

세션 1. 소셜리스닝 결과

# 암환자들을 위한 디지털 정보 검색의 한계

A hand holding a smartphone in the center, displaying a social media interface with the word 'Reels' at the top. The screen shows a profile page with a 'Follow' button and a 'Great job' notification. Surrounding the phone are various social media icons: YouTube, Twitter, Facebook, Instagram, Snapchat, and Pinterest, along with smaller icons like a heart, a speech bubble, and a location pin. The background is a blurred outdoor scene with trees and a building.

# Social Listening이란?

소셜 이용자가 언급한 '날 것'의 정보들을 빨리, 광범위하게 감지 (sensing)하여 그에 대한 적절한 반응을 하거나 사전에 적절한 메시지를 설계해 적절한 채널을 통해 전파하기 위해 하는 활동

# 조사배경

HOME > 의사협회

## 국민 10명 7명 인터넷 통해 부정확한 암정보 획득

송성철 기자 | 승인 2016.04.09 17:54

## 성인남성 10명 중 7명, 전립선암 정보 올바른지 잘 몰라

🕒 2021-03-02 10:47:31



## 암 환자가 찾아본 정보, 70%는 충분하지 못했다

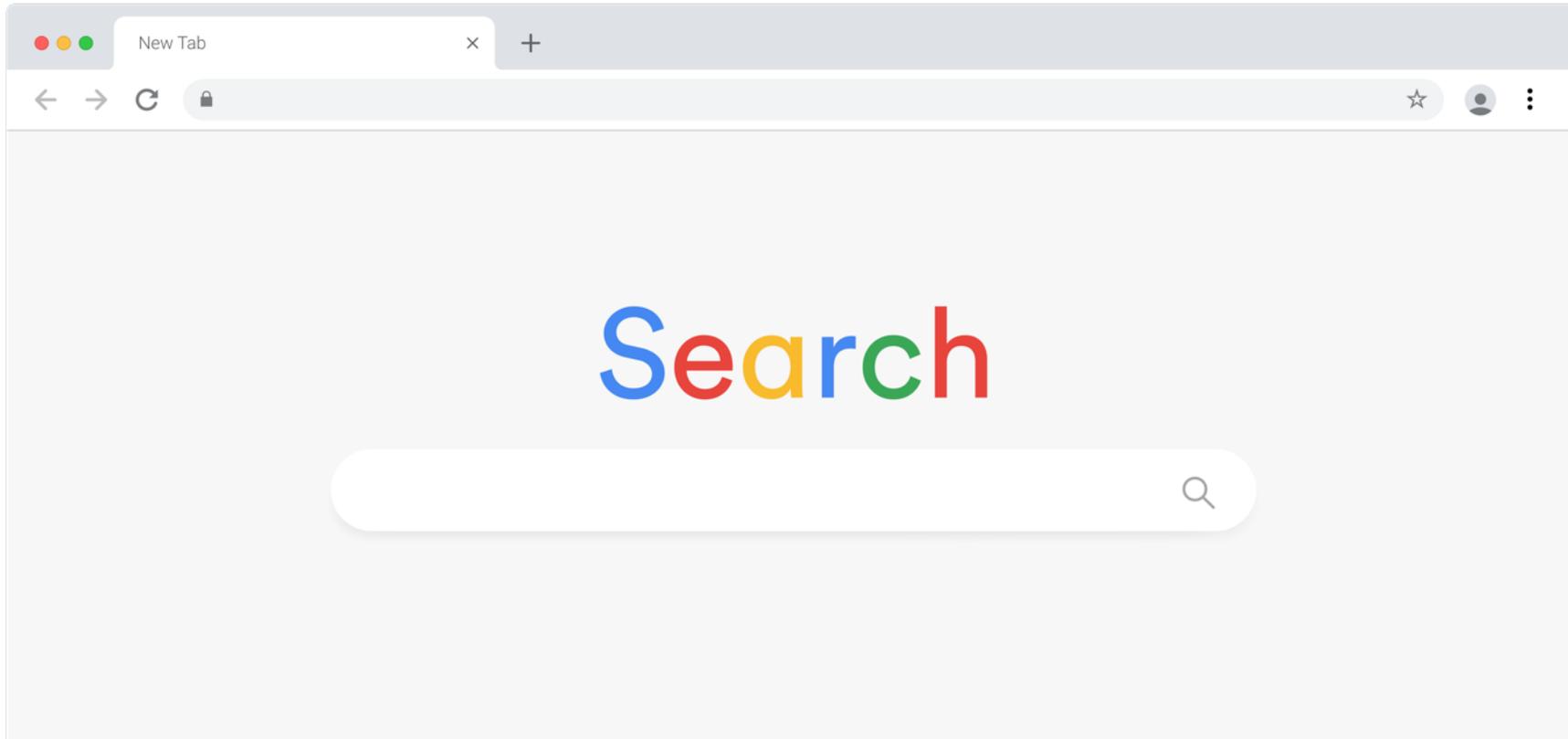
강수경 기자 | 입력 2023.03.20 10:29 | 댓글 0



암 경험자 대상 설문조사 결과 발표...부족 느낀 정보, '치료제와 치료법'  
암 진단 후 가장 궁금했던 정보는 '병원 및 의료진 정보'

# 조사배경

암과 관련된 정보들에 대해 검색했을 때 사람들은 주로 어떤 정보들을 찾아보는가를 알아보던  
과정에서 광고성 게시글의 비중이 높다는 점에 주목함



조사방법

# 데이터 크롤링 & 데이터 마이닝 토픽 모델링



전문가 협업

유 경 한 교수님

전북대학교 / 미디어커뮤니케이션학과



윤 호 영 교수님

이화여자대학교 / 커뮤니케이션·미디어학부

자료의 개요

# 데이터 수집 방법과 데이터의 특성

# 자료의 수집



초기  
건강 관리 생존  
증상  
치료  
검사  
병원 발생  
수술 받다

▲ 7대 암에 공통으로 발견되는 주요 검색어



▲ 개별 암에서 발견되는 특정 검색어

# 데이터 수집

## ▶ 데이터 수집기간

1차

9월 첫째 주 - 둘째 주 (09.07 - 09.16)

2차

10월 둘째 주 - 셋째 주 (10.13 - 10.20)

## ▶ 데이터 수집 범위 및 수집 기준

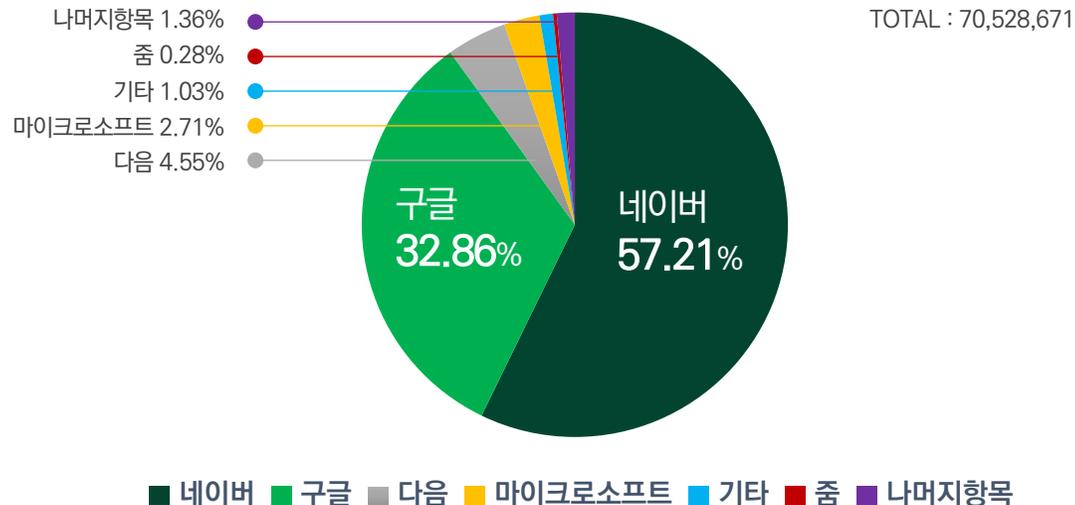
수집 대상 - 네이버, 구글의 검색 페이지 (현색 검색엔진 국내 점유율 1, 2위)

<http://www.internettrend.co.kr/trendForward.tsp>

수집 범위 - 네이버 : 상위 10개 검색페이지 (페이지 당 약 10개 내외)

- 구글 : 검색결과 첫 페이지의 스크롤 허용 범위 내 결과 (첫페이지 기준, 약 100개)

### 국내 주요 포털 검색량 분포



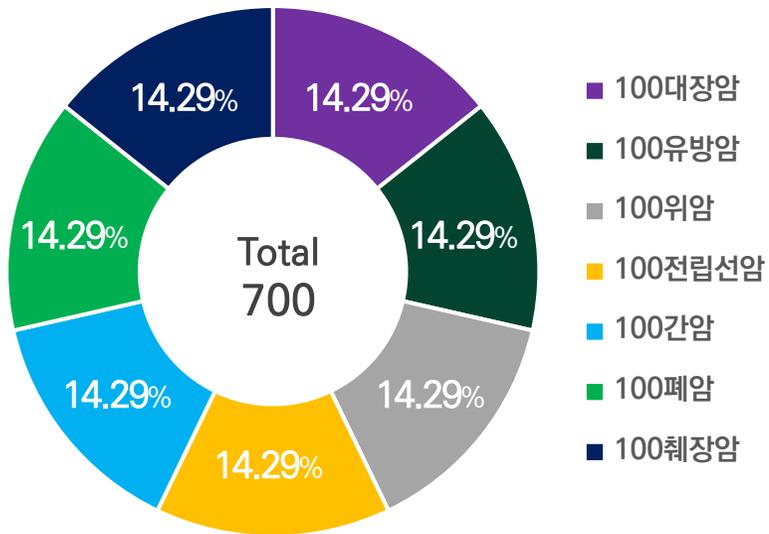
### 수집된 데이터의 분포 (네이버/구글)



