

# HIRA ISSUE

## 코로나19의 세계적 유행에 대응하는 국제협력 연구플랫폼

조도연 주임연구원  
건강보험심사평가원 빅데이터연구부

### | 키워드 |

코로나바이러스감염증-19, 국제협력연구, 온라인 플랫폼, 보건의료빅데이터, 건강보험청구자료, Real World Data(RWD)

### 1. 들어가며

세계보건기구(WHO)는 2020년 3월 11일 코로나19(코로나바이러스감염증-19, COVID-19)를 감염병 경보단계 중 최고 위험 등급인 '팬데믹(pandemic)' 단계로 선언하였다.

전세계적으로 코로나19 확진자는 5월 19일기준 약 470만 명에서 5월 27일기준 약 552만 명으로 8일간 약 82만 명 증가했고 사망자는 31만 7천 명에서 34만 8천 명으로 같은 기간 약 2만 9천 명이 증가했다.

세계 각국이 코로나19에 대응하기 위한 다양한 정책을 추진함에도 불구하고 치료제 개발 및 정책수립의 근거창출을 위한 실제 임상자료(RWD, Real World Data)가 부족한 상황에서, 우리나라는 초기 적극적 진단검사 실시 및 사회적 거리 두기 시행 등의 방역대책을 추진함으로써 질병의 확산과 사망자의 발생을 적극적으로 억제한 결과 국제사회로부터 성공적 방역사례로 평가받고 있으며 더불어 신속한 양방향 진료데이터 분석 시스템 구축으로 국제사회와 협력하고 있다.

국제사회와 협력은 구체적으로 우리나라 건강보험청구자료를 활용하여 코로나19 진료데이터를 정제 및 비식별화하여 구축한 연구용 데이터셋을 3월 27일부터 국내외 연구자에게 개방하는 형태로 진행되고 있다. 이로써 각국은 감염병 정책 수립 근거로 우리나라의 연구용 데이터셋을 활용할 수 있게 되었다. 이에 이 글에서는 보건복지부와 건강보험심사평가원(이하 '심사평가원')에서 제공하는 연구용 데이터셋의 내용과 이용방법 등을 소개하고자 한다.

## 2. 온라인 플랫폼 기반 코로나19 국제협력 연구 개요

### 1) 연구용 데이터셋 구성

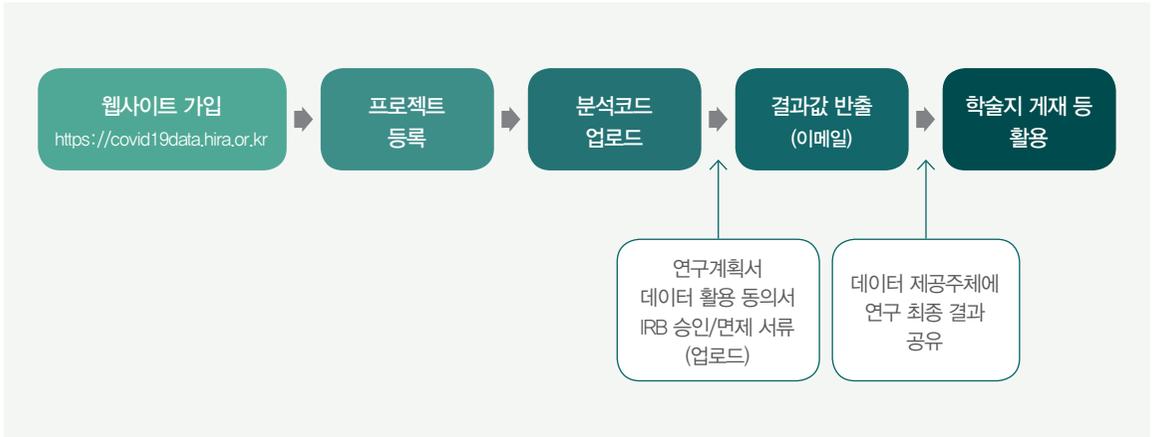
국내외 연구자에게 제공되는 연구용 데이터셋의 자료는 2020년 5월 15일까지 심사평가원에 청구된 코로나19 진료비 청구자료를 활용하여 구축하였다. 연구용 데이터셋은 명세서기본내역(환자 기본정보, 방문 요양기관 정보, 요양급여비용 등), 상병내역, 진료내역(의료행위, 치료재료, 원내의약품 처방내역 등), 원외의약품 처방내역 등 4가지 테이블로 구성되어있다. 분석의 용이성을 위하여 코로나19 진료기록 데이터와 해당 환자들의 과거 3년간(2017년 1월부터 2020년 5월까지)의 코로나19 이외의 일반 진료기록 데이터로 분리하여 각각 4개씩 총 8개의 데이터셋으로 구축하였으며 변수는 총 59개이다. 코로나19 진료기록 데이터셋 기준으로 코로나19 진료내역이 한 번이라도 청구된 사람은 총 234,427명이며, 476,508건의 진료내역이 포함되어 있으며, 모든 데이터셋의 개인정보는 비식별 조치하였다.

