17x10^{-4}$
N02AJ13	Tramadol And Paracetamol	$2.04x10^{-13}$	$5.82x10^{-4}$
R05FA	Opium Derivatives And Expectorant	$2.42x10^{-13}$	$6.47x10^{-4}$
R06AE09	Levocetirizine	$3.35x10^{-13}$	$7.12x10^{-4}$
R01BA02	Pseudoephedrine	$2.49x10^{-12}$	$7.76x10^{-4}$
M01AX	Other Antiinflammatory And Antirheumatic Agents, Non-steroids	$3.17x10^{-12}$	$8.41x10^{-4}$
J01DC04	Cefaclor	$4.91x10^{-12}$	$9.06x10^{-4}$
A02BA	H2-Receptor Antagonists	$5.77x10^{-12}$	$9.70x10^{-4}$
M01AE14	Dexibuprofen	$1.56x10^{-11}$	$1.03x10^{-3}$
H02AB04	Methylprednisolone	$1.85x10^{-11}$	$1.10x10^{-3}$
M01AB16	Aceclofenac	$1.04x10^{-10}$	$1.16x10^{-3}$
A03FA09	Mosapride	$1.67x10^{-10}$	$1.23x10^{-3}$
S01XA20	Artificial Tears And Other Indiffernet Preparations	$1.73x10^{-10}$	$1.29x10^{-3}$
N02BE01	Paracetamol	$2.45x10^{-10}$	$1.36x10^{-3}$
R06AX19	Azelastine	$1.26x10^{-9}$	$1.42x10^{-3}$
A02BA01	Cimetidine	$1.48x10^{-9}$	$1.49x10^{-3}$
J01CR02	Amoxicillin And Beta-Lactamase Inhibitors	$2.33x10^{-9}$	$1.55x10^{-3}$
S01BA07	Fluorometholone	$4.21x10^{-9}$	$1.62x10^{-3}$
A02X	Other Drugs For Acid Related Disordes	$5.17x10^{-9}$	$1.68x10^{-3}$
A03A	Drugs For Functional Gastrointestinal Disorders	$2.98x10^{-8}$	$1.75x10^{-3}$
H02AB06	Prednisolone	$4.97x10^{-8}$	$1.81x10^{-3}$
R05	Cough And Cold Preparations	$7.73x10^{-8}$	$1.88x10^{-3}$
C04AX	Other Peripheral vasodilators	$1.01x10^{-7}$	$1.94x10^{-3}$
A03FA07	Itopride	$1.33x10^{-7}$	$2.01x10^{-3}$
J01MA12	Levofloxacin	$1.89x10^{-7}$	$2.07x10^{-3}$
A02BA02	Ranitidine	$2.54x10^{-7}$	$2.13x10^{-3}$
A02AD03	Almagate	$2.75x10^{-7}$	$2.20x10^{-3}$
M03BX	Other Centrally Acting Agents (Muscle Relaxants)	$2.78x10^{-7}$	$2.26x10^{-3}$
A03FA	Propulsives	$8.02x10^{-7}$	$2.33x10^{-3}$
A03AX58	Alverine, Combinations	$1.39x10^{-6}$	$2.39x10^{-3}$
R05FA02	Opium Derivatives And Expectorant	$1.42x10^{-6}$	$2.46x10^{-3}$
J02AC01	Fluconazole	$1.54x10^{-6}$	$2.52x10^{-3}$
M03BC01	Orphenadrine (Citrate)	$2.15x10^{-6}$	$2.59x10^{-3}$
C10AA05	Atorvastatin	$2.39x10^{-6}$	$2.65x10^{-3}$
R01AD09	Mometasone	$3.99x10^{-6}$	$2.72x10^{-3}$
R05CB15	Erdosteine	$6.49x10^{-6}$	$2.78x10^{-3}$
M01AC06	Meloxicam	$7.63x10^{-6}$	$2.85x10^{-3}$
N05BA01	Diazepam	$8.09x10^{-6}$	$2.91x10^{-3}$
J01FA09	Clarithromycin	$1.16x10^{-5}$	$2.98x10^{-3}$
D07AC18	Prednicarbate	$1.55x10^{-5}$	$3.04x10^{-3}$
D01BA02	Terbinafine	$1.87x10^{-5}$	$3.10x10^{-3}$
J01MA01	Ofloxacin	$2.14x10^{-5}$	$3.17x10^{-3}$
R07AX	Other Respirato	$2.30x10^{-5}$	$3.23x10^{-3}$
A03FA03	Domperidone	$2.38x10^{-5}$	$3.30x10^{-3}$
R05CA12	Hederae Helicis	$2.66x10^{-5}$	$3.36x10^{-3}$
J01CA04	Amoxicillin	$2.88x10^{-5}$	$3.43x10^{-3}$
A03AA05	Trimebutine	$2.88x10^{-5}$	$3.49x10^{-3}$
A04AD	Other Antiemetics	$3.01x10^{-5}$	$3.56x10^{-3}$
A07BC05	Diosmectite	$3.51x10^{-5}$	$3.62x10^{-3}$
R05CA	Expectorants	$3.69x10^{-5}$	$3.69x10^{-3}$
J01FA06	Roxithromycin	$3.92x10^{-5}$	$3.75x10^{-3}$
A03AC05	Tiropramide	$4.29x10^{-5}$	$3.82x10^{-3}$
N05BB01	Hydroxyzine	$6.71x10^{-5}$	$3.88x10^{-3}$
R05CB06	Ambroxol	$6.83x10^{-5}$	$3.95x10^{-3}$
L02AB01	Megestrol	$7.15x10^{-5}$	$4.01x10^{-3}$
R06AA07	Diphenylpyraline	$7.75x10^{-5}$	$4.08x10^{-3}$

Table 4. Continued

ATC code	Name	P-value	FDR
A16AX	Various Alimentary Tract And Metabolism Products	$8.79x10^{-5}$	$4.14x10^{-3}$
S01XA	Other Ophthalmologicals	$1.11x10^{-4}$	$4.20x10^{-3}$
A02BC03	Lansoprazole	$1.54x10^{-4}$	$4.27x10^{-3}$
M01AX21	Diacerein	$1.55x10^{-4}$	$4.33x10^{-3}$
C09AA01	Captopril	$1.84x10^{-4}$	$4.40x10^{-3}$
R03DA07	Theobromine	$1.84x10^{-4}$	$4.46x10^{-3}$
R06AX13	Loratadine	$2.13x10^{-4}$	$4.53x10^{-3}$
D01A	Antifungals For Topical Use	$2.23x10^{-4}$	$4.59x10^{-3}$
J01DB05	Cefadroxil	$2.54x10^{-4}$	$4.66x10^{-3}$
A02BA04	Nizatidine	$2.73x10^{-4}$	$4.72x10^{-3}$
C03CA01	Furosemide	$2.89x10^{-4}$	$4.79x10^{-3}$
M03BX04	Tolperisone	$3.66x10^{-4}$	$4.85x10^{-3}$
G04CA02	Tamsulosin	$6.13x10^{-4}$	$4.92x10^{-3}$
R03DA11	Doxofylline	$6.97x10^{-4}$	$4.98x10^{-3}$
J05AB09	Famciclovir	$7.12x10^{-4}$	$5.05x10^{-3}$
A03	Drugs For Functional Gastrointestinal Disorders	$9.73x10^{-4}$	$5.11x10^{-3}$
M05BA04	Alendronic Acid	$1.03x10^{-3}$	$5.17x10^{-3}$
D01AE15	Terbinafine	$1.25x10^{-3}$	$5.24x10^{-3}$
J01MA02	Ciprofloxacin	$1.32x10^{-3}$	$5.30x10^{-3}$
A10AC01	Insulin (Human)	$1.42x10^{-3}$	$5.37x10^{-3}$
A02BA03	Famotidine	$1.45x10^{-3}$	$5.43x10^{-3}$
A02BA08	Lafutidine	$1.54x10^{-3}$	$5.50x10^{-3}$
M01AB05	Diclofenac	$1.68x10^{-3}$	$5.56x10^{-3}$
D07CC01	Betamethasone And Antibiotics	$1.71x10^{-3}$	$5.63x10^{-3}$
A03AB06	Otilonium Bromide	$1.75x10^{-3}$	$5.69x10^{-3}$
A07FA01	Lactic Acid Producing Organisms	$1.76x10^{-3}$	$5.76x10^{-3}$
A02BC04	Rabeprazole	$2.13x10^{-3}$	$5.82x10^{-3}$
J01GB01	Tobramycin	$2.14x10^{-3}$	$5.89x10^{-3}$
D07AD01	Clobetasol	$2.27x10^{-3}$	$5.95x10^{-3}$
A09AA	Enzyme Preparations	$2.32x10^{-3}$	$6.02x10^{-3}$
N02AD01	Pentazocine	$2.54x10^{-3}$	$6.08x10^{-3}$
R06AD07	Mequitazine	$2.72x10^{-3}$	$6.14x10^{-3}$
A01AD02	Benzydamine	$2.98x10^{-3}$	$6.21x10^{-3}$
J02AC02	Itraconazole	$3.25x10^{-3}$	$6.27x10^{-3}$
R06AX26	Fexofenadine	$3.35x10^{-3}$	$6.34x10^{-3}$
D06AX09	Mupirocin	$3.35x10^{-3}$	$6.40x10^{-3}$
N05BA23	Tofisopam	$3.38x10^{-3}$	$6.47x10^{-3}$
H02AB09	Hydrocortisone	$3.97x10^{-3}$	$6.53x10^{-3}$
D07AC03	Desoximetasone	$4.15x10^{-3}$	$6.60x10^{-3}$
R06AX24	Epinastine	$4.28x10^{-3}$	$6.66x10^{-3}$
J01DC02	Cefuroxime	$4.31x10^{-3}$	$6.73x10^{-3}$
D07AC14	Methylprednisolone aceponate	$4.38x10^{-3}$	$6.79x10^{-3}$
C05BX01	Calcium Dobesilate	$4.40x10^{-3}$	$6.86x10^{-3}$
R06AX22	Ebastine	$4.59x10^{-3}$	$6.92x10^{-3}$
R06AX12	Terfenadine	$5.08x10^{-3}$	$6.99x10^{-3}$
A02BC05	Esomeprazole	$5.22x10^{-3}$	$7.05x10^{-3}$
M01AX22	Morniflumate	$5.36x10^{-3}$	$7.12x10^{-3}$
R03DA	Xanthines	$5.49x10^{-3}$	$7.18x10^{-3}$
R06AE07	Cetirizine	$5.62x10^{-3}$	$7.24x10^{-3}$
N05AX12	Aripiprazole	$6.15x10^{-3}$	$7.31x10^{-3}$
S01ED03	Levobunolol	$6.22x10^{-3}$	$7.37x10^{-3}$
G04BD02	Flavoxate	$6.66x10^{-3}$	$7.44x10^{-3}$
G04BD06	Propiverine	$6.76x10^{-3}$	$7.50x10^{-3}$
A02BA06	Roxatidine	$6.81x10^{-3}$	$7.57x10^{-3}$