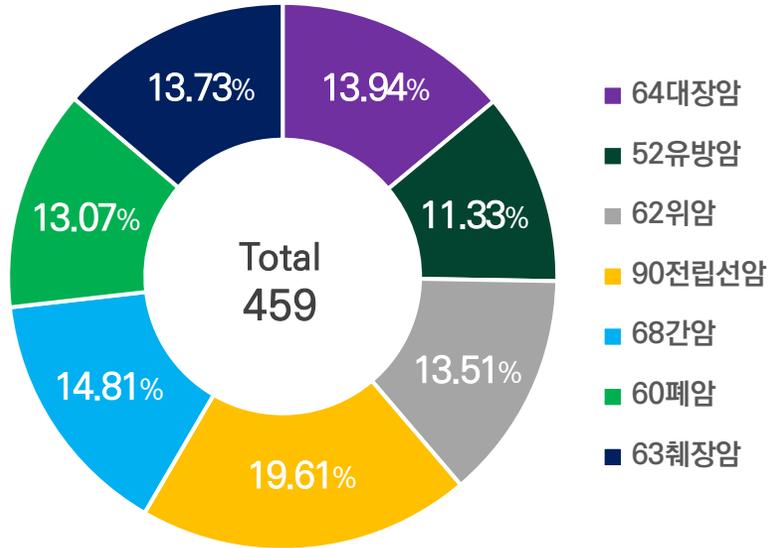
# 데이터 수집

## 7대 암 별 수집 데이터 현황

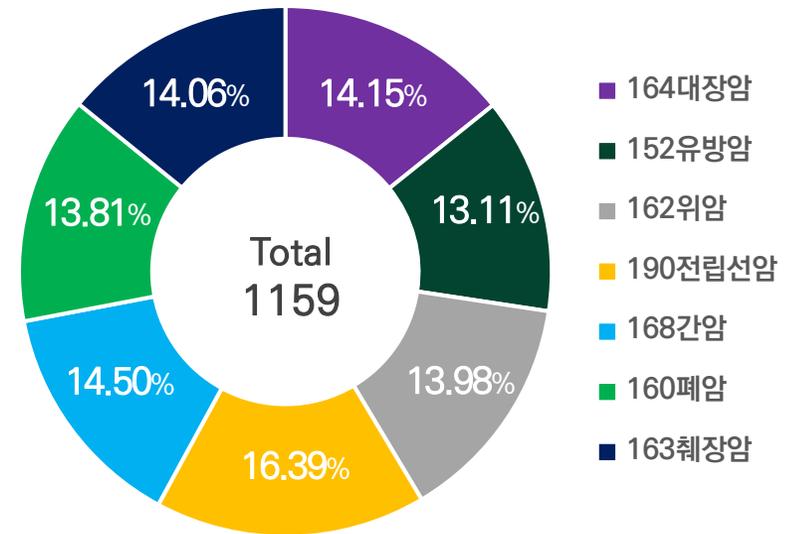
네이버



구글



전체 (네이버+구글)



