



# 항생제 내성 정보 제공 서비스

공조 - 김예린 | 김혜진 | 김희성 | 정현정

# 목차

1

## 프로젝트 개요

- 01. 주제 선정 배경
- 02. 분석 목표

2

## 데이터 전처리

- 01. 데이터 변수 설명
- 02. 데이터 전처리
- 03. 데이터 병합

3

## 모델링 및 성능평가

- 01. 모델 탐구
- 02. 모델 설명
- 03. 모델 선택
- 04. 모델 성능

4

## 서비스 소개 및 결론

- 01. 서비스 구축
- 02. 기대효과 및 활용방안
- 03. 참고문헌





## 김예린

1. 데이터 전처리
2. 논문 및 배경정보 검색
3. 모델링
4. PPT 작성
5. 발표

## 김혜진

1. 데이터 전처리
2. 앱, 웹 인터페이스 구축

## 김희성

1. 데이터 수집
2. 코딩 작성
3. PPT 작성

## 정현정

1. 데이터 전처리
2. 논문 및 배경정보 검색
3. 데이터 분석
4. PPT 작성

# 01

## 주제 선정 배경

- 1. 항생제 오남용
- 2. 항생제 내성 위험성

## 분석 목표

- 1. 항생제 오남용의 최소화
- 2. 사용자에게 정보를 제공

# 유럽에서만 항생제 내성균 때문에 매일 100명 사망

## 모든 항생제 무용지물 ... '다제내성 슈퍼 임질균' 비상

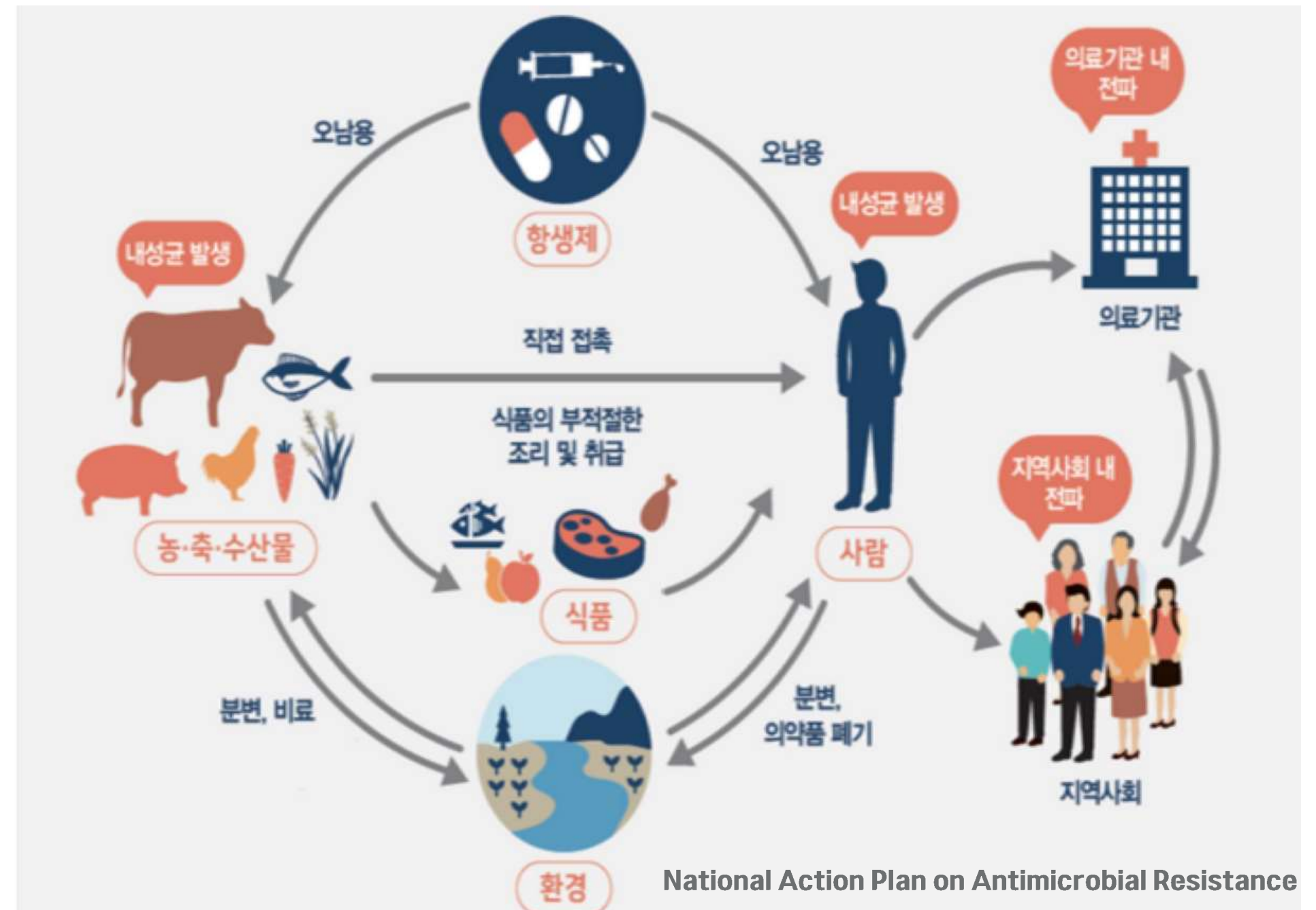
세팔로스포린 계열 항생제 내성 임질균 확산 중

HIV처럼 장기 관리 불치병 발전할 가능성 있어

확산세 비해 수익성 낮아 새로운 항생제 개발 더뎠

2019년 전국 의료기관 대상 항생제 처방 질적 평가에서는 전체 항생제 처방 중 26.1%가 부적절한 것으로 확인된 바 있다.