Table 5. Drug associated with colorectal cancer prognosis in discovery set at 3 year landmark point

ATC code	Name	P-value	FDR
M01AE	Propionic Acid	$7.91x10^{-9}$	$6.61x10^{-5}$
R06AB04	Chlorphenamine	$1.60x10^{-7}$	$1.32x10^{-4}$
N01BB02	Lidocaine	$2.04x10^{-7}$	$1.98x10^{-4}$
R06AE09	Levocetirizine	$3.46x10^{-7}$	$2.65x10^{-4}$
R05FA	Opium Derivatives And Expectorant	$4.80x10^{-7}$	$3.31x10^{-4}$
A06AD15	Macrogol	$1.05x10^{-6}$	$3.97x10^{-4}$
A02BA	H2-Receptor Antagonists	$1.25x10^{-6}$	$4.63x10^{-4}$
R06AX	Other Antihistamines For Systemic Use	$1.96x10^{-6}$	$5.29x10^{-4}$
R05DB27	Levodropropizine	$2.22x10^{-6}$	$5.95x10^{-4}$
R01BA02	Pseudoephedrine	$2.98x10^{-6}$	$6.61x10^{-4}$
N02AJ13	Tramadol And Paracetamol	$8.13x10^{-6}$	$7.28x10^{-4}$
L02AB01	Megestrol	$2.43x10^{-5}$	$7.94x10^{-4}$
R05CB01	Acetylcysteine	$2.55x10^{-5}$	$8.60x10^{-4}$
H02AB04	Methylprednisolone	$3.95x10^{-5}$	$9.26x10^{-4}$
M03BX09	Eperisone	$6.27x10^{-5}$	$9.92x10^{-4}$
J01MA08	Fleroxacin	$1.94x10^{-4}$	$1.06x10^{-3}$
C04AX	Other Peripheral vasodilators	$2.03x10^{-4}$	$1.12x10^{-3}$
A02BX	Other Drugs For Peptic Ulcer And Gastro-Oesophageal Reflux Disease (GORD)	$2.76x10^{-4}$	$1.19x10^{-3}$
R06AX19	Azelastine	$2.86x10^{-4}$	$1.26x10^{-3}$
A03FA09	Mosapride	$3.39x10^{-4}$	$1.32x10^{-3}$
M09AB	Enzymes (Used In Musculo-Skeletal System Inflammatory Conditions)	$4.28x10^{-4}$	$1.39x10^{-3}$
N02AA05	Oxycodone	$5.28x10^{-4}$	$1.46x10^{-3}$
J02AC01	Fluconazole	$1.29x10^{-3}$	$1.52x10^{-3}$
J01AA20	Combinations Of Tetracyclines	$1.35x10^{-3}$	$1.59x10^{-3}$
N02AB03	Fentanyl	$1.40x10^{-3}$	$1.65x10^{-3}$
V03AE	Drugs For Treatment of Hyperkalemia And Hyperphosphatemia	$1.42x10^{-3}$	$1.72x10^{-3}$
R01AD09	Mometasone	$1.62x10^{-3}$	$1.79x10^{-3}$
D07AC18	Prednicarbate	$1.77x10^{-3}$	$1.85x10^{-3}$
R05CB15	Erdosteine	$1.82x10^{-3}$	$1.92x10^{-3}$

3. Validation set에서 MWAS 수행결과

Discovery set 에서 랜드마크 시점을 1 년으로 했을 때 발견된 약물 117 개 및 3 년 시점에서 새롭게 발견된 7 개의 약물을 더해 총 124 개의 약물에 대해 validation set 에서 검증을 진행하였다. discovery set 에서 선정된 124 개의 약물 중 validation set 에서 랜드마크 시점 1 년, 유의수준 0.05 하에서 유의한 약물은 총 95 개였으며 discovery set 에서 랜드마크 시점 1,3 년 모두에서 발견된 약물이 discovery set 에서 시행된 내적타당도 평가에서 재현이 더 잘되는 경향을 보였다. 세부 결과를 살펴보면 discovery set 에서 랜드마크 시점 1,3 년 모두에서 유의했던 약물 22 개는 validation set 에서 랜드마크 시점을 1 년으로 했을 때 21 개가 유의하였고(21/22=0.95), discovery set 에서 랜드마크 시점 3 년에서만 유의했던 약물 7 개의 경우 validation set 에서 랜드마크 시점을 1 년으로 했을 때 1 개만 유의했으며(1/7=0.14), discovery set 에서 랜드마크 시점 1 년에만 유의했던 약물 95 개는 validation set 에서 랜드마크 시점을 1 년으로 했을 때 72 개가 유의하였다. (72/95=0.76) 동일한 분석을 validation set 의 랜드마크 시점을 3 년으로 하여 시행하였을 때 유의한 약물은 61 개였으며 discovery set 에서 랜드마크 시점 모두에서 유의했던 약물 중 17 개(17/22=0.77), 랜드마크 시점 3 년에서만 유의했던 약물 중 2 개(2/7=0.29), 랜드마크 시점 1 년에서만 유의했던 약물 중 42 개가 유의하였다. (42/95=0.44) Discovery set 과 validation set 의 모든 랜드마크 시점에서 유의했던 약물 17 개 중 16 개는 대장암 환자의 사망위험도를 감소시키는 방향으로 작용하였고 해당 약물의 ATC 명칭은 H2-Receptor Antagonists(ATC code: A02BA), Other Drugs For Peptic Ulcer And Gastro-Oesophageal Reflux Disease (GORD)(ATC code: A02BX), Mosapride(ATC code: A03FA09), Prednicarbate(ATC code: D07AC18), Methylprednisolone(ATC code: H02AB04), Fluconazole(ATC code: J02AC01), Eperisone(ATC code: M03BX09), Enzymes (Used In Musculo-Skeletal System Inflammatory Conditions)(ATC code: M09AB), Mometasone(ATC code: R01AD09), Pseudoephedrine(ATC code: R01BA02), Acetylcysteine(ATC code: R05CB01), Erdosteine(ATC code: R05CB15), Levodropropizine(ATC code: R05DB27), Chlorphenamine(ATC code: R06AB04), Levocetirizine(ATC code: R06AE09), Other Antihistamines For Systemic Use(ATC code: R06AX)였다. 대장암 환자의 사망위험도를 증가시키는 방향으로 작용한 약물은 Megestrol(ATC code: L02AB01)였다. (Table 6)

Table 6. Hazard ratio in the validation set based on significance in the discovery set at landmark point