## 분석 데이터

최종 내용 분석 대상

# 919게시글

**NAVER** 598건

**Google** 321건

▶ 내용 확인이 불가능한 포스팅의 예시

1) 회원가입을 해야 확인 가능한 커뮤니티 게시글

---

2) 기간이 지나 삭제되거나 내용이 없는 글

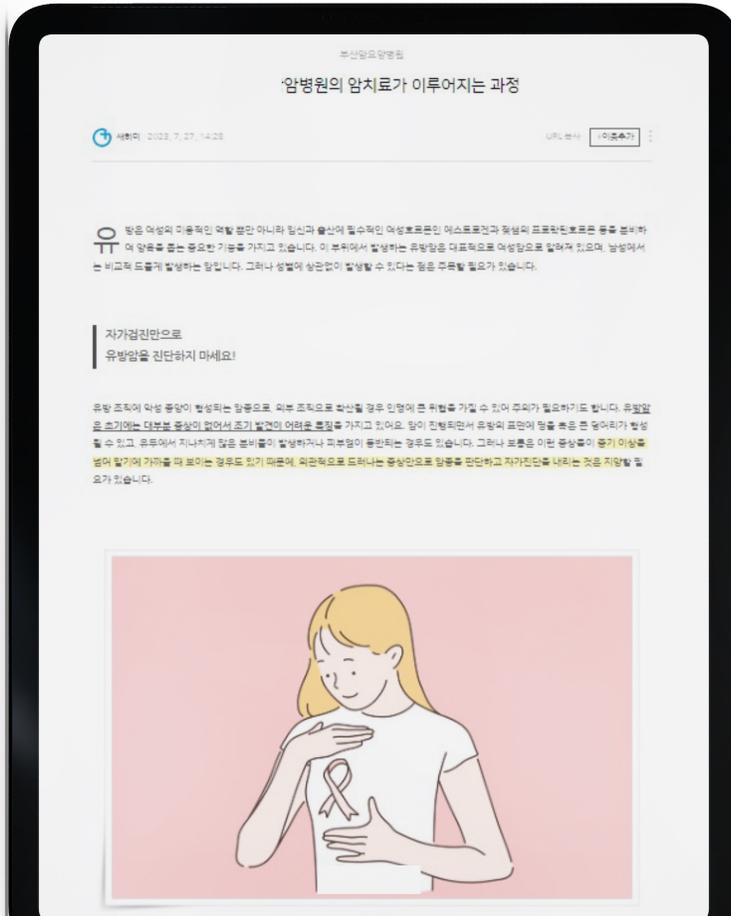
---

3) PDF, 첨부파일 형태의 포스팅 글

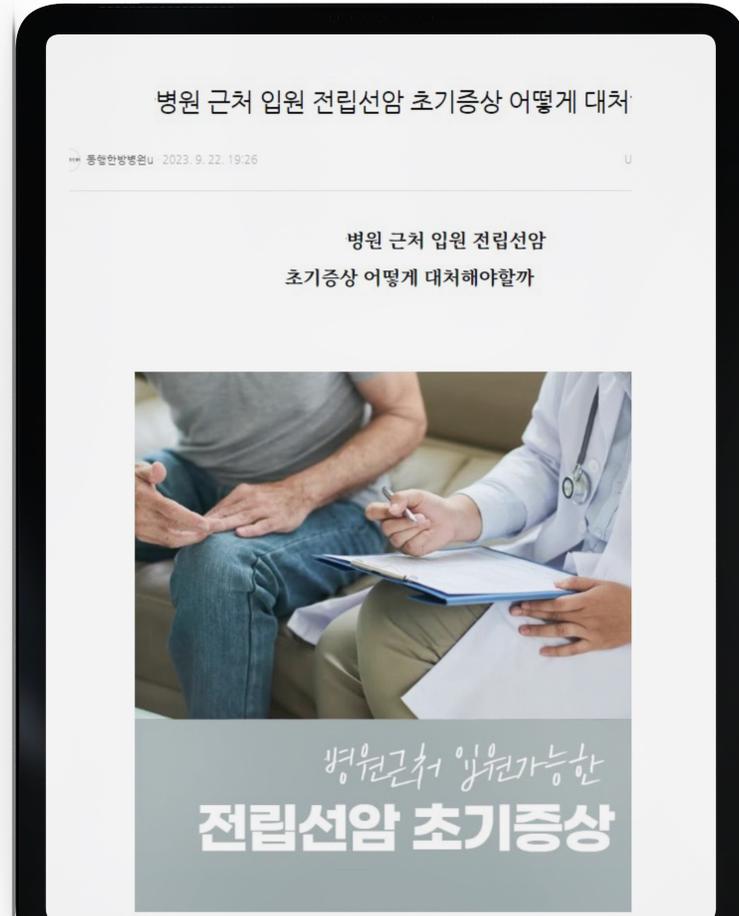
# 네이버와 구글의 광고성 포스팅 분석

# 광고성 게시글 분류 기준

▼ 블로그 게시글 중에 병원 홍보 및  
광고의 목적이 대부분인 글

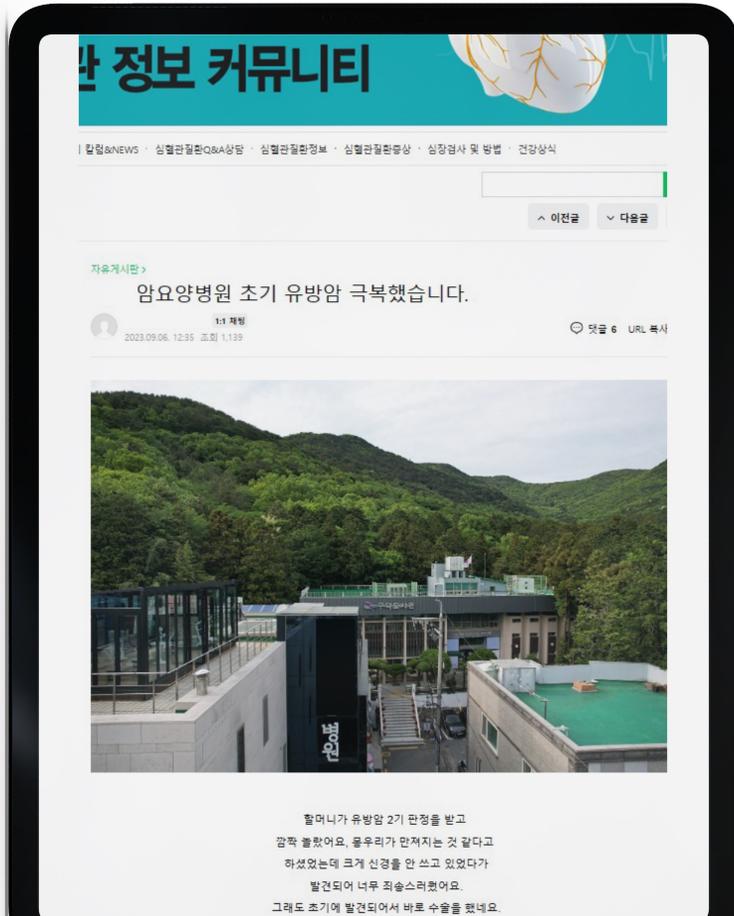


▼ 블로그 게시글로 병원 홍보 및  
광고 목적을 일부 포함하는 글

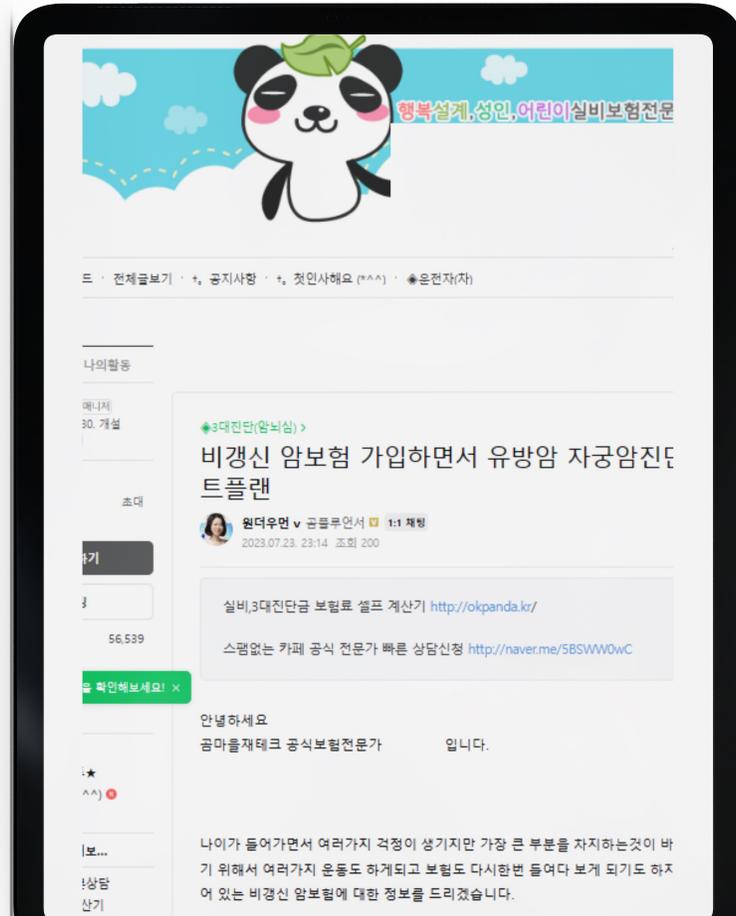


# 광고성 게시글 분류 기준

▼ 카페/커뮤니티 게시글 중에 광고의 목적이 대부분인 글



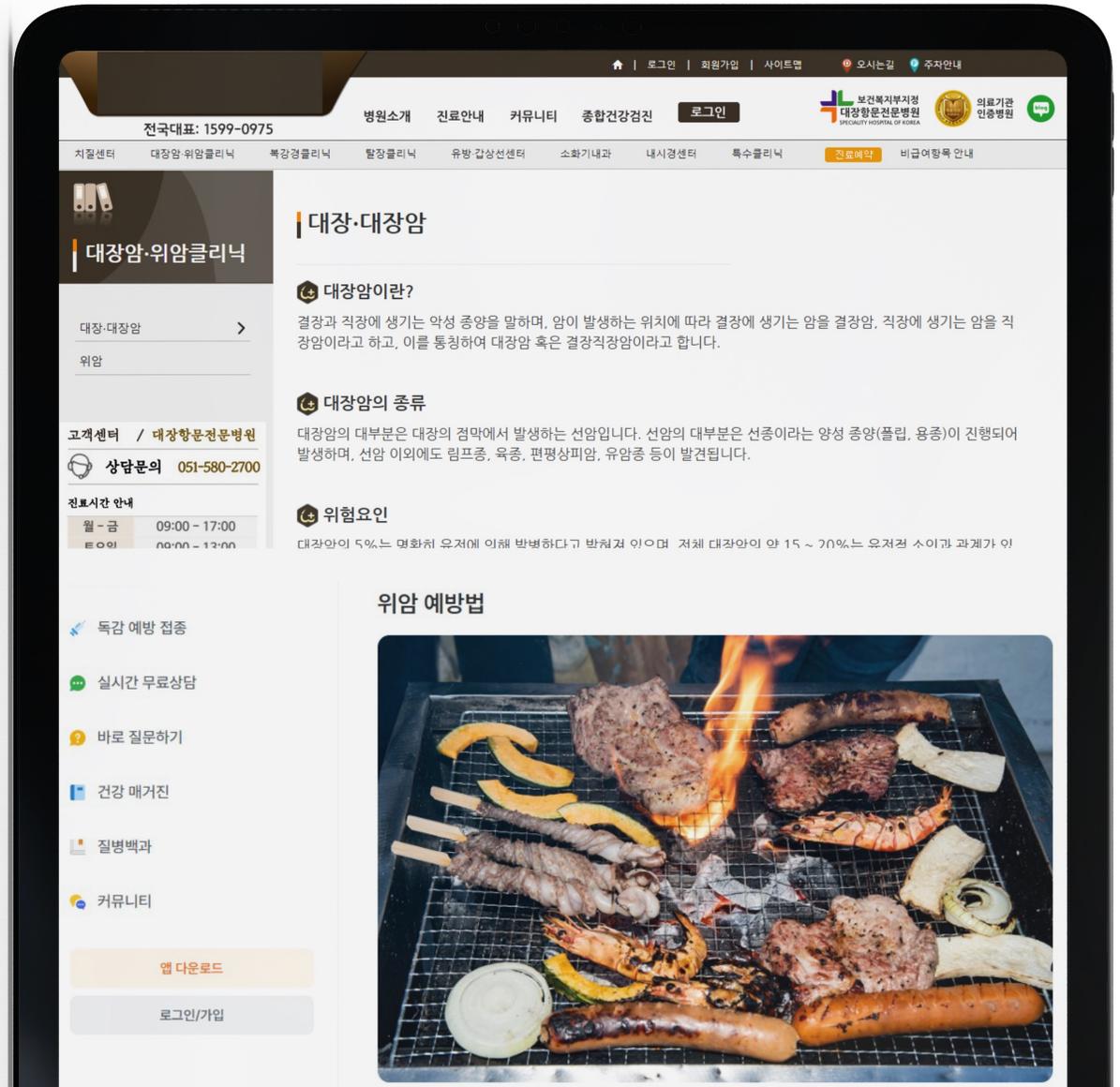
▼ 커뮤니티에 암투병 경험 공유를 빙자하여 병원 홍보 및 광고하는 게시글