### 2) 활용방법

연구등록은 심사평가원이 구축한 온라인 플랫폼을 통해 이루어진다. 연구자는 온라인 플랫폼 웹사이트(<https://covid19data.hira.or.kr>)에 소속기관, 이메일 등 기본 정보를 입력하고 데이터이용동의서를 온라인으로 서명하면 심사평가원에서 등록정보 확인을 위하여 이메일을 송부한다. 연구자는 이메일로 등록 확인 후 연구에 참여할 수 있다.

연구용 데이터셋은 SAS, R 및 공통데이터모델(CDM, Common Data Model) 형태로 심사평가원 내 폐쇄망 서버에 적재되어 있으며, 데이터 레이아웃과 샘플데이터셋을 등록된 연구자에게 웹사이트를 통해 공개하고 있다. 이를 토대로 연구자가 직접 분석 코드를 작성하여 연구수행 페이지에 업로드하면, 심사평가원 연구진이 폐쇄망 서버에서 분석 코드를 실행 후 결과값의 개인정보 포함 여부 등을 검토 후 연구자에게 반출한다.

연구용 데이터셋은 개인정보 보호를 위해 비식별화 조치를 하였을 뿐만 아니라 연구자들이 데이터를 직접 열람하여 분석할 수 없도록 하였으며, 결과값 반출 전에 연구자는 연구계획서, 데이터이용동의서 및 IRB 관련 서류를 업로드해야 한다.



[그림 1] 코로나19 국제협력연구 절차

### 3) 활용 현황

5월 27일 기준으로, 미국과 한국, 영국, 이탈리아, 이스라엘 등 58개국 1,461명의 연구자가 웹사이트에 등록하였고 미국과 한국 연구자가 약 64%를 차지하고 있다. 연구자별 소속은 주로 의과대학 연구자 및 대학병원의 임상 의사이며, 분석업체나 제약사의 연구자도 일부 참여하고 있다.

연구를 위한 데이터셋 활용 요청은 32개국에서 총 371건이 신청되었고, 한국이 147건으로 가장 많았으며, 그 뒤로 미국이 122건, 영국 19건, 이스라엘 13건, 캐나다가 10건을 차지하고 있다. 연구 주제는 주로 코로나19 현황 및 질환의 특성, 기저질환 및 약제와 확진 또는 사망률의 관계, 중증도 환자 예측을 위한 머신러닝 분석 등이다.

## 3. 나가며

전 세계가 당면한 코로나19 치료와 효과적 방역정책 수립을 위한 근거 창출을 위해 한국에서는 온라인 기반으로 코로나19 진료자료를 구축한 후 개인정보의 유출 없이 연구할 수 있는 새로운 방안을 시도하고 있으며 전 세계의 많은 연구자가 연구 참여를 희망하였고 실제 연구가 진행되고 있다.

온라인 기반 코로나19 국제협력연구와 같은 실험적 연구방식이 유의미한 결과를 도출하고 향후 발전적 방향으로 확대 되기 위해서는 몇 가지 고려할 사항이 있다.