ATC Code (name)	Discovery set		Validation set	
	1 year landmark point	3 year landmark point	1 year landmark point Hazard ratio (95% CI)	3 year landmark point Hazard ratio (95% CI)
A02BA(H2-Receptor Antagonists)	S	S	0.31 (0.22-0.45)	0.36 (0.23-0.58)
A02BX(Other Drugs For Peptic Ulcer And Gastro-Oesophageal Reflux Disease (GORD))	S	S	0.37 (0.29-0.49)	0.4 (0.27-0.6)
A03FA09(Mosapride)	S	S	0.27 (0.21-0.37)	0.3 (0.2-0.44)
D07AC18(Prednicarbate)	S	S	0.4 (0.25-0.65)	0.46 (0.25-0.84)
H02AB04(Methylprednisolone)	S	S	0.4 (0.28-0.57)	0.6 (0.38-0.93)
J02AC01(Fluconazole)	S	S	0.36 (0.22-0.62)	0.42 (0.22-0.82)
L02AB01(Megestrol)	S	S	1.85 (1.36-2.52)	2.72 (1.74-4.26)
M03BX09(Eperisone)	S	S	0.38 (0.27-0.52)	0.47 (0.31-0.7)
M09AB(Enzymes (Used In Musculo-Skeletal System Inflammatory Conditions))	S	S	0.27 (0.2-0.36)	0.4 (0.25-0.64)
R01AD09(Mometasone)	S	S	0.31 (0.17-0.56)	0.44 (0.22-0.87)
R01BA02(Pseudoephedrine)	S	S	0.4 (0.29-0.54)	0.52 (0.35-0.78)
R05CB01(Acetylcysteine)	S	S	0.44 (0.33-0.59)	0.53 (0.36-0.78)
R05CB15(Erdosteine)	S	S	0.42 (0.27-0.64)	0.45 (0.26-0.78)
R05DB27(Levodropropizin)	S	S	0.28 (0.2-0.39)	0.35 (0.23-0.54)
R06AB04(Chlorphenamine)	S	S	0.39 (0.29-0.52)	0.58 (0.4-0.85)
R06AE09(Levocetirizine)	S	S	0.27 (0.18-0.41)	0.35 (0.21-0.58)
R06AX(Other Antihistamines For Systemic Use)	S	S	0.26 (0.19-0.38)	0.38 (0.25-0.59)
C04AX(Other Peripheral vasodilators)	S	S	0.7 (0.47-1.05)	0.78 (0.46-1.31)
M01AE(Propionic Acid)	S	S	0.46 (0.35-0.6)	0.71 (0.47-1.06)
N02AJ13(Tramadol And Paracetamol)	S	S	0.48 (0.36-0.65)	0.71 (0.48-1.05)
R05FA(Opium Derivatives And Expectorant)	S	S	0.45 (0.34-0.58)	0.74 (0.49-1.12)

Table 6. Continued

ATC Code (name)	Discovery set		Validation set	
	1 year landmark point	3 year landmark point	1 year landmark point Hazard ratio (95% CI)	3 year landmark point Hazard ratio (95% CI)
R06AX19(Azelastine)	S	S	0.46 (0.32-0.66)	0.73 (0.47-1.12)
N02AA05(Oxycodone)	NS	S	2.59 (1.27-5.29)	5.85 (2.31-14.82)
J01AA20(Combinations Of Tetracyclines)	NS	S	3.09 (0.43-22.54)	4.27 (0.57-31.92)
J01MA08(Fleroxacin)	NS	S	1.93 (0.47-7.87)	4.32 (1.02-18.24)
N01BB02(Lidocaine)	NS	S	1.31 (0.42-4.1)	1.34 (0.19-9.68)
N02AB03(Fentanyl)	NS	S	1.8 (0.79-4.11)	2.45 (0.86-6.93)
A06AD15(Macrogol)	NS	S	-	-
V03AE(Drugs For Treatment of Hyperkalemia And Hyperphosphatemia)	NS	S	-	-
A02AD03(Almagate)	S	NS	0.42 (0.32-0.55)	0.61 (0.42-0.91)
A02BA01(Cimetidine)	S	NS	0.45 (0.35-0.59)	0.61 (0.4-0.94)
A02BA04(Nizatidine)	S	NS	0.43 (0.27-0.7)	0.45 (0.24-0.85)
A02BA08(Lafutidine)	S	NS	0.17 (0.05-0.52)	0.22 (0.05-0.89)
A02BC04(Rabeprazole)	S	NS	0.45 (0.28-0.75)	0.34 (0.16-0.73)
A02X(Other Drugs For Acid Related Disordes)	S	NS	0.5 (0.38-0.65)	0.51 (0.35-0.75)
A03AA05(Trimebutine)	S	NS	0.41 (0.31-0.56)	0.48 (0.32-0.71)
A03AC05(Tiropramide)	S	NS	0.56 (0.42-0.73)	0.67 (0.45-0.98)
A03AX58(Alverine, Combinations)	S	NS	0.39 (0.25-0.61)	0.3 (0.15-0.57)
A03FA03(Domperidone)	S	NS	0.56 (0.42-0.74)	0.63 (0.42-0.94)
A03FA07(Itopride)	S	NS	0.48 (0.36-0.64)	0.47 (0.32-0.7)
A04AD(Other Antiemetics)	S	NS	0.5 (0.33-0.76)	0.58 (0.35-0.97)
A09AA(Enzyme Preparations)	S	NS	0.5 (0.29-0.88)	0.41 (0.19-0.88)

Table 6. Continued

ATC Code (name)	Discovery set		Validation set	
	1 year landmark point	3 year landmark point	1 year landmark point Hazard ratio (95% CI)	3 year landmark point Hazard ratio (95% CI)
A16AX(Various Alimentary Tract And Metabolism Products)	S	NS	0.14 (0.05-0.38)	0.19 (0.06-0.61)
C03CA01(Furosemide)	S	NS	1.51 (1.04-2.18)	1.89 (1.15-3.1)
D01BA02(Terbinafine)	S	NS	0.23 (0.1-0.51)	0.23 (0.09-0.64)
D07AD01(Clobetasol)	S	NS	0.39 (0.25-0.62)	0.55 (0.32-0.95)
H02AB06(Prednisolone)	S	NS	0.53 (0.4-0.7)	0.64 (0.43-0.94)
J01CA04(Amoxicillin)	S	NS	0.42 (0.3-0.59)	0.59 (0.39-0.9)
J01CR02(Amoxicillin And Beta-Lactamase Inhibitors)	S	NS	0.42 (0.32-0.56)	0.48 (0.32-0.7)
J01DC04(Cefaclor)	S	NS	0.38 (0.29-0.5)	0.65 (0.44-0.95)
J01FA06(Roxithromycin)	S	NS	0.46 (0.3-0.7)	0.53 (0.31-0.91)
J01FA09(Clarithromycin)	S	NS	0.38 (0.23-0.64)	0.52 (0.28-0.96)
J01GB01(Tobramycin)	S	NS	0.38 (0.22-0.66)	0.38 (0.18-0.78)
J01MA12(Levofloxacin)	S	NS	0.49 (0.37-0.67)	0.59 (0.4-0.88)
J02AC02(Itraconazole)	S	NS	0.4 (0.21-0.79)	0.39 (0.16-0.95)
M01AB16(Aceclofenac)	S	NS	0.43 (0.33-0.57)	0.58 (0.4-0.86)
M01AE14(Dexibuprofen)	S	NS	0.32 (0.24-0.44)	0.37 (0.25-0.56)
M01AX(Other Antiinflammatory And Antirheumatic Agents, Non-steroids)	S	NS	0.36 (0.27-0.48)	0.47 (0.32-0.69)
M01AX21(Diacerein)	S	NS	0.42 (0.24-0.73)	0.52 (0.28-0.97)
M03BX(Other Centrally Acting Agents (Muscle Relaxants))	S	NS	0.39 (0.28-0.56)	0.5 (0.32-0.77)
N05BA01(Diazepam)	S	NS	0.58 (0.44-0.77)	0.59 (0.4-0.88)
N05BB01(Hydroxyzine)	S	NS	0.51 (0.36-0.71)	0.54 (0.35-0.84)
R05CA(Expectorants)	S	NS	0.25 (0.13-0.49)	0.42 (0.2-0.88)

Table 6. Continued

ATC Code (name)	Discovery set		Validation set	
	1 year landmark point	3 year landmark point	1 year landmark point	3 year landmark point
			Hazard ratio (95% CI)	Hazard ratio (95% CI)
R05FA02(Opium Derivatives And Expectorant)	S	NS	0.43 (0.3-0.62)	0.59 (0.37-0.92)
R06AA07(Diphenylpyralin)	S	NS	0.46 (0.32-0.67)	0.63 (0.4-0.98)
R06AD07(Mequitazine)	S	NS	0.47 (0.31-0.73)	0.56 (0.33-0.96)
R06AE07(Cetirizine)	S	NS	0.44 (0.3-0.66)	0.51 (0.31-0.84)
R07AX(Other Respirato)	S	NS	0.29 (0.16-0.55)	0.37 (0.16-0.83)
S01BA07(Fluorometholone)	S	NS	0.39 (0.28-0.54)	0.49 (0.32-0.75)
S01XA20(Artificial Tears And Other Indiffernet Preparations)	S	NS	0.44 (0.32-0.59)	0.47 (0.31-0.71)
A02BA02(Ranitidine)	S	NS	0.67 (0.51-0.87)	0.88 (0.6-1.29)
A02BC03(Lansoprazole)	S	NS	0.38 (0.22-0.65)	0.55 (0.29-1.03)
A02BC05(Esomeprazole)	S	NS	0.38 (0.19-0.78)	0.44 (0.18-1.08)
A03(Drugs For Functional Gastrointestinal Disorders)	S	NS	0.64 (0.48-0.85)	0.82 (0.54-1.24)
A03A(Drugs For Functional Gastrointestinal Disorders)	S	NS	0.42 (0.3-0.59)	0.69 (0.45-1.04)
A03AB06(Otilonium Bromide)	S	NS	0.64 (0.43-0.95)	0.77 (0.47-1.26)
A03FA(Propulsives)	S	NS	0.61 (0.47-0.8)	0.83 (0.55-1.25)
A07FA01(Lactic Acid Producing Organisms)	S	NS	0.69 (0.53-0.91)	0.89 (0.58-1.39)
C10AA05(Atorvastatin)	S	NS	0.49 (0.3-0.79)	0.7 (0.39-1.25)
D01A(Antifungals For Topical Use)	S	NS	0.37 (0.18-0.75)	0.56 (0.26-1.21)
D01AE15(Terbinafine)	S	NS	0.3 (0.11-0.82)	0.49 (0.18-1.33)
D06AX09(Mupirocin)	S	NS	0.5 (0.33-0.77)	0.64 (0.38-1.07)
D07AC03(Desoximetasone)	S	NS	0.6 (0.36-0.98)	0.78 (0.44-1.4)
D07CC01(Betamethasone And Antibiotics)	S	NS	0.51 (0.32-0.81)	0.62 (0.35-1.09)