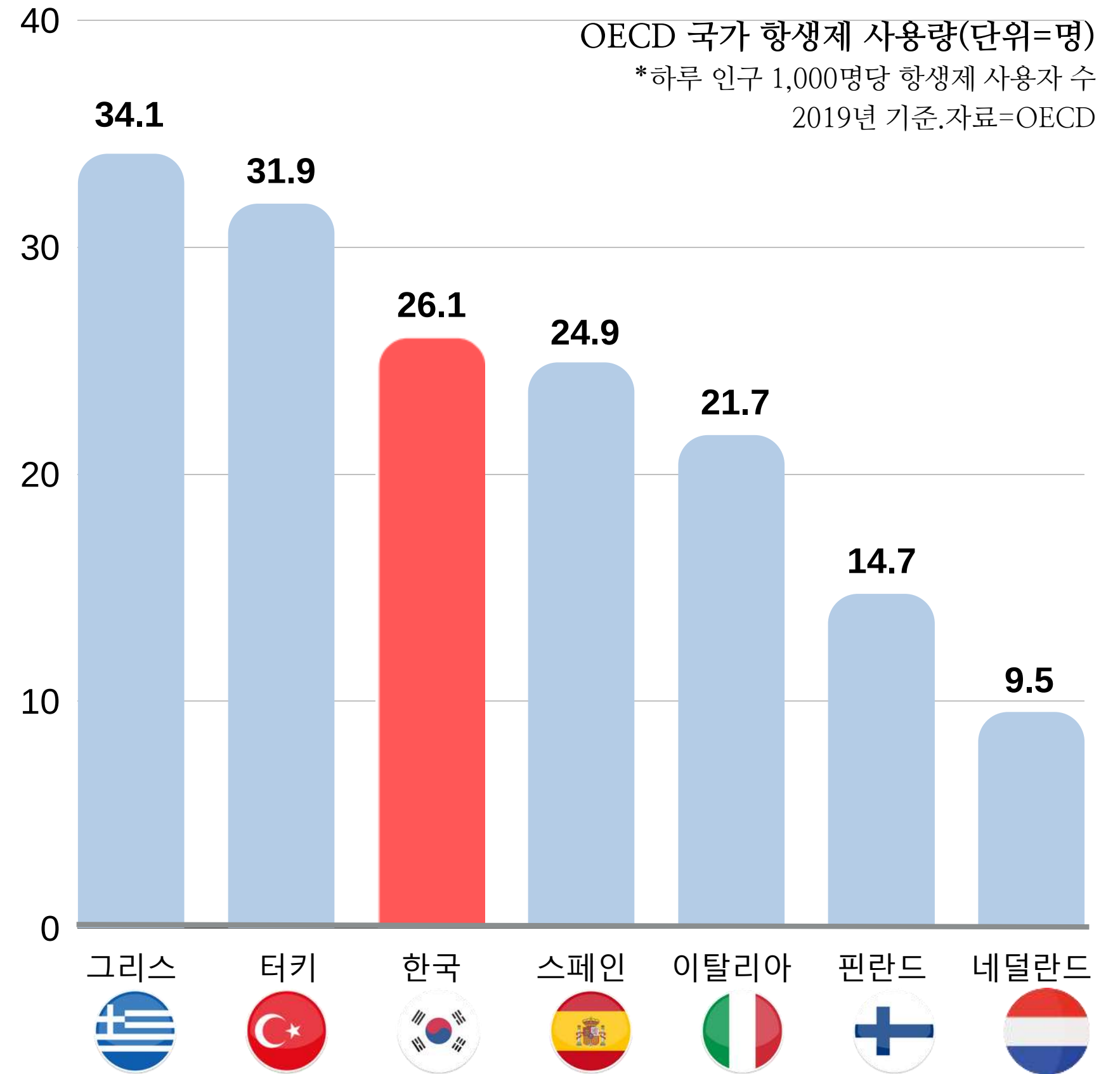
2022년 의사 대상 항생제 인식도 조사에서는 의사의 40% 이상이 항생제가 불필요한 상황임에도 처방한 경험 있다고 답했다.





감기 항생제 처방 20 년새 절반됐지만  
항생제 사용량 OECD 3 위 오명

코로나發 항생제 남용에 내성균도 급증... '조용한 팬데믹' 공포

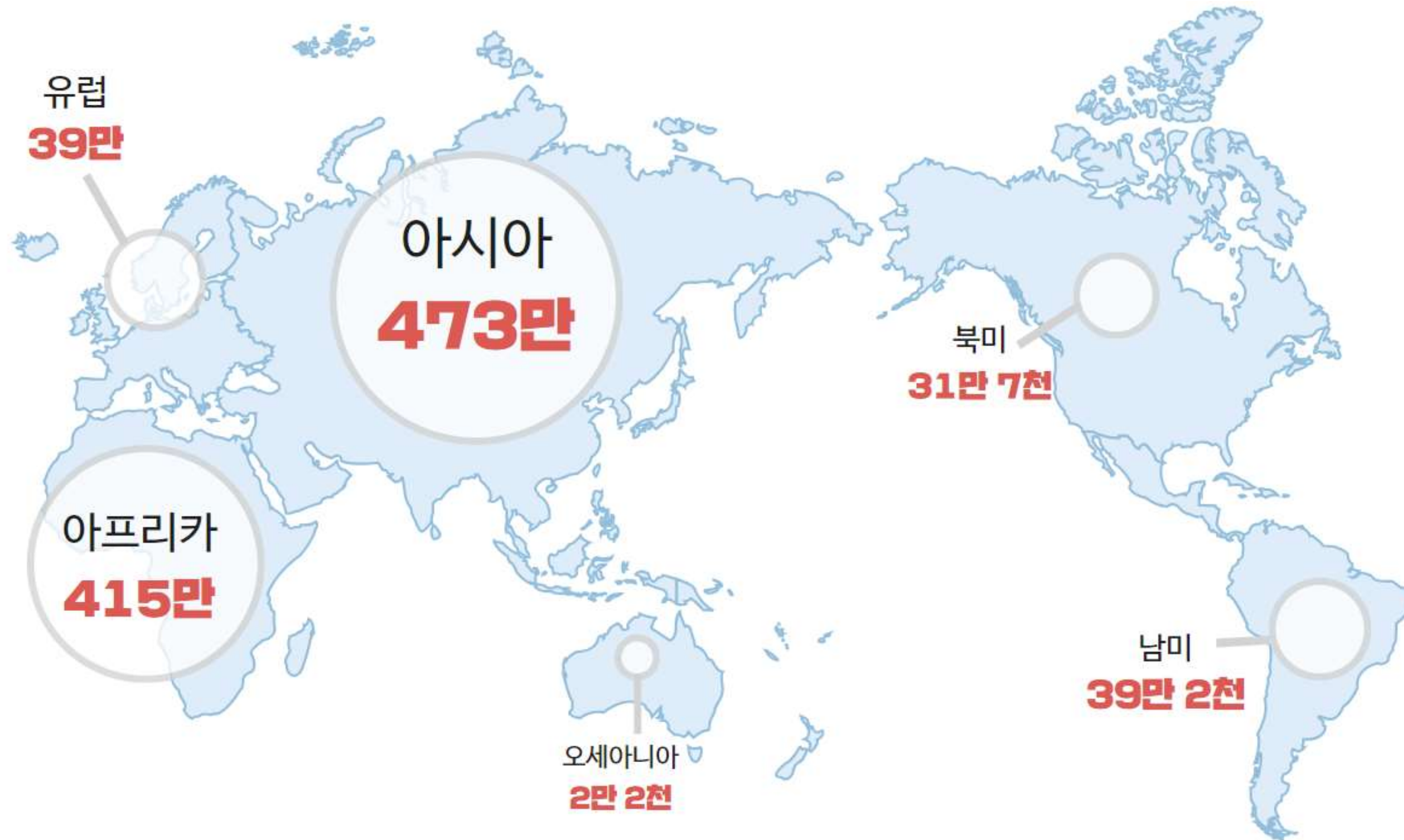


항생제 처방률 OECD 29개 나... 환자 의사를 모조리 무시할 순 없지 않습니까?" "혹시 모를 2차 감염 가능성과

코로나 사태 이후 의사들 사이에서 항생제 처방을 둘러싼 논쟁이 이어지고 있다. 세균 잡는 항생제를 바이러스 질환인 코로나에 처방하는 게 맞느냐는 것이다. 이론적으로는 항생제는 세균에 작용하기 때문에 바이러스에는 효과가 없다. 하지만 의료 현장에선 이를 적용하기가 쉽지 않다. 대면 진료로는 환자의 증상이 코로나 때문인지, 면역력 저하로 2차 세균 감염이 발생한 것인지 구분하기 어렵기 때문이다.