# 광고성 게시글 분류 기준

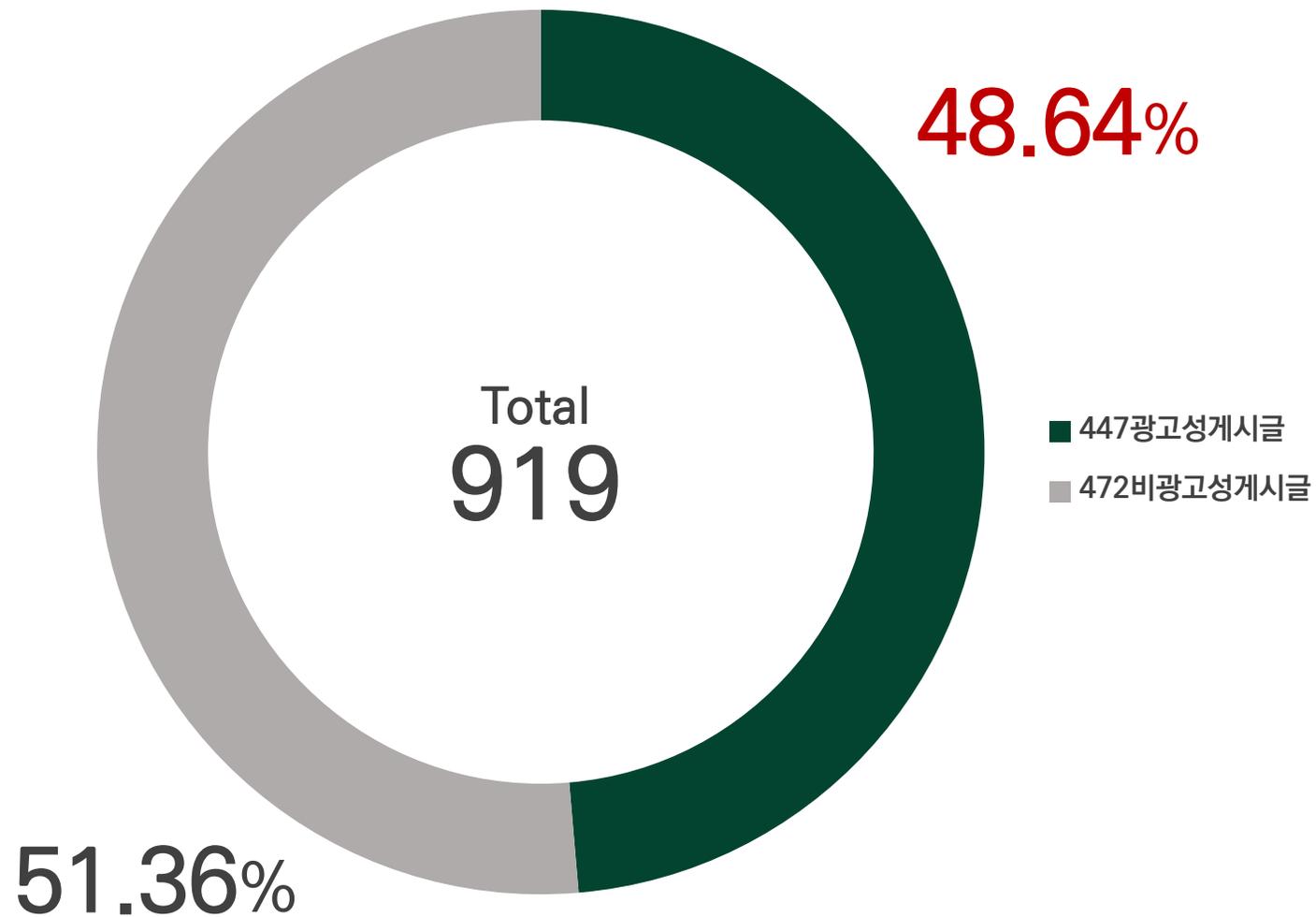
▶ 병원 웹사이트 게시글 중에 병원 홍보,  
광고 목적이 포함된 글

▶ 기사 형식을 빌려 정보를 제공하고 시작 또는  
말미에 광고 홍보글을 게시하는 형태



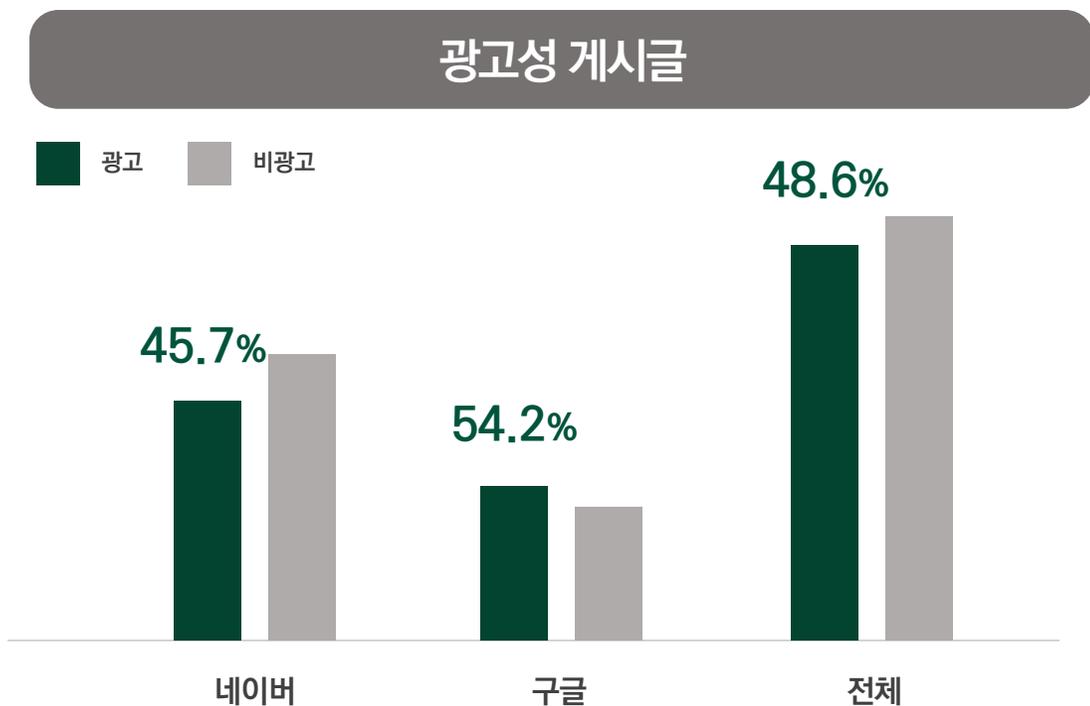
## 광고성 포스팅

광고성 포스팅은 전체의 48.6%로 포털에서 검색을 통해 접할 수 있는 글 중 약 전반에 가까운 게시글이 광고성 포스팅인 것으로 조사됨.



# 광고성 포스팅

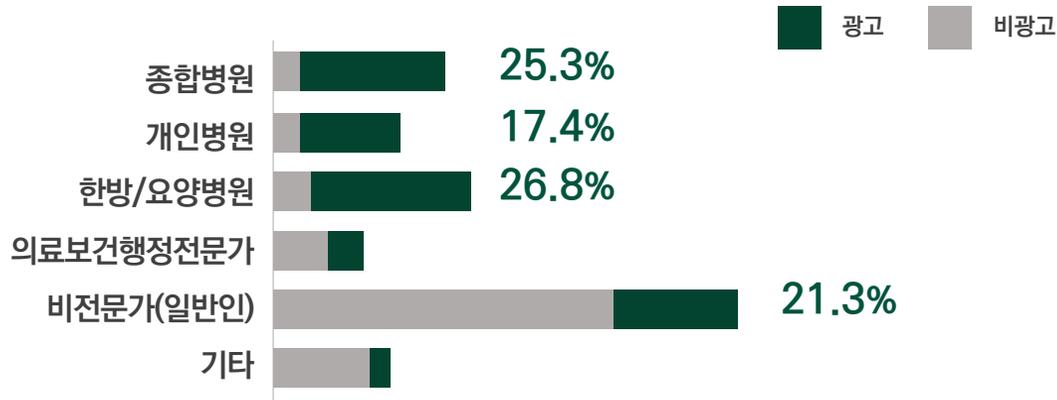
광고성 포스팅을 포털 사이트별로 조사한 결과, 네이버는 45.7%, 구글은 54.2%가 광고성 게시글로 나타나, 구글의 광고성 글의 비중이 더 높게 나타남



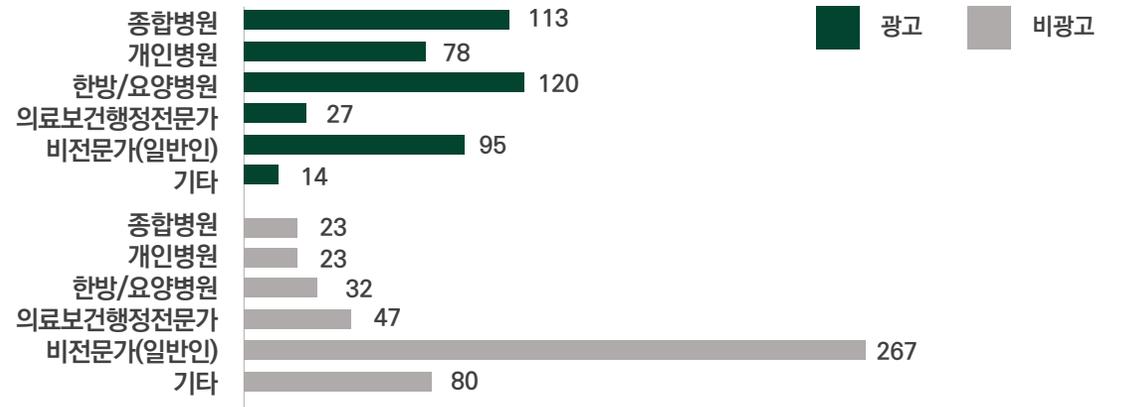
		광고	비광고	전체
네이버	빈도	273	325	598
	site 중 %	45.7%	54.3%	100.0%
구글	빈도	174	147	321
	site 중 %	54.2%	45.8%	100.0%
전체	빈도	447	472	919
	site 중 %	48.6%	51.4%	100.0%

# 광고성 글의 주요 게시자

## 광고성 글의 주요 게시자 1



## 광고성 글의 주요 게시자 2

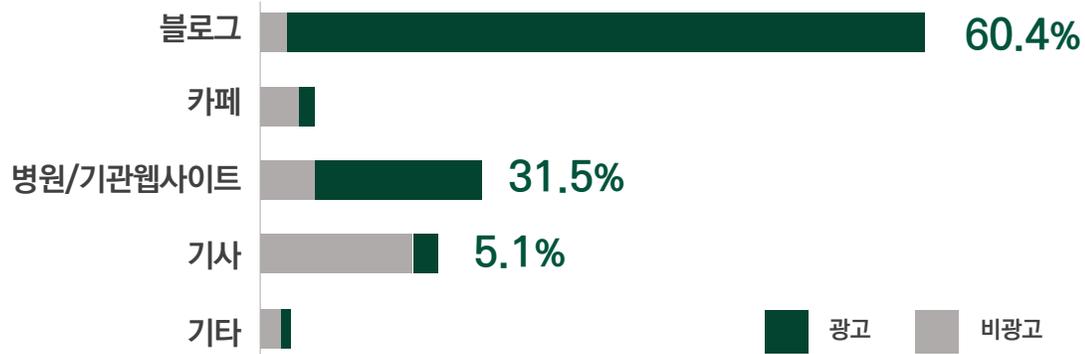


광고성 게시글을 가장 많이 게시하는 행위자를 유형별로 살펴본 결과, **한방요양병원이 26.8%로 가장 많았음.**

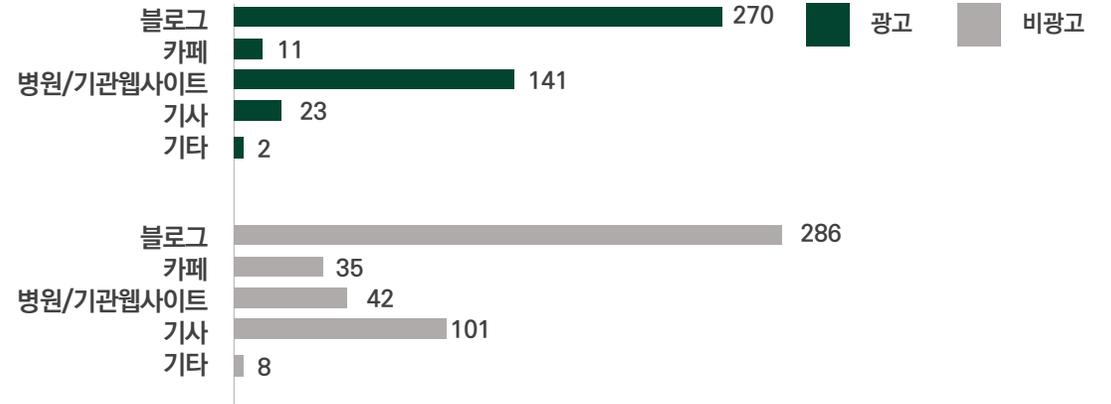
대학종합병원 25.3%, 비전문가(일반인) 21.3%, 중소 개인병원 17.4% 순으로 조사됨.