Table 6. Continued

ATC Code (name)	Discovery set		Validation set	
	1 year landmark point	3 year landmark point	1 year landmark point Hazard ratio (95% CI)	3 year landmark point Hazard ratio (95% CI)
G04CA02(Tamsulosin)	S	NS	0.7 (0.5-0.99)	0.76 (0.48-1.21)
H02AB09(Hydrocortisone)	S	NS	0.53 (0.34-0.82)	0.6 (0.34-1.06)
J01DB05(Cefadroxil)	S	NS	0.63 (0.41-0.97)	0.95 (0.58-1.55)
J01MA01(Ofloxacin)	S	NS	0.62 (0.47-0.82)	0.88 (0.6-1.29)
M01AC06(Meloxicam)	S	NS	0.46 (0.32-0.66)	0.65 (0.42-1.03)
M01AX22(Morniflumate)	S	NS	0.49 (0.27-0.9)	0.63 (0.3-1.3)
M03BC01(Orphenadrine (Citrate))	S	NS	0.54 (0.36-0.83)	0.54 (0.31-0.93)
M05BA04(Alendronic Acid)	S	NS	0.27 (0.11-0.66)	0.41 (0.16-1.03)
N02BE01(Paracetamol)	S	NS	0.52 (0.39-0.69)	0.65 (0.4-1.08)
N05BA23(Tofisopam)	S	NS	0.63 (0.42-0.95)	0.86 (0.53-1.39)
R03DA(Xanthines)	S	NS	0.34 (0.18-0.62)	0.55 (0.28-1.06)
R05(Cough And Cold Preparations)	S	NS	0.58 (0.42-0.82)	0.67 (0.44-1.04)
R05CA12(Hederae Helicis)	S	NS	0.59 (0.43-0.81)	0.73 (0.49-1.1)
R05CB06(Ambroxol)	S	NS	0.5 (0.35-0.7)	0.68 (0.44-1.04)
R06AX13(Loratadine)	S	NS	0.55 (0.33-0.91)	0.7 (0.38-1.29)
R06AX24(Epinastine)	S	NS	0.16 (0.04-0.65)	0.32 (0.08-1.31)
R06AX26(Fexofenadine)	S	NS	0.47 (0.3-0.75)	0.64 (0.37-1.12)
S01XA(Other Ophthalmologicals)	S	NS	0.55 (0.37-0.81)	0.73 (0.45-1.16)
A01AD02(Benzydamine)	S	NS	0.74 (0.46-1.19)	0.85 (0.46-1.55)
A02BA03(Famotidine)	S	NS	0.73 (0.53-1.02)	0.95 (0.62-1.45)
A02BA06(Roxatidine)	S	NS	0.53 (0.22-1.3)	0.98 (0.39-2.45)

Table 6. Continued

ATC Code (name)	Discovery set		Validation set	
	1 year landmark point	3 year landmark point	1 year landmark point Hazard ratio (95% CI)	3 year landmark point Hazard ratio (95% CI)
A07BC05(Diosmectite)	S	NS	0.8 (0.6-1.07)	0.98 (0.66-1.46)
C05BX01(Calcium Dobesilate)	S	NS	1.02 (0.42-2.48)	1.12 (0.35-3.56)
D07AC14(Methylprednisolone aceponate)	S	NS	0.47 (0.21-1.07)	0.78 (0.31-1.95)
G04BD02(Flavoxate)	S	NS	0.39 (0.15-1.06)	0.32 (0.08-1.32)
G04BD06(Propiverine)	S	NS	0.99 (0.63-1.56)	0.83 (0.45-1.54)
J01DC02(Cefuroxime)	S	NS	0.77 (0.49-1.21)	1.01 (0.57-1.78)
J01MA02(Ciprofloxacin)	S	NS	0.8 (0.58-1.09)	1.19 (0.8-1.77)
J05AB09(Famciclovir)	S	NS	0.62 (0.36-1.07)	0.81 (0.42-1.56)
M01AB05(Diclofenac)	S	NS	0.7 (0.46-1.09)	0.91 (0.55-1.53)
M03BX04(Tolperisone)	S	NS	0.61 (0.35-1.05)	0.89 (0.48-1.66)
N02AD01(Pentazocine)	S	NS	1.16 (0.16-8.33)	1.93 (0.26-14.23)
R03DA07(Theobromine)	S	NS	0.66 (0.4-1.1)	0.74 (0.37-1.47)
R03DA11(Doxofylline)	S	NS	0.61 (0.36-1.03)	0.56 (0.27-1.18)
R06AX12(Terfenadine)	S	NS	0.79 (0.32-1.94)	0.6 (0.19-1.96)
R06AX22(Ebastine)	S	NS	0.67 (0.39-1.16)	1.19 (0.64-2.21)
S01ED03(Levobunolol)	S	NS	2.38 (0.33-17.33)	2.88 (0.39-21.4)
C09AA01(Captopril)	S	NS	3.35 (0.46-24.23)	-
A10AC01(Insulin (Human))	S	NS	-	-
N05AX12(Aripiprazole)	S	NS	-	-

(*N.S. = not significant, S=significant; significance indicated by results of discovery set analysis)

4. 재현성에 대한 분석

본 연구에서는 discovery set에서 MWAS를 통해 선정된 약물이 validation set에서 시행된 내적타당도 평가에서도 잘 재현 되는지를 판단하기 위해 분석에 포함된 773개의 약물에서 MWAS를 통해 선정된 약물 124개를 제외한 649개의 약물 중 랜드마크 시점 1,3년 시점에서 p값은 0.05 미만이었지만 선정되지 않았던 약물을 validation set에서 분석에 포함하여 비교하였다. 분석결과 discovery set에서 랜드마크 시점을 1년으로 하였을 때 p값이 0.05보다는 작았지만 FDR 기준에는 해당하지 않아 선정되지 않았던 약물은 59개였으며 validation set에서 랜드마크 시점을 각각 1,3년으로 했을 때 유의한 결과를 보인 약물은 각각 23개, 8개였다. 동일한 분석을 discovery set 랜드마크 시점을 3년으로 하여 시행하였을 때 선정되지 않았던 약물은 70개였으며 validation set에서 랜드마크 시점을 각각 1,3년으로 했을 때 유의한 결과를 보인 약물은 각각 45개, 29개였다. (Table 7)

Table 7. Number of significant drugs in relation to the selection at the landmark point in discovery set

Discovery set		1 year landmark point			3 year landmark point		
		S**, n (%)	N.S.**, n (%)	p-value***	S, n (%)	N.S., n (%)	p-value
1 year landmark point	Selected (n=117)	94 (80.3)	23 (19.7)	<0.01	59 (50.4)	58 (49.6)	<0.01
	Not Selected*(n=59)	23 (39.0)	36 (61.0)		8 (13.6)	51 (86.4)	
3 year landmark point	Selected (n=29)	22 (75.9)	7 (24.1)	0.26	19 (65.5)	10 (34.5)	0.03
	Not Selected (n=70)	45 (64.3)	25 (35.7)		29 (41.4)	41 (58.6)	

(*p-value is less than 0.05 in discovery set analysis, but not selected)
(** S=significant N.S. = not significant,; significance indicated by results of validation set analysis.)
(*** p-values were calculated by chi-square test)

5. MWAS 결과와 기존에 대장암 예후와 관련이 있었던 약물과의 비교

기존에 대장암 예후와 관계가 있다고 보고된 약물인 metformin, statin, aspirin이 MWAS 분석에서도 유의한 관계가 있는지를 살펴보았다. 분석에 포함된 약물 중 metformin, statin, aspirin에 해당하는 약물은 각각 6개, 10개, 2개 였으며 각각의 약물을 하나로 묶은 것을 포함한 총 21개에 대해서 분석을 진행하였다. Discovery set에서 위발견율을 기준으로 유의한 약물은 Atorvastatin(ATC code: C10AA05)이 유일하였으며 metformin과 aspirin의 경우 discovery set에서 위발견율을 기준으로 통계적으로 유의한 관련성을 보이는 약물은 없었다. 하지만 일부 약물에서 유의수준 0.05하에서는 통계적으로 유의한 관련성을 보였으며 통계적으로 유의하지 않은 경우에도 위험비(hazard ratio)의 방향은 선행연구 결과와 대체로 유사하였다. (Table 8)

Table 8. Comparison between the result of MWAS findings and candidate colorectal cancer prognosis associated drugs in meta-analyses