전문가들은 "그래도 최대한 항생제 처방을 줄여야 한다"고 목소리를 높인다. 의원급에서는 코로나 비대면 처방으로 항생제가 과도하게 사용되고, 대형 병원급에서도 코로나 입원 환자가 폭증하면서 항생제 사용이 급증한 상황. 이로 인해 여러 항생제에 내성이 생긴 세균, 이른바 '다제내성균'이 늘어 더 큰 문제가 생길 수 있다는 우려다.

## 2050년 연간 대륙별 항생제 예측 사망자 전망치



# 02

## 데이터 변수 설명

1. 배양결과 데이터
2. 항생제 처방 리스트 데이터
3. 치료결과 데이터

## 데이터 병합

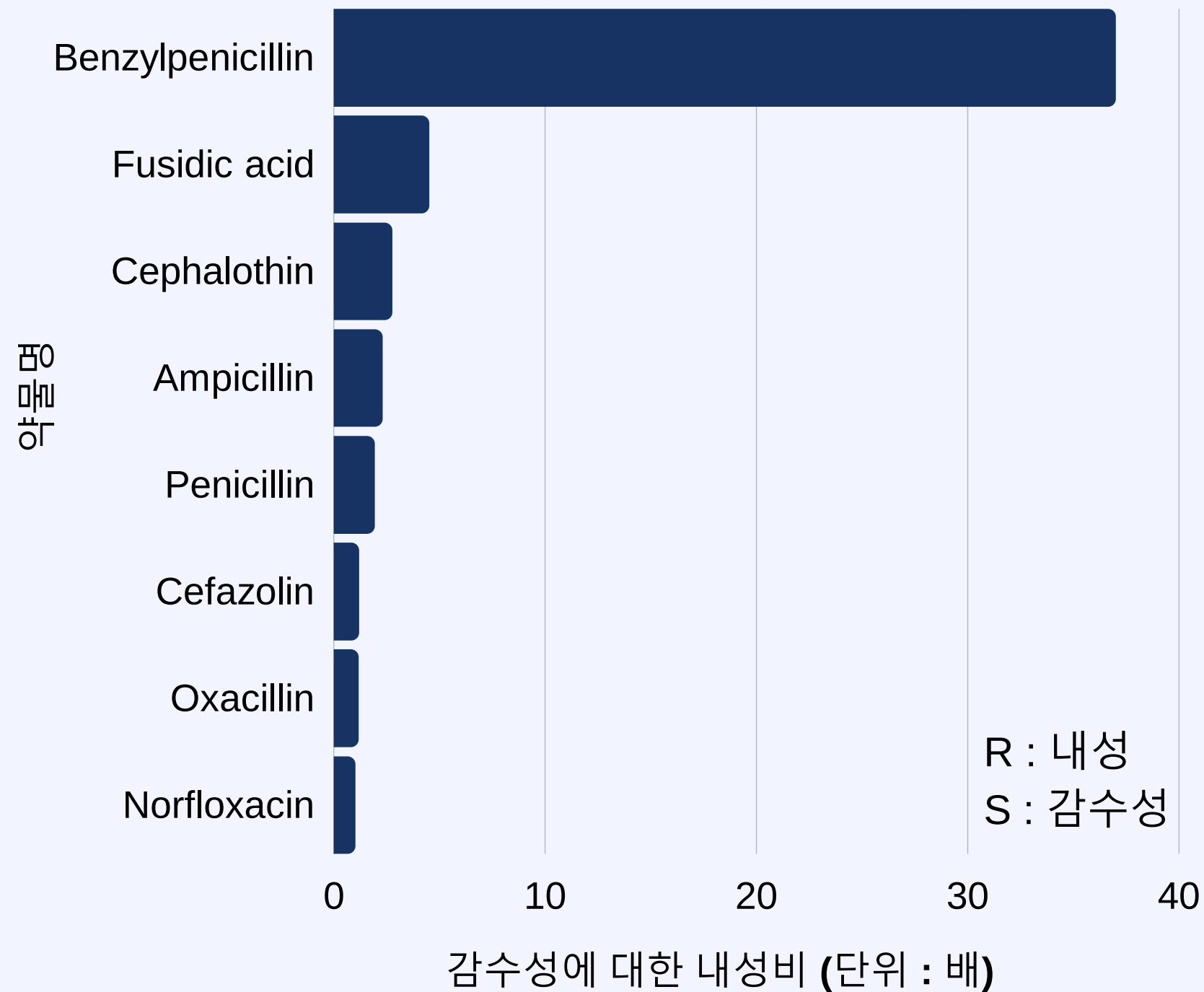
1. 내원정보&환자정보를 기준  
(배양결과 & 항생제 처방 리스트 & 치료결과)  
데이터 병합

## 데이터 전처리

1. 필터링
2. 불필요한 데이터 제거
3. 단위 mg 통일



## 항생제에 따른 내성 비율(R/S)



## 연관분석 결과 (항생제와 내성의 연관성)

항생제 이름	지지도	신뢰도	향상도
ANTBT_NM	Support_R	Confidence_R	Lift_R
Benzylpenicillin	0.000187859	0.973684211	2.47717633
Cephalothin	0.004874185	0.732824427	1.864398442
Ampicillin	0.037975487	0.695217735	1.768722239
Penicillin	0.019557668	0.655882854	1.668649304
Cefazolin	0.016823554	0.538386546	1.369723771
Oxacillin	0.008669449	0.532678209	1.355201037

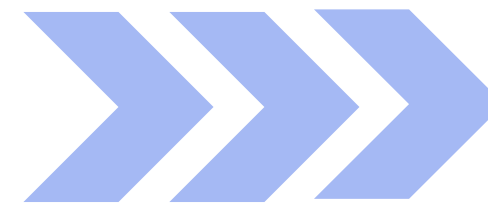
### • Lift (ANTBT -> R)

$$= \frac{\text{항생제를 사용하지 않았을 때 내성(R)이 걸릴 확률}}{\text{항생제를 사용했을 때 내성(R)이 걸릴 확률의 증가 비율}}$$

★ ★ ★  
★ Lift > 1 : 항생제 사용과 내성 간에  
양의 연관성이 존재

약 성분명	1회 처방량	1회 처방횟수	처방일수
INGR_NM	PRSC_CAPA	PRSC_NT	PRSC_DCNT
Cefpodoxime proxetil	100	2	3
Ciprofloxacin	500	2	7
Ciprofloxacin	500	1	1
Levofloxacin	500	1	1
Azithromycin	250	1	3
lincomycin	150	1	1
Cefadroxil	500	2	3
Meropenem	1000	3	1
Moxifloxacin HCl	400	1	1