# 광고성 글의 주요 게시경로

광고성 글의 주요 게시경로



광고성 글의 주요 게시경로



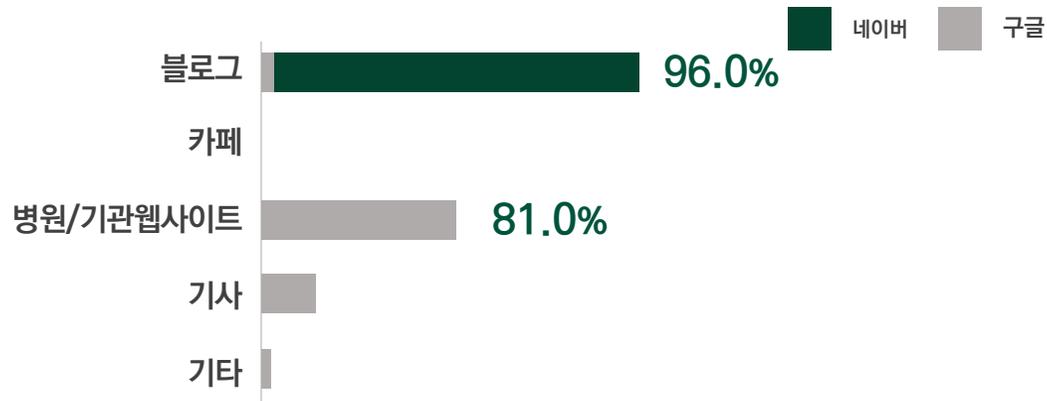
광고성 글이 가장 많이 게시되는 경로는 **블로그**

전체 광고성 게시글의 60.4%였고, 병원/기관 웹사이트가 31.5%, 언론보도 5.1% 순으로 조사됨.

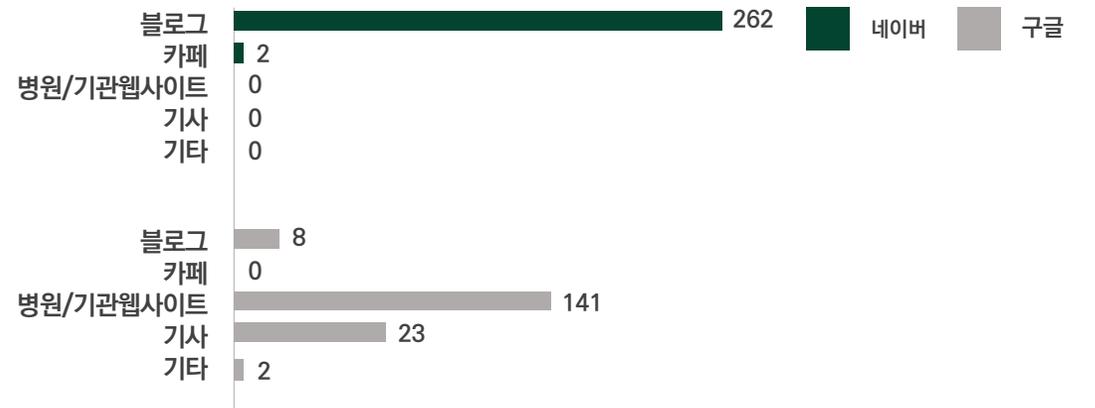
→ 블로그를 통한 광고성 게시글이 압도적으로 많고, 네이버 검색을 통한 블로그 검색 비중이 높아 대책이 요구됨

# 광고성 포스팅 (네이버 vs 구글)

## 광고성 포스팅 (네이버 vs 구글)



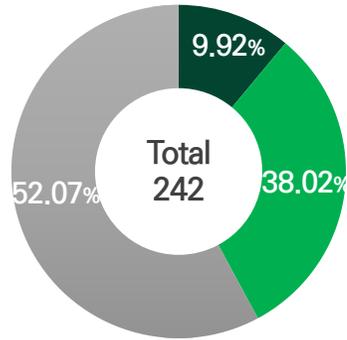
## 광고성 포스팅 (네이버 vs 구글)



포털사이트 마다 검색시 발견되는 광고성 게시글의 경로가 다름  
네이버 검색시 블로그에 광고성 게시글이 압도적으로 많고(96.0%), 구글 검색시 병원/기관웹사이트가 압도적으로 높음(81.0%)

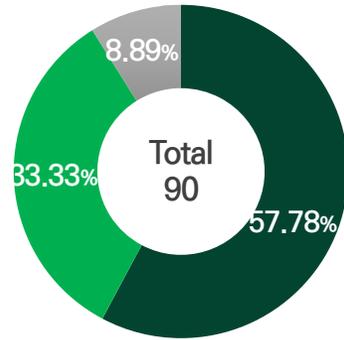
# 광고성 게시글의 홍보 내용 비중

네이버



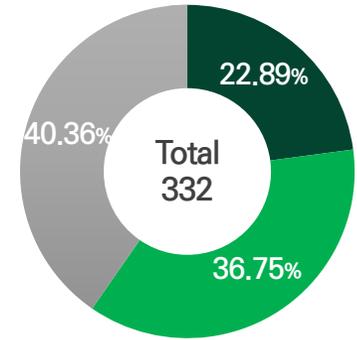
■ 24단순표시 ■ 92일부광고 ■ 126전체광고

구글



■ 52단순표시 ■ 30일부광고 ■ 8전체광고

전체



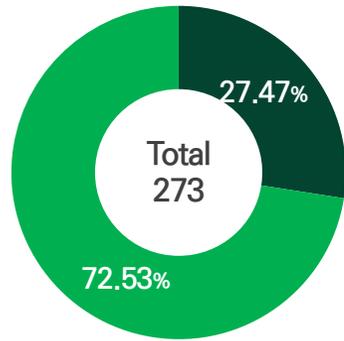
■ 76단순표시 ■ 122일부광고 ■ 134전체광고

네이버 : 전체광고 >> 일부광고 >> 단순표시    구글: 단순표시 >> 일부광고 >> 전체 광고

광고성 포스팅 중 글자수 카운트가 가능한 332건을 대상으로 분석한 결과, 네이버는 전체 광고의 비중이 52.1%로 가장 높았고, 구글은 단순 광고 표시가 57.78%로 가장 높게 조사됨.  
네이버에서 암 정보를 검색할 경우 전체 광고인 게시글을 접할 가능성이 매우 높음

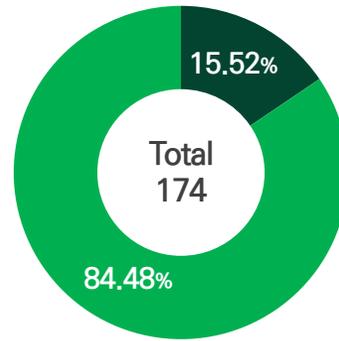
## 광고성 게시글의 전문 데이터 활용 여부

네이버



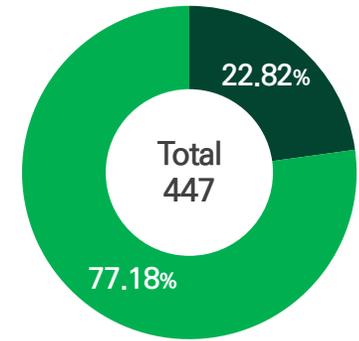
■ 75있음 ■ 198없음

구글



■ 27있음 ■ 147없음

전체



■ 102있음 ■ 345없음

학술논문이나 검증된 데이터를 소개 또는 사용하는 경우는 22.8%로, 낮은 수준으로 조사됨.