Candidate drugs	Previous Meta-analysis Result	ATC code	Discovery set (1 year landmark) Hazard Ratio (95% CI)	Discovery set (3 year landmark) Hazard Ratio (95% CI)	Validation set (1 year landmark) Hazard Ratio (95% CI)	Validation set (3 year landmark) Hazard Ratio (95% CI)
Metformin	Overall survival (HR 0.6, 95% CI: 0.53–0.67)	Combined	0.71 (0.46-1.09)	0.52 (0.28-0.95)	0.94 (0.60-1.48)	1.38 (0.76-2.51)
		A10BA02	0.69 (0.44-1.07)	0.49 (0.26-0.92)	0.86 (0.54-1.36)	1.19 (0.65-2.19)
		A10BD02	0.76 (0.37-1.58)	0.28 (0.07-1.16)	0.58 (0.23-1.44)	0.54 (0.17-1.76)
		A10BD03	1.11 (0.15-8.07)	1.36 (0.18-10.07)	-	-
		A10BD07	-	-	0.44 (0.06-3.15)	0.68 (0.09-4.90)
		A10BD08	0.48 (0.07-3.44)	1.02 (0.14-7.53)	1.36 (0.19-9.81)	4.56 (0.62-33.53)
		A10BD11	-	-	-	-
Statin	Overall survival (HR 0.86, 95% CI: 0.76–0.98)	Combined	0.36 (0.24-0.55)*	0.45 (0.26-0.80)	0.44 (0.30-0.67)*	0.62 (0.37-1.03)
		C10AA01	0.41 (0.17-1.01)	0.55 (0.20-1.52)	0.41 (0.18-0.93)*	0.23 (0.06-0.93)*
		C10AA02	1.18 (0.29-4.80)	2.56 (0.62-10.59)	0.38 (0.05-2.76)	0.50 (0.07-3.68)
		C10AA03	0.30 (0.04-2.14)	0.86 (0.12-6.24)	1.32 (0.42-4.21)	1.82 (0.56-5.93)
		C10AA04	1.22 (0.30-4.94)	2.67 (0.65-10.95)	-	-
		C10AA05	0.31 (0.19-0.51)*	0.37 (0.19-0.73)	0.49 (0.30-0.79)*	0.70 (0.39-1.25)
		C10AA07	0.71 (0.31-1.60)	0.63 (0.15-2.61)	0.62 (0.25-1.52)	1.02 (0.37-2.81)
		C10AA08	0.29 (0.04-2.10)	-	-	-
		C10BA02	-	-	0.26 (0.04-1.85)	0.59 (0.08-4.29)
		C10BA03	-	-	-	-
		C10BX03	0.36 (0.05-2.61)	0.68 (0.09-4.96)	0.35 (0.05-2.48)	1.04 (0.14-7.53)
Aspirin	Overall survival (HR 0.84, 95% CI: 0.75–0.94)	Combined	0.69 (0.51-0.93)	0.72 (0.46-1.14)	0.92(0.66-1.28)	0.99 (0.64-1.56)
		B01AC06	0.68 (0.50-0.94)	0.68 (0.42-1.08)	0.93 (0.66-1.29)	1.11 (0.71-1.73)
		N02BA01	0.89 (0.47-1.68)	0.90 (0.36-2.23)	0.69 (0.32-1.49)	0.44 (0.14-1.40)

(* Significant after false discovery rate correction)

IV. 고찰 및 결론

1. 대장암 예후와 관련된 약물의 해석

관찰연구는 임상시험이 윤리성 또는 비용 등의 문제로 제한될 때에 새로운 약물 또는 중재법이 실제로 적용되었을 때 효과 또는 안전성을 측정함에 있어 유용한 정보를 제공한다. 그러나 이러한 관찰연구에서 직면하게 되는 문제는 재현성의 문제이다. 물론 임상시험도 동일한 문제가 발생할 수 있지만 최근에는 임상시험 등록을 통해 모든 과정을 투명하게 공개함으로써 재현성 확보에 힘쓰고 있다. 본 연구는 우리나라 인구에 대한 대표성을 가지는 국민건강보험공단 표본DB를 통해 MWAS라는 방법을 사용하여 대장암 예후와 관계된 약물을 데이터주도 방식을 통한 체계적 분석으로 선택적 보고를 줄이고자 하였고 추가적으로 대장암 환자에서 예후와 관계가 있는 약물에 대한 새로운 가설을 발견하고자 하였다. 연구 설계 과정에서 랜드마크 시점으로 설정한 1,3년에서 discovery set과 validation set에서 모두 유의했던 약물은 17개였다. 17개의 약물 중 H2-Receptor Antagonists(ATC code: A02BA), Other Drugs For Peptic Ulcer And Gastro-Oesophageal Reflux Disease (GORD)(ATC code: A02BX), Mosapride(ATC code: A03FA09), Prednicarbate(ATC code: D07AC18), Methylprednisolone(ATC code: H02AB04), Fluconazole(ATC code: J02AC01), Eperisone(ATC code: M03BX09), Enzymes (Used In Musculo-Skeletal System Inflammatory Conditions)(ATC code: M09AB), Mometasone(ATC code: R01AD09), Pseudoephedrine(ATC code: R01BA02), Acetylcysteine(ATC code: R05CB01), Erdosteine(ATC code: R05CB15), Levodropropizine(ATC code: R05DB27), Chlorphenamine(ATC code: R06AB04), Levocetirizine(ATC code: R06AE09), Other Antihistamines For Systemic Use(ATC code: R06AX) 등 총 16개의 약물은 대장암 환자의 사망위험도를 감소시키는 방향으로 작용하였고 Megestrol(ATC code: L02AB01)는 대장암 환자의 사망위험도를 증가시키는 방향으로 작용하였다. 발견된 약물 중 인간을 대상으로 한 연구에서 대장암과의 관련성 여부가 보고된 약물은 H2-Receptor Antagonists(ATC code: A02BA), Levocetirizine(ATC code: R06AE09) 등 히스타민 길항제였다. 각각의 결과를 요약하면 H2-Receptor Antagonists의 경우 코크란 리뷰에서 수술을 시행한 대장암 환자에서 H2-Receptro Antagonists를 사용한 경우 생존에 있어 이익이 있는 것으로 나타났으며 Levocetirizine의 경우 임상 2상 연구로서 전이가 된 대장암 환자에서 Capecitabine과 Bevacizumab의 내성을 나타내는 경우 Levocetirizine을 사용하는 경우 생존에 있어 이익이 있다는 연구였다. [56, 57]

2. 재현성에 대한 분석결과 해석

본 연구에서는 대장암 예후와 관련된 약물에 대한 분석에 더해서 내적타당도 평가 결과 비교를 통한 MWAS 방법론의 재현성에 대한 분석을 시행하였다. Table 7에서는 이에 대한 결과를 제시하고 있다. Discovery set에서 랜드마크 시점 1년에 선택된 약물은 선택되지 않은 약물에 비해 validation set에서 랜드마크 시점 1,3년에 모두 통계적으로 유의하게 재현성이 높은 것으로 나타났으며 discovery set에서 랜드마크 시점 3년에서 선택된 약물은 validation set에서 랜드마크 시점 3년에서만 통계적으로 유의하게 재현성이 높은 것으로 나타났다. 이상을 종합하면 MWAS 방법론이 통계적 검정을 수행하는 연구에서 흔히 사용되는 유의수준 0.05를 기준으로 연구를 진행하는 것에 비해 대체적으로 연구의 재현성을 높이는데 도움을 줄 수 있는 것으로 사료된다. 연구에서 재현성이 중요한 이유는 Bradford Hill이 인과성을 평가하기 위해 제시한 기준 중 관련성의 일관성과 맞닿아 있기 때문이다. 즉, 일관되는 관련성이 연구를 통해 보고될 경우 인과관계일 가능성이 높다는 것이다. [58] MWAS 방법론이 어떻게 연구의 재현성을 높이는지를 살펴보기 위하여 기존에 수행된 MWAS 연구들을 고찰하였으나 연구의 재현성 제고보다는 주로 새로운 가설 생성에 중점을 두고 있어 MWAS 방법론이 어떻게 연구의 재현성을 높이는지에 대해서는 서술된 바가 없었다. [10-14] 이에 MWAS와 동일한 과정으로 수행되는 GWAS에서 그 이유를 찾기 위해 문헌을 고찰한 결과 GWAS를 통해 연구 재현성을 높일 수 있는 이유는 다음과 같다. 연구의 재현성을 높이기 위해서는 분석과정에서 연구자의 의도를 배제하고 모든 과정이 표준화된 상태에서 체계적으로 연구를 수행하는 것이 중요하다. GWAS 방법론이 활성화 되기 이전에 수행되었던 단일유전자(single gene approach), 후보유전자 접근법(Candidate gene approach) 등의 연구들에서 연구의 재현성이 높지 않았던 이유는 크게 두 가지로 첫 번째는 다중검정에 대한 부적절한 통제, 두 번째는 선택적 보고에 따른 출판편향이다. 이와는 반대로 GWAS에서는 다중검정을 통제하기 위해 엄격한 통계방법을 적용하였으며 이에 따라 위양성의 가능성이 낮아졌다. 또한 가설 기반하에 수행되는 연구들과는 달리 모든 SNP에 대한 결과를 분석하고 투명하게 보고하여 출판편향의 가능성을 줄였으며 그 결과 연구 재현성이 높아졌다. [32]