항생제 필요한 열만 추출



약 성분명	1일 총처방량	일수
INGR_NM	TPRSC_CAPA	DAY
Cefpodoxime proxetil	200	3
Ciprofloxacin	1000	5.4
Levofloxacin	750	7
Ciprofloxacin	800	6
Moxifloxacin HCl	400	7
Gentamicin	80	1
Sultamicillin	1125	2
Cefixime	200	5
lincomycin	120	1

항생제 데이터 전처리

## 1회 복용량

약 성분명	PRSC_CAPA	1회 처방횟수	처방일수
INGR_NM		PRSC_NT	PRSC_DCN
Cefpodoxime pr	100	2	1
Ciprofloxacin	500	2	1
Ciprofloxacin	500	1	1
Levofloxacin	250	1	1
Azithromycin	150	1	1
lincomycin	500	1	2
Cefadroxil	500	2	3
Meropenem	1000	3	1
Moxifloxacin HC	400	1	1

## 1일 복용량

PRSC_NT
2
2
1
1
1
1
2
3
1

## 1일 총 복용량

TPRSC_CAPA	량	일수
INGR_NM	CAPA	DAY
Cefpodox	200	3
Ciproflo	1000	5.4
Levofloxa	750	7
Ciproflo	800	6
Moxifloxa	400	7
Gentamic	80	1
Sultamici	1125	2
Cefixime	200	5
lincomyc	120	1

항생제 필요한 열만 추출

항생제 데이터 전처리

**딥러닝(DNN)모델 & 시계열(SARIMA) 모델 학습에 사용**

배양 결과 TABLE

항생제 이름	내성 여부
ANTBT_NM	RSLT_CONT
Amikacin	S
Ampicillin/Sulbactam	S
Ceftazidime	R

항생제 TABLE

약 성분명	1일 총처방량	일수
INGR_NM	TPRSC_CAPA	DAY
Cefpodoxime proxetil	200	3
Azithromycin	250	3
lincomycin	150	1

치료결과 TABLE

상태
antibiotics
경쾌
완쾌
완쾌

VTHOS\_NO & RGNO  
기준으로 파일 합치기

항생제  
+ 배양  
+ 치료결과  
TABLE

약 성분명	1일 총처방량	일수	상태	항생제 이름	내성 여부
INGR_NM	TPRSC_CAPA	DAY	antibiotics	ANTBT_NM	RSLT_CONT
Cefpodoxime proxetil	200	3	경쾌	Amikacin	S
Azithromycin	250	3	완쾌	Ampicillin/Sulbactam	S
lincomycin	150	1	경쾌	Ceftazidime	R



# 03

## 모델 탐구

1. 랜덤 포레스트
2. 로지스틱 회귀모형
3. 그레디언트 부스트
4. 에이다부스트
5. 딥러닝 DNN

## 모델 성능

1. 기존 연구와 비교

## 모델 선택

1. 딥러닝 DNN 모델 선택
2. 하이퍼파라미터 튜닝
  - Hidden layer
  - Epochs

## 시계열

1. 월별 항생제 사용량의 합 예측 및 실제 비교
2. SARIMA
3. 성능 확인

## 기존 연구

'탐색적 데이터 분석과 기계학습을 통한 상부 요로감염  
환자 Ciprofloxacin 항생제 내성 예측 연구', 2023.03  
한국정보과학회 논문지

	AUROC	Accuracy	Precision	F1 score
SVM SHAP	0.775	0.641	0.606	0.707

## 모델 탐구

용인시 데이터 전처리 후  
모델링 결과

모델	AUROC	accuracy	Precision	Recall	F1 Score
랜덤포레스트	0.732	0.784	0.778	0.784	0.779
로지스틱 회귀	0.625	0.728	0.713	0.728	0.698
그레디언트 부스트	0.629	0.738	0.731	0.738	0.705
AdaBoost	0.654	0.741	0.728	0.741	0.721
딥러닝 DNN	0.754	0.732	0.718	0.732	0.704



AUROC값이 높은  
딥러닝(DNN) 모델을 선택

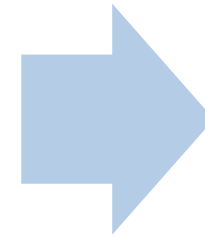
## 제안하는 딥러닝(DNN) 모델 구조

### Hyper-parameter tuning

- hidden layer1 units: 8, 'relu'
- hidden layer2 units: 4, 'relu'
- output layer units: 1, 'sigmoid'
- epochs: 30

### Model compile

- loss : binary\_crossentropy
- optimizer : adam
- metrics : accuracy



	AUROC	accuracy	Precision	Recall	F1 Score
test1	0.766228336	0.731798933	0.722796159	0.722796159	0.722796159
test2	0.744936102	0.703084568	0.705222953	0.703084568	0.704100345
test3	0.74971528	0.731693201	0.72188679	0.731693201	0.697288548
test4	0.74719652	0.712650458	0.706378628	0.712650458	0.708892773
test5	0.759073584	0.739372625	0.726410505	0.726410505	0.726410505
test6	0.761632444	0.735766635	0.735766635	0.735766635	0.735766635
test7	0.765571743	0.735822283	0.72262562	0.735822283	0.712634177
test8	0.763685021	0.728554655	0.717999145	0.728554655	0.692498229
test9	0.76155331	0.729862383	0.723312331	0.729862383	0.68973772
test10	0.727272843	0.728671515	0.721608099	0.728671515	0.688114359
평균	0.754686518	0.727727725	0.720400686	0.725531236	0.707823945

## 제안하는 딥러닝(DNN) 모델 구조

### Hyper-parameter tuning

- hidden layer1 units: 8, 'relu'
- hidden layer2 units: 4, 'relu'
- output layer units: 1, 'sigmoid'
- epochs: 30

### Model compile

- loss : binary\_crossentropy
- optimizer : adam
- metrics : accuracy



Accuracy Mean (정확도 평균)

**0.7277**



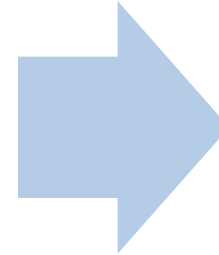
## 제안하는 딥러닝(DNN) 모델 구조

### Hyper-parameter tuning

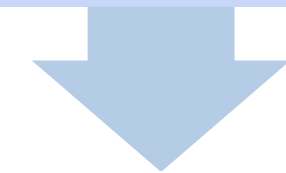
- hidden layer1 units: 8, 'relu'
- hidden layer2 units: 4, 'relu'
- output layer units: 1, 'sigmoid'
- epochs: 30

### Model compile

- loss : binary\_crossentropy
- optimizer : adam
- metrics : accuracy



units8/4,epchs30	AUROC	accuracy	Precision	Recall	F1 Score
test1	0.766228336	0.731798933	0.722796159	0.722796159	0.722796159
test2	0.744936102	0.703084568	0.705222953	0.703084568	0.704100345
test3	0.74971528	0.731693201	0.72188679	0.731693201	0.697288548
test4	0.74719652	0.712650458	0.706378628	0.712650458	0.708892773
test5	0.759073584	0.739372625	0.726410505	0.726410505	0.726410505
test6	0.761632444	0.735766635	0.735766635	0.735766635	0.735766635
test7	0.765571743	0.735822283	0.72262562	0.735822283	0.712634177
test8	0.763685021	0.728554655	0.717999145	0.728554655	0.692498229
test9	0.76155331	0.729862383	0.723312331	0.729862383	0.68973772
test10	0.727272843	0.728671515	0.721608099	0.728671515	0.688114359
평균	0.754686518	0.727727725	0.720400686	0.725531236	0.707823945