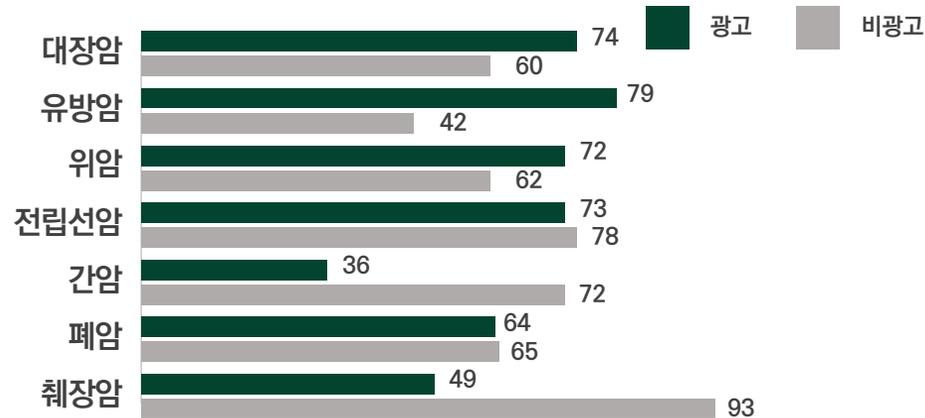
구글(15.5%)보다 네이버의 광고성 게시글(27.5%)에서 더 많이 발견됨

→ 학술논문이나 검증데이터를 사용하면 정보의 정확성과 신뢰도를 높일 수 있음

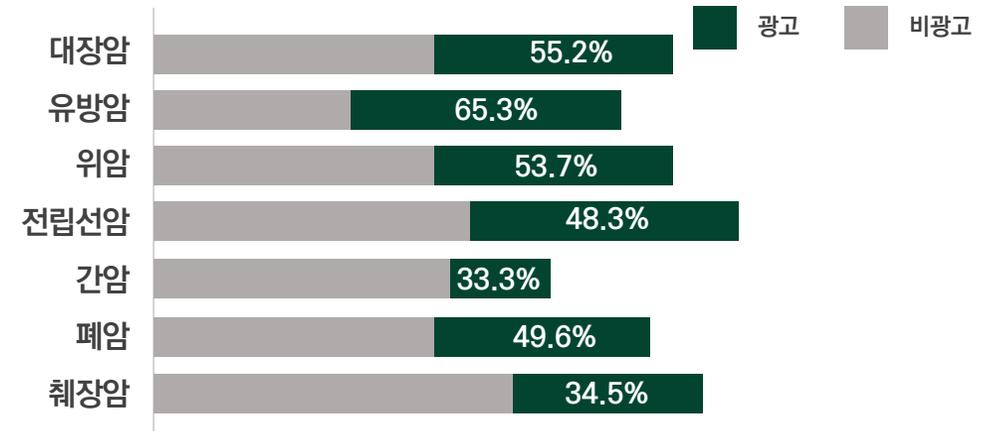
# 7대 암별 광고성 게시글의 비중

# 7대 암별 광고성 게시물

7대 암별 광고성 게시물



7대 암별 광고성 게시물



7대 암별 광고성 게시물 분석결과, 유방암의 광고성 게시물이 가장 많고(65.3%), 대장암(55.2%), 위암(53.7%) 순으로 조사됨.

반면, 간암(33.3%), 췌장암(34.5%)은 광고성 게시물이 상대적으로 적음

광고도 순수 광고 내용만 채우지 않고  
매우 기초적이고 원론적인 정보를 제공하는 식의 유도를 통해  
본래 광고 내용으로 이끌고 내려가는 방식임



광고 콘텐츠의 내용을 보다 구체적으로 살펴봄으로써  
기초 정보 내용을 벗어나는 유도 형식과 광고 내용의 비율을 살펴보기로 함

# 토픽 모델링

토픽모델링은 각 단어들이 통계적으로 특정 토픽에 포함될 확률을 파악하여 문서의 주제를 추정하는 기법

## Topics

gene	0.04
dna	0.02
genetic	0.01
...	

life	0.02
evolve	0.01
organism	0.01
...	

brain	0.04
neuron	0.02
nerve	0.01
...	

data	0.02
number	0.02
computer	0.01
...	

## Documents

### Seeking Life's Bare (Genetic) Necessities

COLD SPRING HARBOR, NEW YORK— How many **genes** does an **organism** need to **survive**? Last week at the genome meeting here,\* two genome researchers with radically different approaches presented complementary views of the basic genes needed for **life**. One research team, using **computer** analyses to compare known **genomes**, concluded that today's **organisms** can be sustained with just 250 genes, and that the earliest life forms required a mere 128 **genes**. The other researcher mapped genes in a simple parasite and estimated that for this organism, 800 genes are plenty to do the job—but that anything short of 100 wouldn't be enough. Although the numbers don't match precisely, those **predictions**

"are not all that far apart," especially in comparison to the 75,000 **genes** in the human genome, notes Siv Andersson of Uppsala University in Sweden, who arrived at the 800 number. But coming up with a consensus answer may be more than just a **genetic numbers** game, particularly as more and more **genomes** are completely mapped and sequenced. "It may be a way of organizing any newly **sequenced genome**," explains Arcady Mushegian, a **computational** molecular biologist at the National Center for Biotechnology Information (NCBI) in Bethesda, Maryland. Comparing an

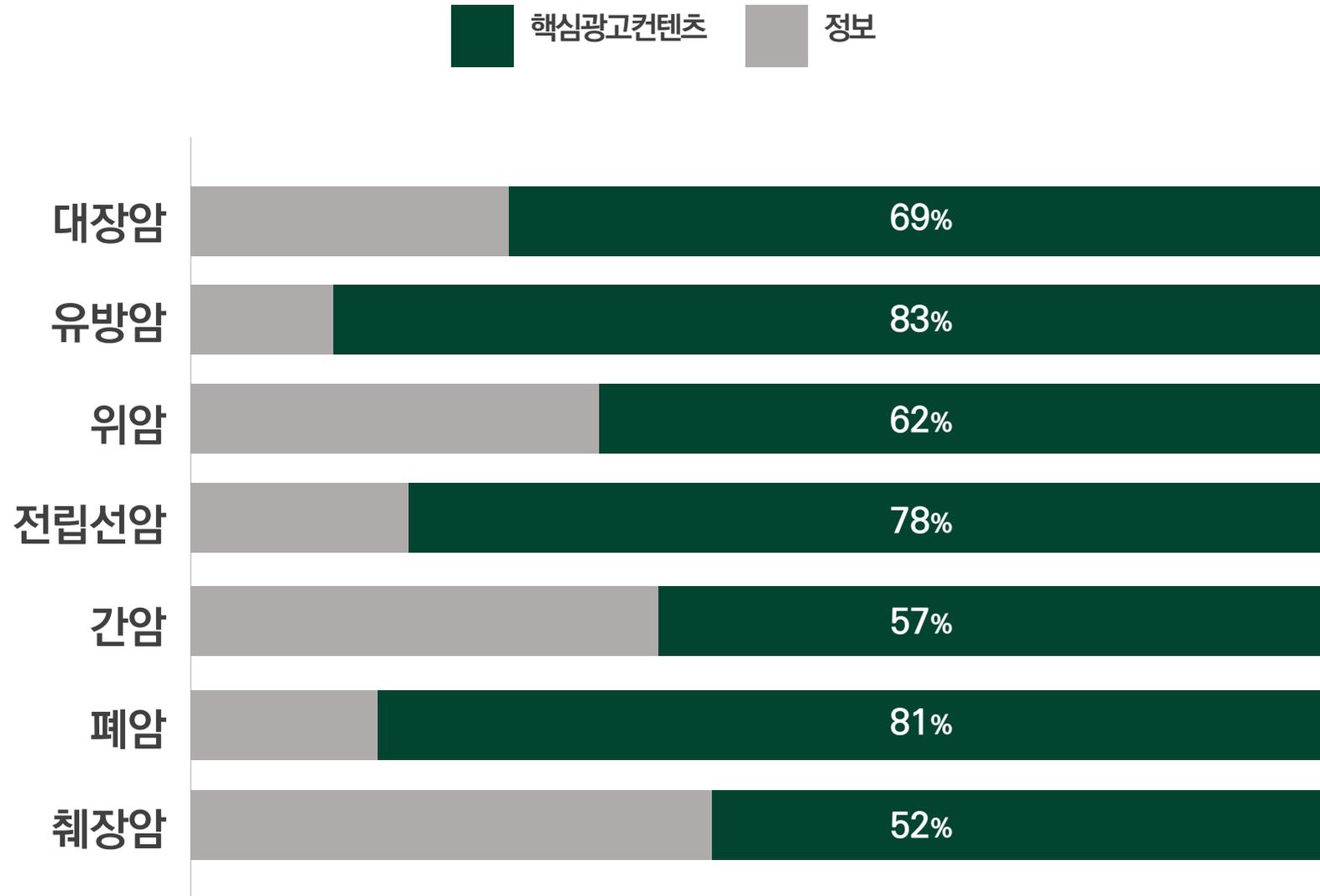
\* Genome Mapping and Sequencing, Cold Spring Harbor, New York, May 8 to 12.

Stripping down. Computer analysis yields an estimate of the minimum modern and ancient genomes.