3. 기존 연구에서 대장암 환자의 예후와 관련이 있었던 약물과의 비교결과 해석

기존 메타분석 등에서 대장암 환자의 예후와 관계가 있다고 알려진 metformin, statin, aspirin이 MWAS를 이용한 본 연구에서도 유사하게 대장암 환자의 예후와 관계가 있는지를 비교하였다. 위험비의 방향은 대체적으로 선행연구와 유사하였으나 statin 중 Atorvastatin(ATC code: C10AA05) 외에는 discovery set에서 위발견율을 기

준으로 유의한 결과를 나타낸 약물이 없었다. 기존 연구에서 metformin, statin, aspirin 등이 어느정도 일관되게 대장암 환자의 예후와 관계가 있다고 보고되고 있기 때문에 이 관계가 참임을 가정할 때 우리 연구에서 위음성의 결과가 나타난 이유는 2가지의 해석이 가능하다. 첫 번째는 본 연구에서 다중검정을 통제하기 위하여 본페로니 교정법에 비해 검정력이 높은 위발견율을 사용하여 분석을 진행하였음에도 불구하고 기존에 알려진 약물들의 효과값(effect size)이 크지 않을 경우 위음성의 결과가 발생할 수 있다는 것이다. 실제로 metformin, aspirin 중 일부 약물들의 p값은 0.05 미만이었으나 위발견율을 기준으로 했을 때는 유의하지 않은 것으로 나타났다. 두 번째는 연구설계에 있어서 기존에 수행된 가설주도 분석과의 차이로 인한 위음성의 가능성이다. 본 연구의 경우 모든 모델에서 공통된 공변량을 포함하여 한번에 여러가지 약물을 분석하는 방식으로 연구를 진행하여 연구의 표준화 및 체계적 연구 수행이 가능하다는 장점이 있으나 가설주도 분석에서 시행하는 모델 선정, 하위집단(subgroup)분석 등이 이루어지지 않는다는 단점이 있다. 실제로 metformin과 대장암 환자와의 예후를 보는 연구들은 대장암 환자 중 당뇨 환자로 대상군을 제한하거나 metformin에 대한 환자의 순응도(adherence)를 모델에 포함하는 방식으로 연구를 진행하는 경우가 있었다. [7] 또한 aspirin과 대장암 환자와의 예후를 보는 연구의 경우에도 대장암 환자 중 PIK3CA 변이가 있는 환자를 대상으로 하위집단분석을 진행하는 경우가 많았다. [59, 60]

4. 연구의 장점 및 한계

본 연구의 장점은 다음과 같다. 첫 번째 우리나라 국민에 대한 대표성을 가지는 건강보험공단 표본DB를 통해 MWAS의 가능성을 제시하였다. 특히 기존에 수행되었던 MWAS 연구들과는 달리 연구방법론에 있어서 재현성에 중점을 두었고 MWAS 방법론이 전통적인 통계적 검정 수행연구에 비해 내적타당도 평가에 있어서 대체적으로 우위를 보였다. 두 번째는 기존 수행된 MWAS연구와 마찬가지로 새로운 가설의 제시이다. 본 연구에서는 총 17개의 약물에서 대장암 예후와 관계된 실마리정보(signal)를 찾아냈다. 추후에 잘 설계된 연구를 통해 이러한 관련성이 인과적인 관련성인지에 대한 추가 연구가 필요하다. 본 연구의 한계는 다음과 같다. 먼저 본질적으로 관찰 연구가 가지는 한계이다. 특히 MWAS의 경우 한 번에 여러 가설을 검증하기 때문에 연구의 타당도가 떨어질 경우 몇 가지 가설만을 검증하는 가설 기반의 연구에 비해 이러한 문제가 더 크게 작용할 수 있다. 무작위배정을 통해 약물 사용군, 비사용군을 분류하는 임상시험과는 달리 약물 사용군과 비사용군의 기본적 차이로 인해 비교성에 문제가 발생할 수 있다. 특히 특정약물의 적응증이 환자의 예

후와 관련된 경우에는 이 문제는 심각하게 작용할 수 있다. 본 연구에서는 건강보험공단 표본DB에는 병기에 대한 정보가 없어 대장암 환자가 받은 치료의 종류를 공변량으로 하여 이를 보정하였으나 여전히 잔여교란(residual confounding)의 문제가 남아 있을 수 있다. 하지만 대장암 병기의 대리변수(proxy variable)로서 대장암 치료를 사용했던 연구들에서 대장암 치료법이 환자의 병기를 어느 정도 대리할 수 있음을 보고하였으며 실제로 결과에는 제시하지 않았지만 치료 종류에 따라 생존분석을 시행하였을 때 환자의 예후에 유의한 차이가 존재하였다. [7, 44] 두 번째는 본 연구에 사용된 자료는 연구목적 자료가 아니라 건강보험 청구를 위한 실사용 자료의 성격을 가진다는 것이다. 실사용 자료의 이용으로 인해 실제 임상 현실 및 진료과정을 직접적으로 반영한다는 장점을 가지지만 이로 인해 진단명 오분류 등에 따르는 정보편향(information bias)이 존재할 수 있다. 본 연구에서는 이러한 정보편향을 줄이기 위해서 진단명에 더해 치료 정보를 추가하여 이러한 편향을 줄이는데 힘썼다.

5. 결론

건강보험공단 표본DB에서 MWAS 방법론을 이용하여 대장암 환자의 예후에 영향을 주는 약물을 체계적으로 분석하였으며 새롭게 17개의 실마리 정보를 발견하였다. 특히 재현성에 대한 분석결과에서는 MWAS를 통해 발견된 약물이 전통적으로 수행되는 통계적 검정 연구에 비해 내적타당도 평가에 있어서 대체적으로 우위를 가짐을 보여주었다. 본 연구는 MWAS 방법론이 우리나라 국민을 대표하는 건강보험공단 표본DB에서 적용될 수 있는 가능성을 제시하였다.

참고문헌

1. 중앙암등록본부. 국가암등록사업 연례 보고서 (2017년 암등록통계). 2020.
2. Cho S, Shin A, Park SK, Shin HR, Chang SH, Yoo KY. Alcohol Drinking, Cigarette Smoking and Risk of Colorectal Cancer in the Korean Multi-center Cancer Cohort. J Cancer Prev. 2015;20(2):147-52. doi: 10.15430/JCP.2015.20.2.147. PubMed PMID: 26151048; PubMed Central PMCID: PMCPMC4492359.
3. Sturmer T, Buring JE, Lee IM, Gaziano JM, Glynn RJ. Metabolic abnormalities and risk for colorectal cancer in the physicians' health study. Cancer Epidemiol Biomarkers Prev. 2006;15(12):2391-7. doi: 10.1158/1055-9965.EPI-06-0391. PubMed PMID: 17164361; PubMed Central PMCID: PMCPMC1764627.
4. Strom BL, Kimmel SE, Hennessy S. Textbook of pharmacoepidemiology. Chichester, West Sussex, UK ; Hoboken, NJ: Wiley Blackwell,; 2013.
5. Rothwell PM, Wilson M, Elwin CE, Norrving B, Algra A, Warlow CP, et al. Long-term effect of aspirin on colorectal cancer incidence and mortality: 20-year follow-up of five randomised trials. Lancet. 2010;376(9754):1741-50. doi: 10.1016/S0140-6736(10)61543-7. PubMed PMID: 20970847.
6. Voorneveld PW, Reimers MS, Bastiaannet E, Jacobs RJ, van Eijk R, Zanders MMJ, et al. Statin Use After Diagnosis of Colon Cancer and Patient Survival. Gastroenterology. 2017;153(2):470-+. doi: 10.1053/j.gastro.2017.05.011. PubMed PMID: WOS:000406040600026.
7. Choe S, Lee J, Park JW, Jeong SY, Cho YM, Park BJ, et al. Prognosis of Patients with Colorectal Cancer with Diabetes According to Medication Adherence: A Population-Based Cohort Study. Cancer Epidemiol Biomarkers Prev. 2020;29(6):1120-7. doi: 10.1158/1055-9965.EPI-19-1455. PubMed PMID: 32482636.
8. Ioannidis JPA. Why most published research findings are false. Plos Med. 2005;2(8):696-701. doi: ARTN e124
10.1371/journal.pmed.0020124. PubMed PMID: WOS:000231676900008.
9. Ioannidis JPA. Why most discovered true associations are inflated. Epidemiology. 2008;19(5):640-8. doi: 10.1097/EDE.0b013e31818131e7. PubMed PMID: WOS:000258712000001.
10. Harpaz R, DuMouchel W, Shah NH, Madigan D, Ryan P, Friedman C. Novel data-mining methodologies for adverse drug event discovery and analysis. Clin Pharmacol Ther. 2012;91(6):1010-21. doi: 10.1038/clpt.2012.50. PubMed PMID: 22549283; PubMed Central PMCID: PMCPMC3675775.
11. Ryan PB, Madigan D, Stang PE, Schuemie MJ, Hripcsak G. Medication-wide association studies. CPT Pharmacometrics Syst Pharmacol. 2013;2:e76. doi: 10.1038/psp.2013.52. PubMed PMID: 24448022; PubMed Central PMCID: PMCPMC4026636.
12. Ryan PB, Madigan D, Stang PE, Overhage JM, Racoosin JA, Hartzema AG. Empirical assessment of methods for risk identification in healthcare data: results from the experiments of the Observational Medical Outcomes Partnership. Stat Med. 2012;31(30):4401-15. doi: 10.1002/sim.5620. PubMed PMID: 23015364.
13. Pottegard A, Friis S, Christensen R, Habel LA, Gagne JJ, Hallas J. Identification of Associations Between Prescribed Medications and Cancer: A Nationwide Screening Study. EBioMedicine. 2016;7:73-9. doi: 10.1016/j.ebiom.2016.03.018. PubMed PMID: 27322460; PubMed Central PMCID: PMCPMC4909325.
14. Patel CJ, Ji JG, Sundquist J, Ioannidis JPA, Sundquist K. Systematic assessment of pharmaceutical prescriptions in association with cancer risk: a method to conduct a population-wide medication-wide longitudinal study. Sci Rep-Uk. 2016;6. doi: ARTN 31308
10.1038/srep31308. PubMed PMID: WOS:000381107400001.
15. Szklo M NF. Epidemiology: Beyond theBasics, 3rd edn. Burlington,MA; : Jones and Bartlett Learning; 2014.
16. 안윤옥 유, 박병주, 김동현, 배종면, 강대희, 신명희, 이무송. 역학의 원리와 응용.