모델의 성능 검증을 위한  
Test 분류 결과

test1	69.14%
test2	67.85%
test3	67.63%
test4	68.28%
test5	67.64%
test6	68.27%
test7	67.74%
test8	67.73%
test9	67.86%
test10	68.00%
평균값	68.01%

## 제안하는 딥러닝(DNN) 모델 구조

### Hyper-parameter tuning

- hidden layer1 units: 8, 'relu'
- hidden layer2 units: 4, 'relu'
- output layer units: 1, 'sigmoid'
- epochs: 30

### Model compile

- loss : binary\_crossentropy
- optimizer : adam
- metrics : accuracy



units8/4,epochs30	AUROC	accuracy	Precision	Recall	F1 Score
test1	0.766228336	0.731798933	0.722796159	0.722796159	0.722796159
test2	0.744936102	0.703084568	0.705222953	0.703084568	0.704100345
test3	0.74971528	0.731693201	0.72188679	0.731693201	0.697288548
test4	0.74719652	0.712650458	0.706378628	0.712650458	0.708892773
test5	0.759073584	0.739372625	0.726410505	0.726410505	0.726410505
test6	0.761632444	0.735766635	0.735766635	0.735766635	0.735766635
test7	0.765571743	0.735822283	0.72262562	0.735822283	0.712634177
test8	0.763685021	0.728554655	0.717999145	0.728554655	0.692498229
test9	0.76155331	0.729862383	0.723312331	0.729862383	0.68973772
test10	0.727272843	0.728671515	0.721608099	0.728671515	0.688114359
평균	0.754686518	0.727727725	0.720400686	0.725531236	0.707823945



모델의 성능 검증을 위한  
Test 분류 정확도

68%

## 항생제 TABLE

1일 총처방량      월      일

TPRSC_CAPA	Month	Day
2000	11	3
240	11	30
18000	2	17
6000	5	9
13500	1	2

Train data : 0.9

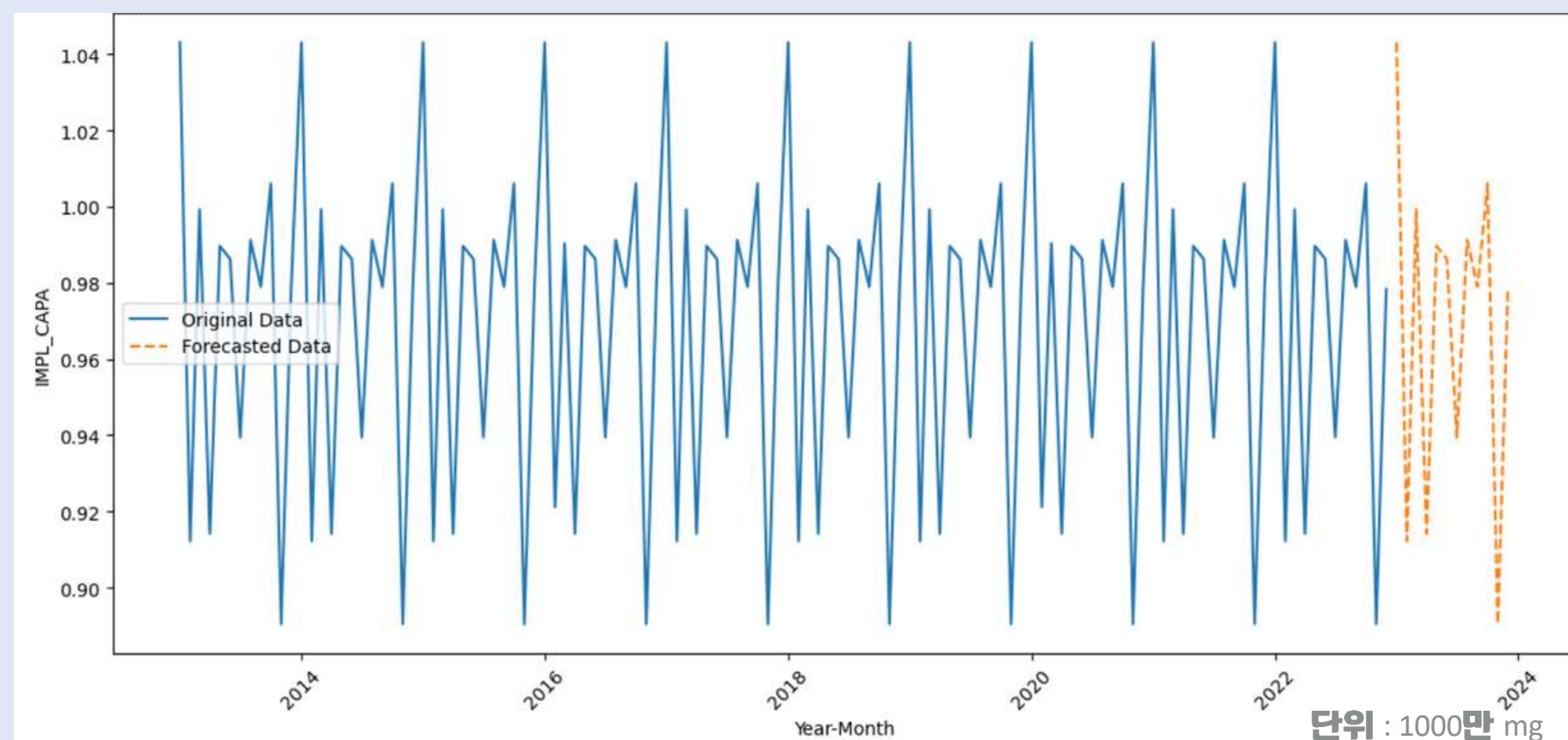
(월-일) 기준 총화추출  
2013년 ~2022년 가정

본 데이터는 약 10년간

1일 총처방량      날짜

TPRSC_CAPA	Date
2000	2013-11-03
240	2022-11-30
18000	2017-02-17
6000	2021-05-09
13500	2018-01-02