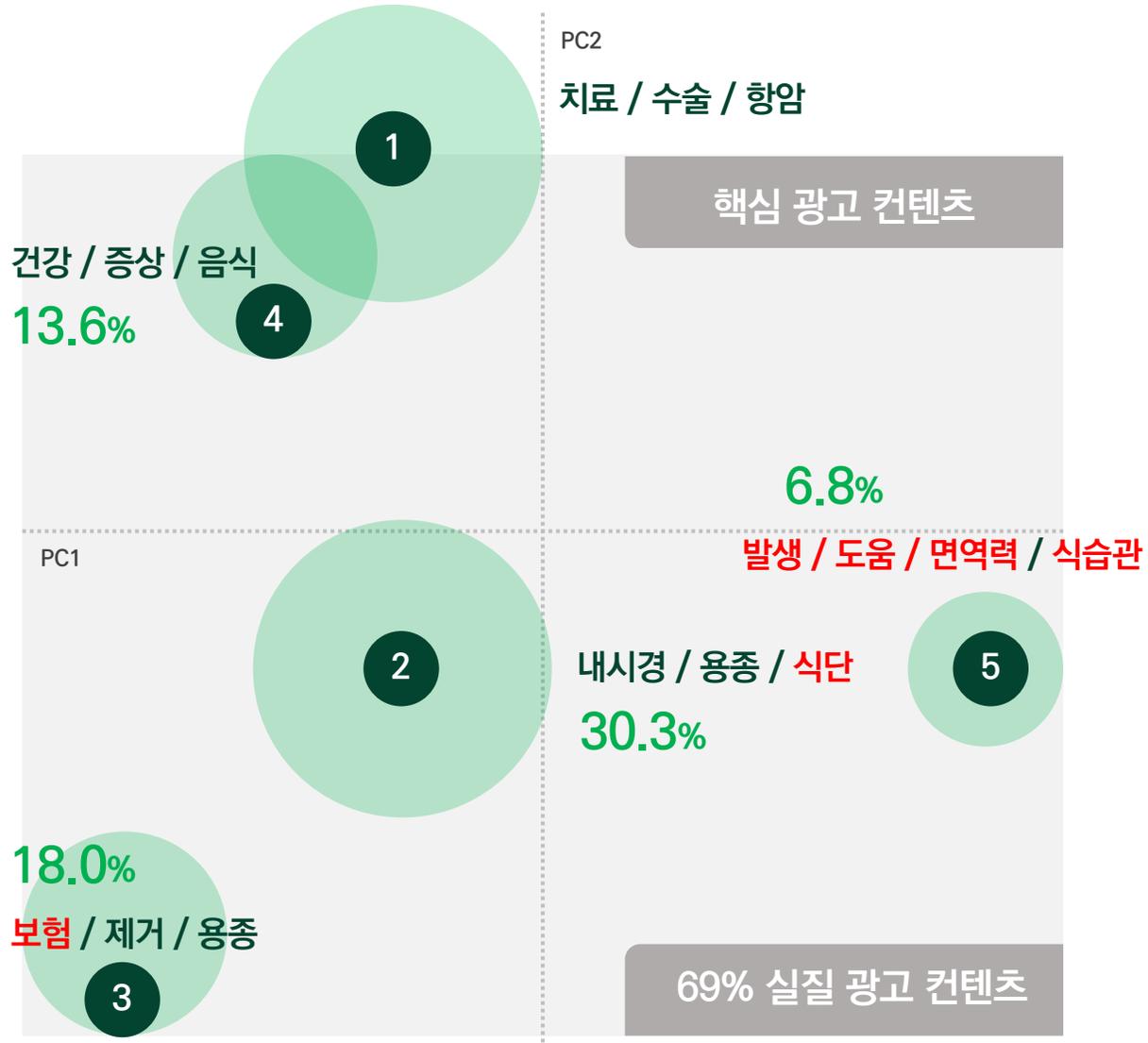
SCIENCE • VOL. 272 • 24 MAY 1996

## Topic proportions & assignments

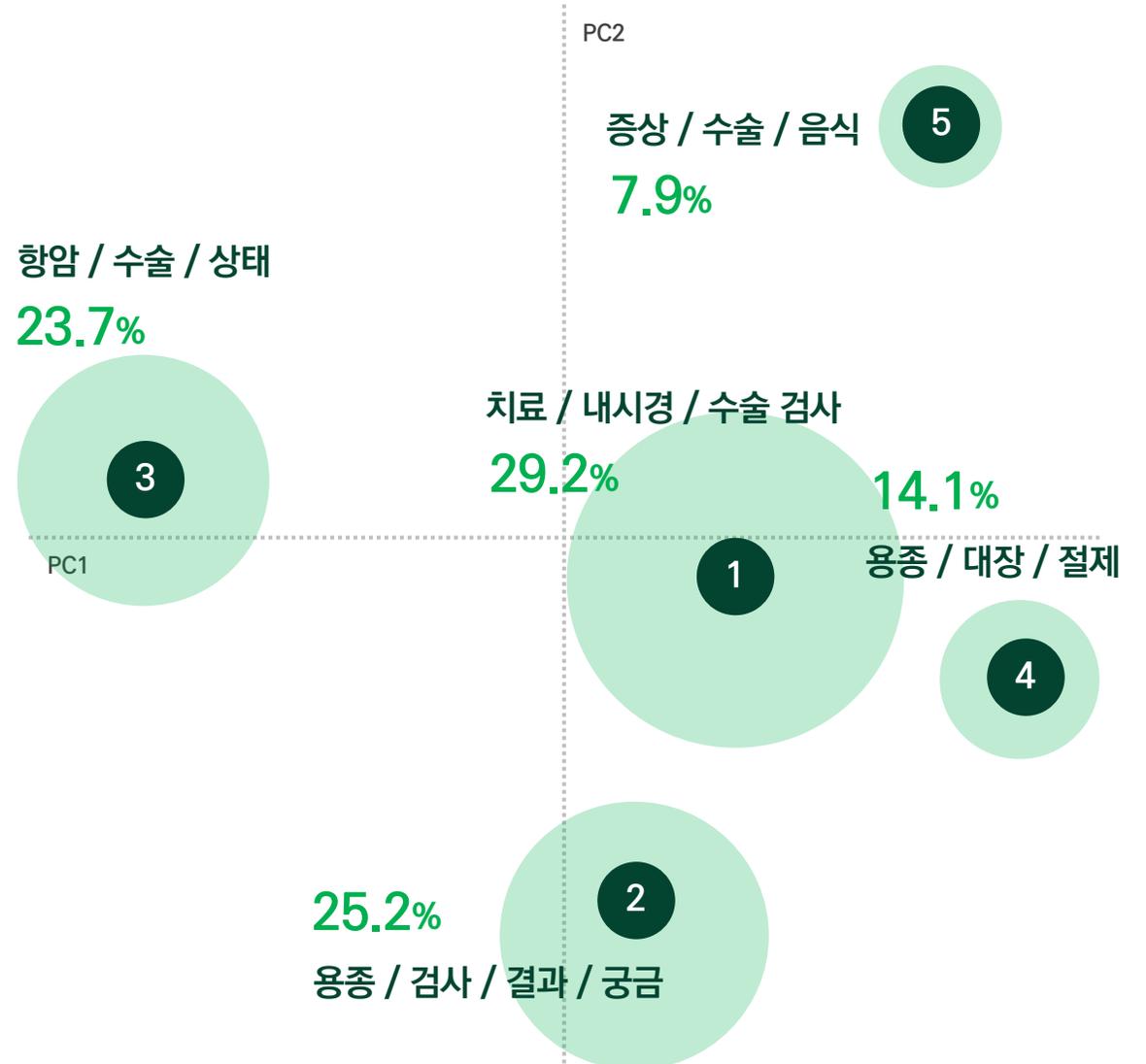
## 7대 암종별 광고성 게시글의 토픽모델링 분석결과



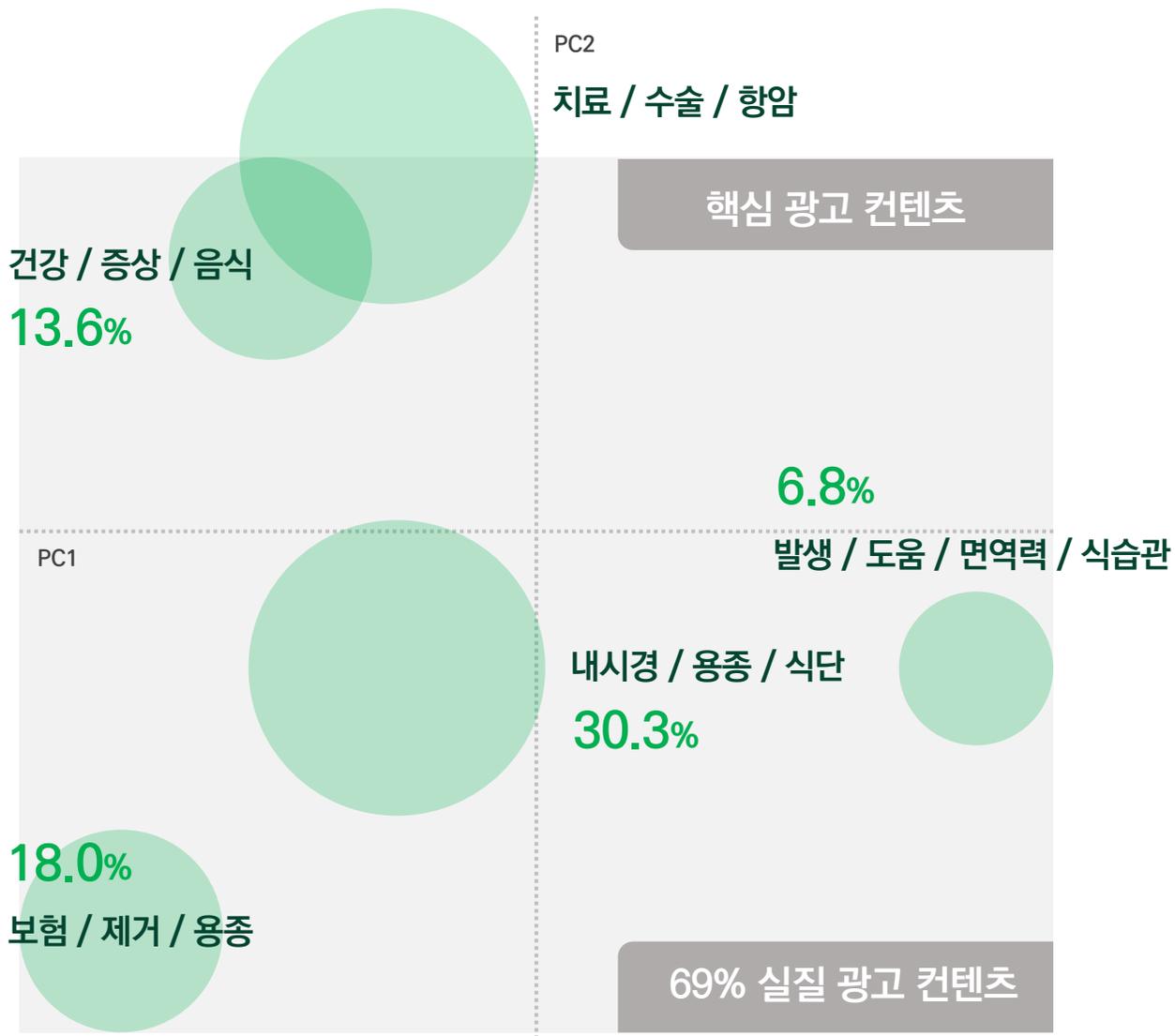
## 광고 콘텐츠 : 대장암



## 비광고 콘텐츠 : 대장암



# 광고 콘텐츠 : 대장암



## 광고 형식 예시

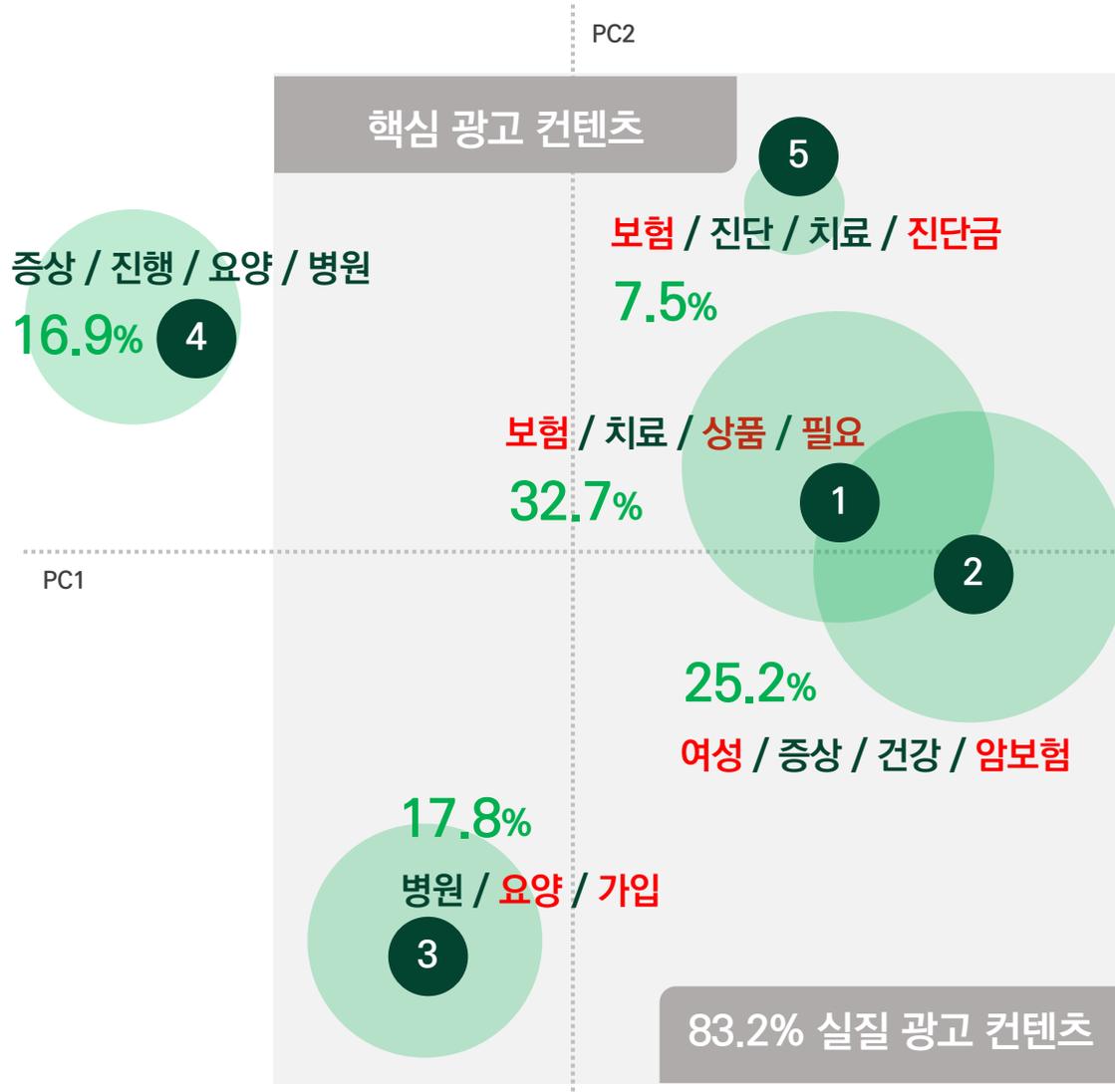