서울: 서울대학교출판부; 2005.
17. L. G. Epidemiology. Forth edition. . United States of America: : Elsevier; ; 2009.
18. Hennekens CH, DeMets D. Statistical association and causation: contributions of different types of evidence. JAMA. 2011;305(11):1134-5. doi: 10.1001/jama.2011.322. PubMed PMID: 21406653.
19. Rothman KJG, Sander; Lash, Timothy L. Modern Epidemiology. 3rd Ed. ed: Lippincott Williams & Wilkins (LWW); 2008.
20. Greenland S, Senn SJ, Rothman KJ, Carlin JB, Poole C, Goodman SN, et al. Statistical tests, P values, confidence intervals, and power: a guide to misinterpretations. Eur J Epidemiol. 2016;31(4):337-50. doi: 10.1007/s10654-016-0149-3. PubMed PMID: 27209009; PubMed Central PMCID: PMCPMC4877414.
21. Dwan K, Gamble C, Williamson PR, Kirkham JJ, Reporting Bias G. Systematic review of the empirical evidence of study publication bias and outcome reporting bias - an updated review. PLoS One. 2013;8(7):e66844. doi: 10.1371/journal.pone.0066844. PubMed PMID: 23861749; PubMed Central PMCID: PMCPMC3702538.
22. Page MJ, McKenzie JE, Kirkham J, Dwan K, Kramer S, Green S, et al. Bias due to selective inclusion and reporting of outcomes and analyses in systematic reviews of randomised trials of healthcare interventions. Cochrane Db Syst Rev. 2014;(10). doi: ARTN MR000035
10.1002/14651858.MR000035.pub2. PubMed PMID: WOS:000347646000085.
23. You B, Gan HK, Pond G, Chen EX. Consistency in the analysis and reporting of primary end points in oncology randomized controlled trials from registration to publication: a systematic review. J Clin Oncol. 2012;30(2):210-6. doi: 10.1200/JCO.2011.37.0890. PubMed PMID: 22162583.
24. Baker M. Is There a Reproducibility Crisis? Nature. 2016;533(7604):452-4. doi: DOI 10.1038/533452a. PubMed PMID: WOS:000376443100012.
25. von Elm E, Altman DG, Egger M, Pocock SJ, Gotzsche PC, Vandenbroucke JP, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Epidemiology. 2007;18(6):800-4. doi: 10.1097/EDE.0b013e3181577654. PubMed PMID: 18049194.
26. Head ML, Holman L, Lanfear R, Kahn AT, Jennions MD. The extent and consequences of p-hacking in science. PLoS Biol. 2015;13(3):e1002106. doi: 10.1371/journal.pbio.1002106. PubMed PMID: 25768323; PubMed Central PMCID: PMCPMC4359000.
27. De Angelis C, Drazen JM, Frizelle FA, Haug C, Hoey J, Horton R, et al. Clinical trial registration: a statement from the International Committee of Medical Journal Editors. Lancet. 2004;364(9438):911-2. doi: 10.1016/S0140-6736(04)17034-7. PubMed PMID: 15364170.
28. Loder E, Groves T, Macauley D. Registration of observational studies. BMJ. 2010;340:c950. doi: 10.1136/bmj.c950. PubMed PMID: 20167643.
29. Wasserstein RL, Lazar NA. The ASA's Statement on p-Values: Context, Process, and Purpose. Am Stat. 2016;70(2):129-31. doi: 10.1080/00031305.2016.1154108. PubMed PMID: WOS:000378462300001.
30. Ozaki K, Ohnishi Y, Iida A, Sekine A, Yamada R, Tsunoda T, et al. Functional SNPs in the lymphotoxin-alpha gene that are associated with susceptibility to myocardial infarction. Nat Genet. 2002;32(4):650-4. doi: 10.1038/ng1047. PubMed PMID: WOS:000179593000019.
31. Visscher PM, Wray NR, Zhang Q, Sklar P, McCarthy MI, Brown MA, et al. 10 Years of GWAS Discovery: Biology, Function, and Translation. American Journal of Human Genetics. 2017;101(1):5-22. doi: 10.1016/j.ajhg.2017.06.005. PubMed PMID: WOS:000404886800001.
32. Marigorta UM, Rodriguez JA, Gibson G, Navarro A. Replicability and Prediction: Lessons and Challenges from GWAS. Trends Genet. 2018;34(7):504-17. doi: 10.1016/j.tig.2018.03.005. PubMed PMID: WOS:000435027200003.
33. Hirschhorn JN, Lohmueller K, Byrne E, Hirschhorn K. A comprehensive review of genetic association studies. Genet Med. 2002;4(2):45-61. doi: Doi 10.1097/00125817-

200203000-00002. PubMed PMID: WOS:000174373800002.
34. Rakyan VK, Down TA, Balding DJ, Beck S. Epigenome-wide association studies for common human diseases. Nat Rev Genet. 2011;12(8):529-41. doi: 10.1038/nrg3000. PubMed PMID: WOS:000292813400010.
35. Kozlitina J, Smagris E, Stender S, Nordestgaard BG, Zhou HH, Tybjaerg-Hansen A, et al. Exome-wide association study identifies a TM6SF2 variant that confers susceptibility to nonalcoholic fatty liver disease. Nat Genet. 2014;46(4):352-+. doi: 10.1038/ng.2901. PubMed PMID: WOS:000334510100010.
36. Hu H, Pearson TA, Zhao JY, Prosperi M, Rasmussen SA, Hong X, et al. Geographic Disparities in Hypertensive Disorders of Pregnancy: An Environment-Wide Association Study. Circulation. 2019;139. doi: 10.1161/circ.139.suppl_1.P100. PubMed PMID: WOS:000478079000052.
37. Kim JH, Son KY, Shin DW, Kim SH, Yun JW, Shin JH, et al. Network analysis of human diseases using Korean nationwide claims data. J Biomed Inform. 2016;61:276-82. doi: 10.1016/j.jbi.2016.05.002. PubMed PMID: 27179758.
38. Chon MW, Lee J, Chung S, Kim Y, Kim HW. Prescription Pattern of Antidepressants for Children and Adolescents in Korea Based on Nationwide Data. J Korean Med Sci. 2017;32(10):1694-701. doi: 10.3346/jkms.2017.32.10.1694. PubMed PMID: 28875616; PubMed Central PMCID: PMCPMC5592186.
39. Lee H, Myung W, Lee C, Choi J, Kim H, Carroll BJ, et al. Clinical epidemiology of long-term suicide risk in a nationwide population-based cohort study in South Korea. J Psychiatr Res. 2018;100:47-55. doi: 10.1016/j.jpsychires.2018.01.018. PubMed PMID: 29486402.
40. Na EJ, Lee H, Myung W, Fava M, Mischoulon D, Paik JW, et al. Risks of Completed Suicide of Community Individuals with ICD-10 Disorders Across Age Groups: A Nationwide Population-Based Nested Case-Control Study in South Korea. Psychiatry Investig. 2019;16(4):314-24. doi: 10.30773/pi.2019.02.19. PubMed PMID: 31042694; PubMed Central PMCID: PMCPMC6504769.
41. Ko K, Lee CW, Nam S, Ahn SV, Bae JH, Ban CY, et al. Epidemiological Characterization of a Directed and Weighted Disease Network Using Data From a Cohort of One Million Patients: Network Analysis. J Med Internet Res. 2020;22(4):e15196. doi: 10.2196/15196. PubMed PMID: 32271154; PubMed Central PMCID: PMCPMC7180516.
42. Lee J, Lee JS, Park SH, Shin SA, Kim K. Cohort Profile: The National Health Insurance Service-National Sample Cohort (NHIS-NSC), South Korea. Int J Epidemiol. 2017;46(2):e15. doi: 10.1093/ije/dyv319. PubMed PMID: 26822938.
43. Kim JA, Yoon S, Kim LY, Kim DS. Towards Actualizing the Value Potential of Korea Health Insurance Review and Assessment (HIRA) Data as a Resource for Health Research: Strengths, Limitations, Applications, and Strategies for Optimal Use of HIRA Data. J Korean Med Sci. 2017;32(5):718-28. doi: 10.3346/jkms.2017.32.5.718. PubMed PMID: 28378543; PubMed Central PMCID: PMCPMC5383602.
44. Lee J, Choe S, Park JW, Jeong SY, Shin A. The Risk of Colorectal Cancer After Cholecystectomy or Appendectomy: A Population-based Cohort Study in Korea. J Prev Med Public Health. 2018;51(6):281-8. doi: 10.3961/jpmph.18.105. PubMed PMID: 30514058; PubMed Central PMCID: PMCPMC6283741.
45. Kang G, Liu W, Cheng C, Wilson CL, Neale G, Yang JJ, et al. Evaluation of a two-step iterative resampling procedure for internal validation of genome-wide association studies. J Hum Genet. 2015;60(12):729-38. doi: 10.1038/jhg.2015.110. PubMed PMID: 26377241; PubMed Central PMCID: PMCPMC4859941.
46. Streiner DL. Best (but oft-forgotten) practices: the multiple problems of multiplicity-whether and how to correct for many statistical tests. Am J Clin Nutr. 2015;102(4):721-8. doi: 10.3945/ajcn.115.113548. PubMed PMID: 26245806.
47. Johnson RC, Nelson GW, Troyer JL, Lautenberger JA, Kessing BD, Winkler CA, et al. Accounting for multiple comparisons in a genome-wide association study (GWAS). Bmc Genomics. 2010;11:724. doi: 10.1186/1471-2164-11-724. PubMed PMID: 21176216; PubMed Central PMCID: PMCPMC3023815.
48. Yoav Benjamini YH. Controlling the False Discovery Rate: A Practical and