## < 2023년 월별 항생제 사용량의 합 예측 >



**특정한 패턴과 일정 주기로 값이 변동**

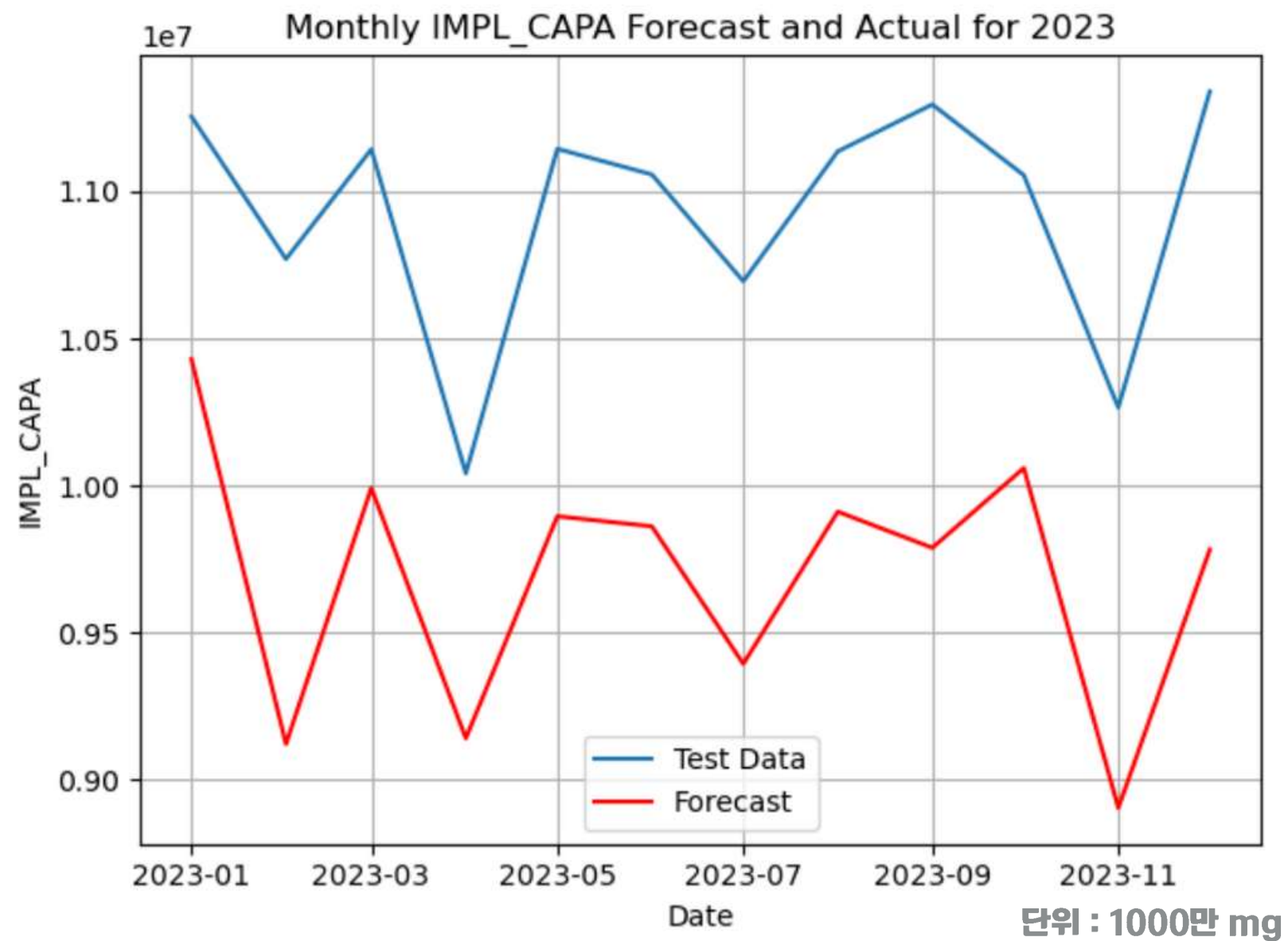
=> 12개월 주기의 **계절성 확인** 가능

=> **SARIMA** 모델 사용



## Test data 0.1

## &lt; 2023년 월별 항생제 사용량의 합 예측 및 실제 비교 &gt;



## SARIMA 모델

Test data를 예측 및 실제값 비교

모델의 정확성 : 88.63%

```
# 상대적 오차 계산
relative_errors = (test_monthly_sum.values - forecasted_caps) / test_monthly_sum.values

# 상대적 오차의 절대값을 취한 후 평균을 계산
mean_relative_error = np.mean(np.abs(relative_errors))

# 평균 상대적 오차를 백분율로 변환
accuracy_percentage = (1 - mean_relative_error) * 100

print(f"모델의 정확성: {accuracy_percentage:.2f}%")
```