먼저, 중요한 정책 주제 혹은 많은 사람들이 관심을 갖고 있는 연구주제에 대한 데이터는 민감한 정보(개인정보 등)를 포함하고 있는 경우가 많아, 이를 어떻게 연계·통합하고 비식별화한 후 개방해야 할지는 여전히 풀어야 할 숙제이다.

이와 함께, 최근 빅데이터 분석을 위한 다양한 오픈소스 분석툴이 활용되고 있으나, 심사평가원과 같은 공공기관 내 빅데이터 플랫폼은 인터넷 연결 없는 폐쇄망을 활용하기 때문에 오픈소스 분석툴의 원활한 업데이트가 이루어지기 어려운 상황이다. 이러한 문제의 해결 및 활발한 빅데이터 분석 환경 제공을 위해서는 관련 보안정책 변경 등 정부의 법적·제도적 개선 노력이 필요하다.

마지막으로, 코로나19 국제협력연구에 참여하는 국가 및 기관이 동일한 데이터셋을 활용한 분산연구 수행으로 국가 단위 임상근거를 생산하고 비교할 수 있는 새로운 방법론의 가능성을 확인하고 새로운 모델을 구축하기 위한 지속적인 연구와 기술개발이 이뤄져야 할 것이다.

#### 참고문헌

코로나19 국제협력연구 홈페이지 <http://covid19data.hira.or.kr>

건강보험심사평가원 보도자료. "심사평가원 코로나19 환자의 연구용 임상데이터셋 구축·공개"(2020.03.27.)

보건복지부 홈페이지. 코로나바이러스감염증-19현황, 국외발생동향. <http://ncov.mohw.go.kr>(검색일자 2020.05.19.)

한국질병관리본부 홈페이지 <http://www.cdc.go.kr>

세계보건기구 홈페이지 <http://www.who.int>

## 평범한 우리가 세상을 바꾸는 방법 **공익신고**

**보호**

- 비밀보장, 신분보호, 불이익조치 금지, 책임감면

**보상**

- 내부 공익신고자에게 최대 30억원의 보상금 지급
- 공익에 기여한 경우 최대 2억원의 포상금 지급
- 구조금(치료비, 이사비, 소송비용 등) 지원

**상담**

- 국번없이 **110** 또는 **1398**

**신고**

- 홈페이지 **1398.acrc.go.kr**
- 우편(서울시 서대문구 통일로 87)

**신고대상** : 6대 분야, 284개 법률 위반행위

<p><b>건강</b></p> <ul style="list-style-type: none"> <li>불량식품 제조 판매</li> <li>무면허 의료행위</li> </ul>	<p><b>안전</b></p> <ul style="list-style-type: none"> <li>부실시공</li> <li>소방시설 미설치</li> </ul>
<p><b>환경</b></p> <ul style="list-style-type: none"> <li>폐수 무단방류</li> <li>폐기를 불법 매립</li> </ul>	<p><b>소비자이익</b></p> <ul style="list-style-type: none"> <li>개인정보 무단 유출</li> <li>허위 과장광고</li> </ul>
<p><b>공정경쟁</b></p> <ul style="list-style-type: none"> <li>기업 간 담합</li> <li>불법 허도급</li> </ul>	<p><b>기타 공공의 이익</b></p> <ul style="list-style-type: none"> <li>가짓 채용광고</li> <li>방위산업기술 불법사용</li> </ul>

국민권익위원회

# HIRA ISSUE

발행일 2020.6.25.

발행처 건강보험심사평가원 심사평가연구소

발행인 송재동

HIRA ISSUE는 국내외 보건의료 현안에 대한 정보제공을 위해 제작되었습니다.  
본 내용은 심사평가연구소 연구진의 견해로 건강보험심사평가원의 공식 입장과 다를 수 있습니다.

강원도 원주시 혁신로 60(반곡동)

Tel. 033-739-0906 | www.hira.or.kr