Powerful Approach to Multiple Testing. J R Stat Soc Ser B Methodol. 1995;(57:):289-300.
49. Suissa S. Immortal time bias in pharmaco-epidemiology. Am J Epidemiol. 2008;167(4):492-9. doi: 10.1093/aje/kwm324. PubMed PMID: 18056625.
50. Levesque LE, Hanley JA, Kezouh A, Suissa S. Problem of immortal time bias in cohort studies: example using statins for preventing progression of diabetes. BMJ. 2010;340:b5087. doi: 10.1136/bmj.b5087. PubMed PMID: 20228141.
51. Anderson JR, Cain KC, Gelber RD. Analysis of survival by tumor response. J Clin Oncol. 1983;1(11):710-9. doi: 10.1200/JCO.1983.1.11.710. PubMed PMID: 6668489.
52. Mi X, Hammill BG, Curtis LH, Lai EC, Setoguchi S. Use of the landmark method to address immortal person-time bias in comparative effectiveness research: a simulation study. Stat Med. 2016;35(26):4824-36. doi: 10.1002/sim.7019. PubMed PMID: 27350312.
53. Ng CW, Jiang AA, Toh EMS, Ng CH, Ong ZH, Peng S, et al. Metformin and colorectal cancer: a systematic review, meta-analysis and meta-regression. Int J Colorectal Dis. 2020;35(8):1501-12. doi: 10.1007/s00384-020-03676-x. PubMed PMID: 32592092.
54. Li P, Wu H, Zhang H, Shi Y, Xu J, Ye Y, et al. Aspirin use after diagnosis but not prediagnosis improves established colorectal cancer survival: a meta-analysis. Gut. 2015;64(9):1419-25. Epub 2014/09/23. doi: 10.1136/gutjnl-2014-308260. PubMed PMID: 25239119.
55. Li Y, He X, Ding Y, Chen H, Sun L. Statin uses and mortality in colorectal cancer patients: An updated systematic review and meta-analysis. Cancer Med. 2019;8(6):3305-13. Epub 2019/05/10. doi: 10.1002/cam4.2151. PubMed PMID: 31069997; PubMed Central PMCID: PMCPMC6558478.
56. Amin M, Desai M, Trinkaus K, Brown A, Wang-Gillam A, Tan B, et al. Phase II trial of levocetirizine with capecitabine and bevacizumab to overcome the resistance of antiangiogenic therapies in refractory metastatic colorectal cancer. J Gastrointest Oncol. 2019;10(3):412-20. doi: 10.21037/jgo.2019.02.01. PubMed PMID: 31183190; PubMed Central PMCID: PMCPMC6534727.
57. Deva S, Jameson M. Histamine type 2 receptor antagonists as adjuvant treatment for resected colorectal cancer. The Cochrane database of systematic reviews. 2012;(8):Cd007814. Epub 2012/08/17. doi: 10.1002/14651858.CD007814.pub2. PubMed PMID: 22895966.
58. Fedak KM, Bernal A, Capshaw ZA, Gross S. Applying the Bradford Hill criteria in the 21st century: how data integration has changed causal inference in molecular epidemiology. Emerg Themes Epidemi. 2015;12. doi: ARTN 14
10.1186/s12982-015-0037-4. PubMed PMID: WOS:000210285500001.
59. Liao X, Lochhead P, Nishihara R, Morikawa T, Kuchiba A, Yamauchi M, et al. Aspirin use, tumor PIK3CA mutation, and colorectal-cancer survival. N Engl J Med. 2012;367(17):1596-606. Epub 2012/10/26. doi: 10.1056/NEJMoa1207756. PubMed PMID: 23094721; PubMed Central PMCID: PMCPMC3532946.
60. Kothari N, Kim R, Jorissen RN, Desai J, Tie J, Wong HL, et al. Impact of regular aspirin use on overall and cancer-specific survival in patients with colorectal cancer harboring a PIK3CA mutation. Acta Oncol. 2015;54(4):487-92. doi: 10.3109/0284186X.2014.990158. PubMed PMID: 25549537; PubMed Central PMCID: PMCPMC4743650.

Abstract

Systematic assessment of pharmaceutical prescriptions association with Colorectal Cancer prognosis in a nationwide population-based cohort study in South Korea

Hyeongtaek Woo
Department of Preventive Medicine
The Graduate School
Seoul National University

Objectives

Previous epidemiologic studies examining the prognosis of colorectal cancer and treatment drugs for noncancerous chronic diseases commonly prescribed in colorectal cancer patients were hypothesis-driven studies. Since the relationship with the prognosis of colorectal cancer is studied by limiting to one or several drugs, biased results may result from selective reporting. Therefore, in this study, to prevent the problem of selective reporting, a Medication wide association study(MWAS) was conducted to systematically examine the relationship between all drugs prescribed in colorectal cancer patients and the prognosis of colorectal cancer.

Methods

Using the National Health Insurance Service Sample Cohort DB, followed from 2002 to 2013, all drugs prescribed in patients diagnosed with colorectal cancer were analyzed for drugs related to the patient's prognosis. A total of 2,178 subjects were divided into 1:1 through random sampling and used as a drug discovery set and a drug validation set. The increase in false positives according to the multiple test was controlled using the false detection rate, and the immortal time bias that occurred as drug use changed over time was controlled by increasing the landmark time point to 1, 3 years. Additional analysis was conducted to see if there was a difference in reproducibility of internal validation between the unselected drug and the selected drug.

Results

In the drug discovery set, a total of 773 drugs were analyzed, and the number of drugs found to be related to the prognosis of colorectal cancer according to the landmark time point was 117 at 1 year and 29 at 3 years. Regardless of the time point, 22 drugs showed significant relevance. When 124 drugs that were significant at least once in the drug discovery set were verified in the drug verification set, the significant drugs at 1 and 3 years were 95 and 61 respectively. A total of 17 drugs, including H2-Receptor

Antagonists(ATC code: A02BA), Other Drugs For Peptic Ulcer And Gastro-Oesophageal Reflux Disease (GORD)(ATC code: A02BX), Mosapride(ATC code: A03FA09), Prednicarbate(ATC code: D07AC18), Methylprednisolone(ATC code: H02AB04), Fluconazole(ATC code: J02AC01), Eperisone(ATC code: M03BX09), Enzymes (Used In Musculo-Skeletal System Inflammatory Conditions)(ATC code: M09AB), Mometasone(ATC code: R01AD09), Pseudoephedrine(ATC code: R01BA02), Acetylcysteine(ATC code: R05CB01), Erdosteine(ATC code: R05CB15), Levodropropizine(ATC code: R05DB27), Chlorphenamine(ATC code: R06AB04), Levocetirizine(ATC code: R06AE09), Other Antihistamines For Systemic Use(ATC code: R06AX) were significant in both the drug discovery set and drug validation set. In further analysis of reproducibility through the internal validity evaluation conducted in the drug validation set, the overall reproducibility of drugs found through the MWAS methodology in the drug discovery set was higher.

Conclusions

In this study a total of 17 signals related to the prognosis of colorectal cancer were found and it was confirmed that the MWAS methodology is not only useful for finding new hypotheses, but also can be used to increase the reproducibility of the study.

Keywords : Colorectal cancer, prescription drugs, multiple tests, landmark analysis, reproducibility, selective reporting

Student Number : **2015-23212**