모델의 정확성 : 88.63%



# 04

## 서비스 구축

1. AnT 앱 서비스
2. AnT+ 웹 서비스

## 참고문헌

## 기대효과 및 활용방안

# 서비스 소개



환자를 위한 서비스

아함<sup>+</sup>



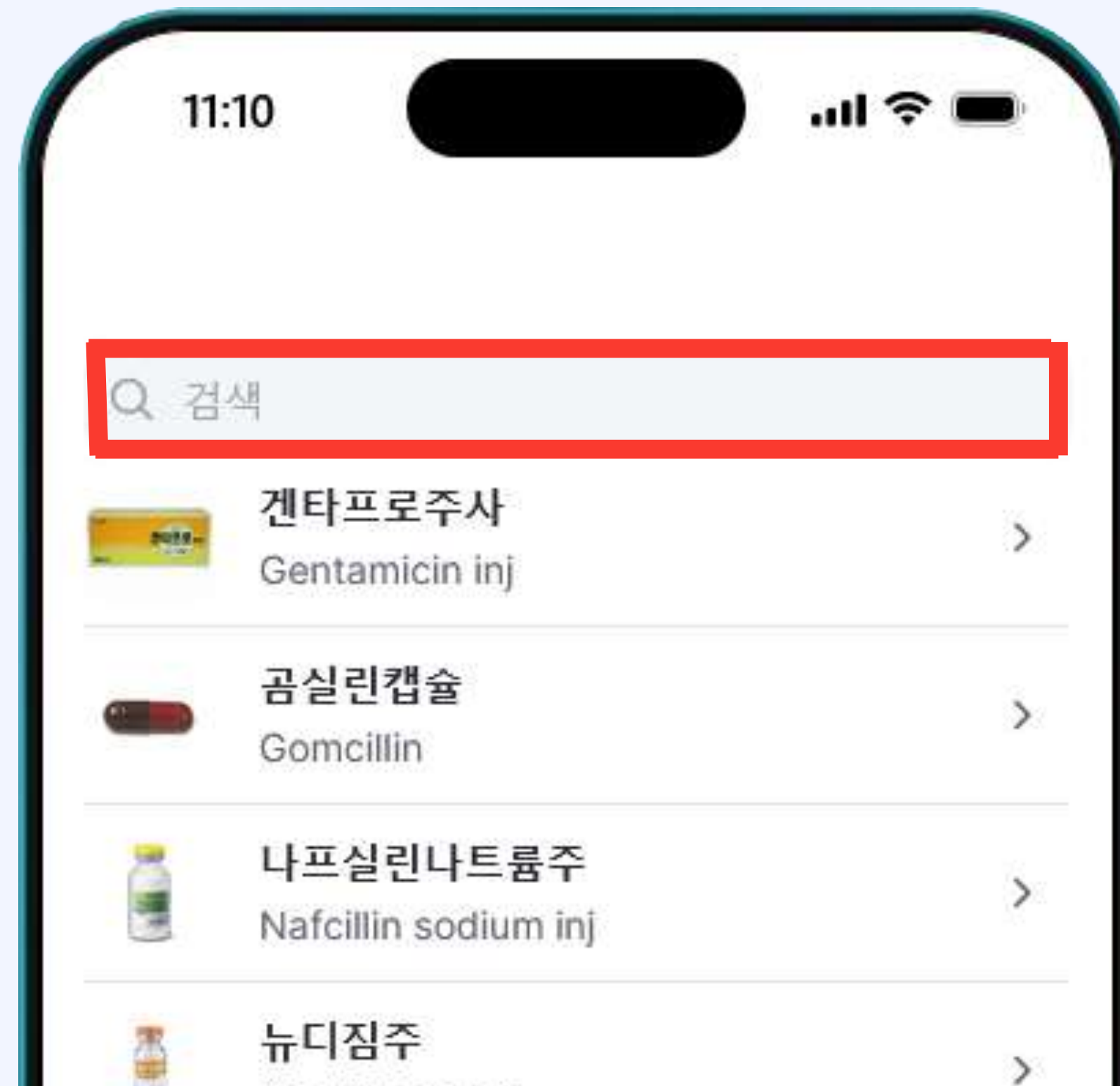
의료 종사자를  
위한 서비스

아함<sup>+</sup>

## " 검색 기능 "

약물 정보 입력

약이름, 성분, ATC코드



## " 정보 기능 "

### 정보 출력

#### 1) 약 이름

#### 2) 약 사진

#### 3) 내성 확률

#### 4) 복용법

#### 5) 추가 정보



#### 복용법

##### 성인

1일 체중 Kg당 3mg(역가)을 8시간마다 분할 정맥 혹은 근육주사  
 생명이 위독한 경우: 1일 체중 Kg당 5mg(역가)을 3 ~ 4회 분할 주사  
 증상의 개선에 따라 감량, 최고혈중농도는 12 $\mu$ g/mL, 최저혈중농도는 2 $\mu$ g/mL가 넘지 않아야 함

##### 소아

미숙아 및 생후 1주 미만의 신생아: 1일 체중 Kg당 5mg(역가)을 12시간마다 분할 정맥주사  
 1일 체중 Kg당 6 ~ 7.5mg(역가)을 8시간마다 분할 정맥 혹은 근육주사  
 유 · 소아의 경우 희석액의 약을 적절히 감량 조절

##### 신장장애인

신기능부전 환자에는 정도에 따라 용량을 조절





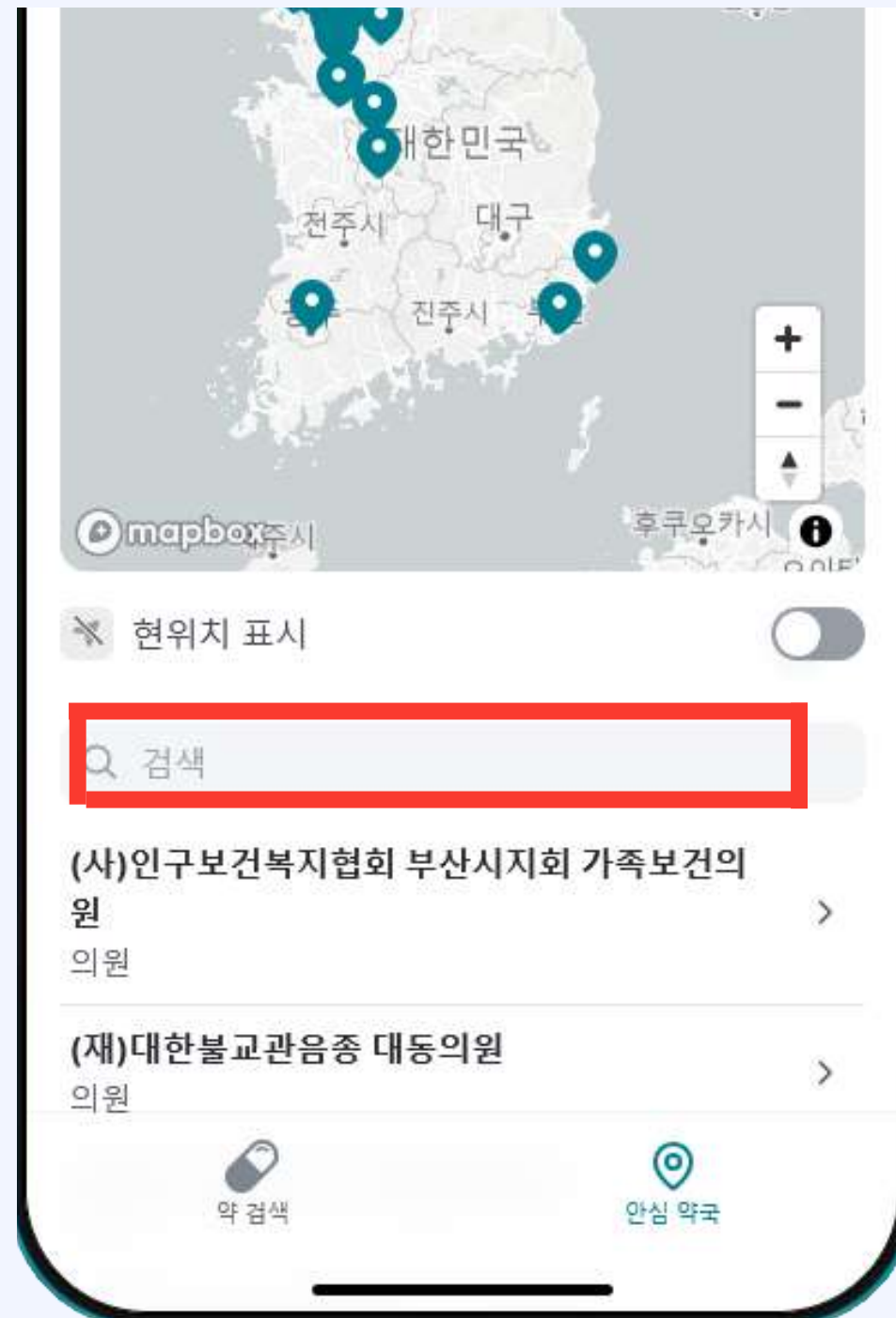
# "안심 약국"

## 그린 처방 의원

보건복지부가 선정한  
의약품 '적정 처방' 의원

## 주소 입력

~도, ~시



## 정보 출력

### 1) 의원 명

### 2) 전화번호

### 3) 주소



# " 시연 영상 "

## 아함



# " 정보 기능 "

## 1. 정보 입력

- 1) 1회 복용량
- 2) 하루 복용 횟수
- 3) 복용 기간

## 2. 정보 출력

- 1) 총 처방량
- 2) 내성 확률



**겐타프로주사**  
Gentamicin inj  
  
성분명  
Gentamicin  
  
ATC코드  
J01GB03  
  
의약품 안전나라

**복용 정보 입력**

1회 복용량

하루 복용 횟수

복용 기간

입력

## 분석 결과

입력한 정보를 기반으로 계산된 결과와 모델 예측 값은 다음과 같습니다:

총 처방량: 5400

내성 확률: 34.35 %

## " 예측 기능 "

### 1. 저장 및 예측

1) 사용량

2) 예측 사용량

3) 비교

약물 검색      사용량 예측

약물 사용 정보 입력

1회 복용량(mg)  
2000

하루 복용 횟수  
3

날짜  
2023-09-01

입력

예측과 비교




# " 예측 기능 "


## 1. 저장 및 예측

1) 사용량

2) 예측 사용량

3) 비교

약물 검색

사용량 예측

### 약물 사용 정보 입력

1회 복용량(mg)

하루 복용 횟수

날짜

입력

### 예측 결과

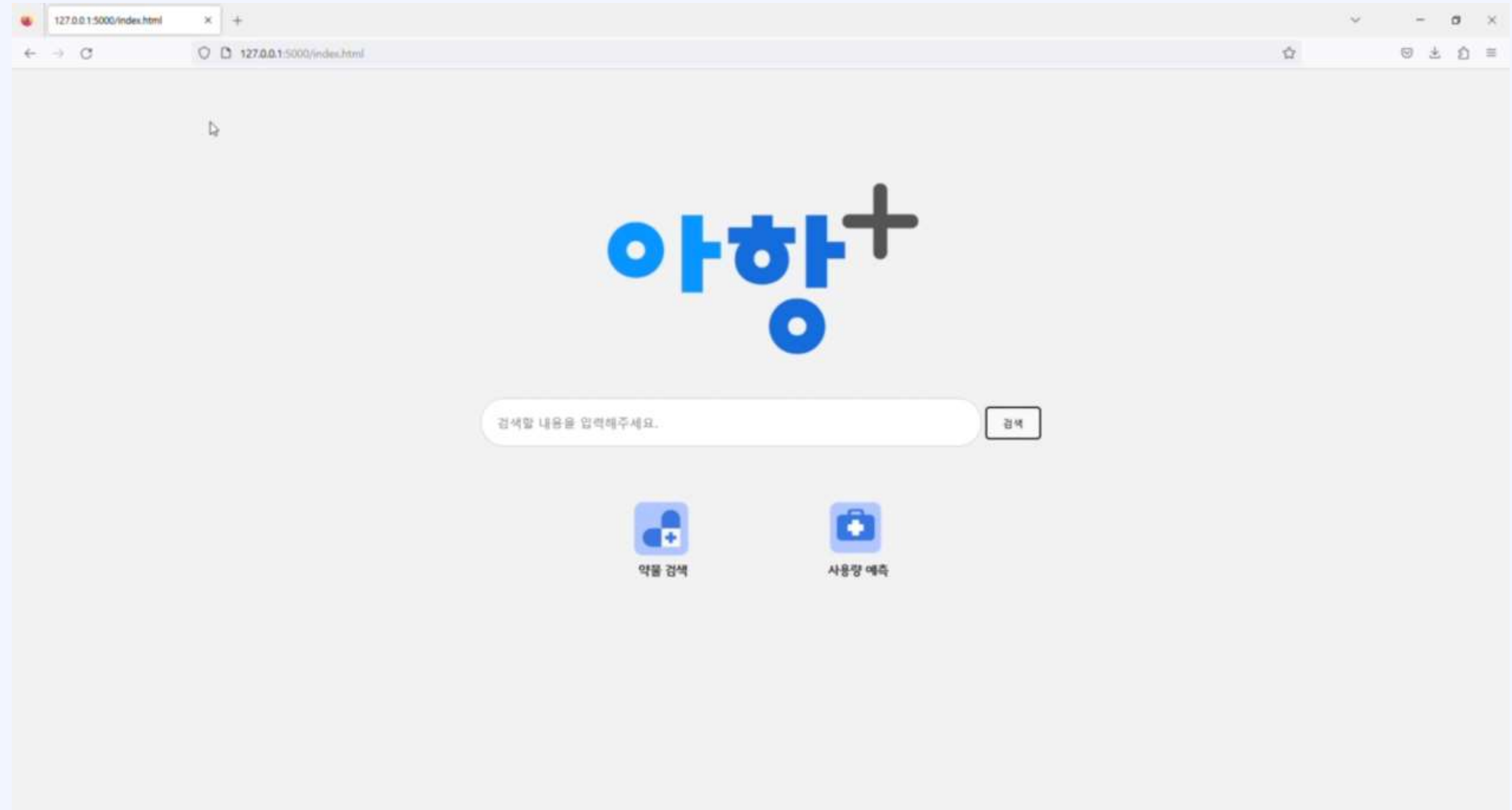
이번 달 실제 항생제 사용량: 1034000.0

이번 달 예측 항생제 사용량: 9789142.0

2023-09월에는 항생제가 예상보다 0.11 배 만큼 더 쓰였습니다.

" 시연 영상 "

아함+



## 서비스 활용방안 및 기대효과



- 그런 처방의원 목록을 제공 해줌으로써 보다 안전한 약 처방을 받을 수 있을 것이라 기대
- 항생제에 대한 정보를 쉽게 얻을 수 있는점
- 항생제 내성에 대한 경각심



- 진료과와 같은 병원의 특성을 고려하면 전반적인 항생제 관리 시스템으로 확장시킬 수 있을 것으로 생각

# 출처

## [해외 논문]

Antibiotic Resistance Threats in the United States, 2019

OECD Health Statistics, 2021

## [국내 논문]

딥러닝 기반 항생제 내성균 감염 예측, 2019

탐색적 데이터 분석과 기계학습을 통한 상부 요로감염 환자 Ciprofloxacin 항생제 내성 예측 연구, 2023

## [자료]

유럽에서만 항생제 내성균 때문에 매일 100명 사망|한의학신문|2023.07.19

한국, 항생제 사용량 OECD서 3위...2025년까지 20% 줄인다|연합뉴스|2021.11.07

모든 항생제 무용지물 ... '다제내성 슈퍼 임질균' 비상|헬스코리아뉴스|2023.02.07

National Action Plan on Antimicrobial Resistance|국무조정실|(2016~2020)

국민 1000명 중 26명 매일 항생제 복용..."내성 심각성 알아야"|news1|2021.11.18

의약품 복용법 : 의약품 안전나라 의약품 통합 정보 시스템

보건 복지부

용인시 산업 진흥원

항생제 사용량, OECD 3위|후생신보|박원빈기자|2021.10.20

'항생제 사용량' OECD 3위인 한국...2025년까지 20% 감소 목표|브릿지경제|2021.11.07

국내 인체 '항생제' 사용량 OECD 3위...4년 뒤 20% 감축|한스경제|홍성익 기자|2021.11.08

1000명 중 26명꼴 매일 항생제 처방, OECD 3위|조선일보|2019.11.18

"韓, 감기 환자 10명 중 4명에 항생제 처방"|뉴시스|2020.10.20

논문 내용 상의 출처는 제출 논문 참조





**THANK YOU**

**Q & A**