암 발생위치에 따라 증상도 다양해서 혹시나 하는 사이에 암이 진행될 수 있습니다. 대장암 위치에 대한 증상과 원인 예방법에 대해 알아보겠습니다. 정기적인 대장 내시경 검사나 가족성 용종증이 있는 분들은 유전자 이상소견을 검사해서 초기에 종양을 치료하거나 예방하는 방법도 있습니다. 평소 식습관 조절과식이섬유식을 권장합니다.

...

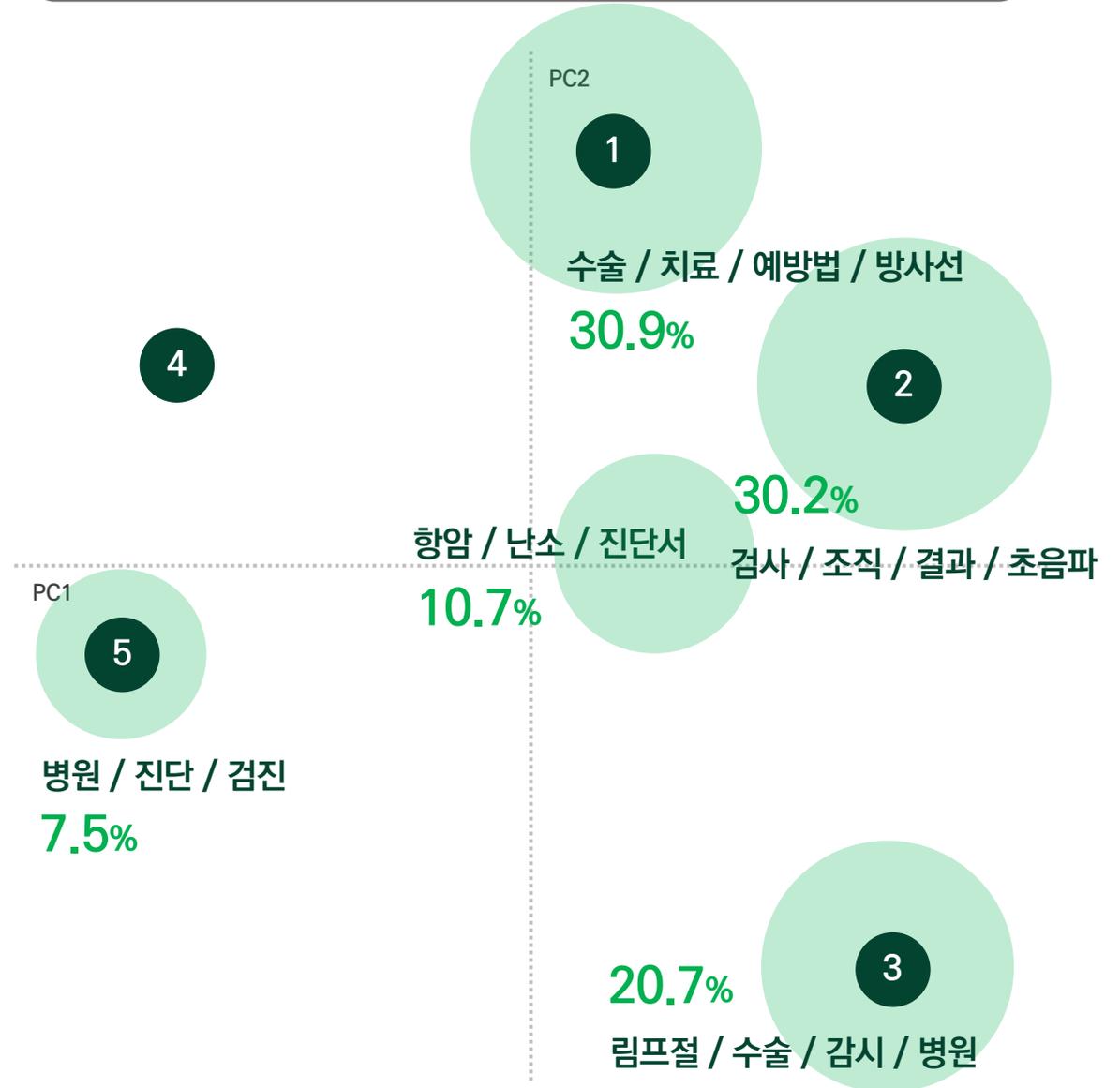
초기 증상을 발견하기 위해 변상태를 평소에 잘 체크하시고 정기적인 건강검진을 통해 관리하시는 것이 중요합니다. 암후유증 치료를 위한 \*\*요양병원. \*\*요양병원에서 암수술 후 회복치료와 후유증 치료를 도와드립니다.

## 설명이후 광고 형식

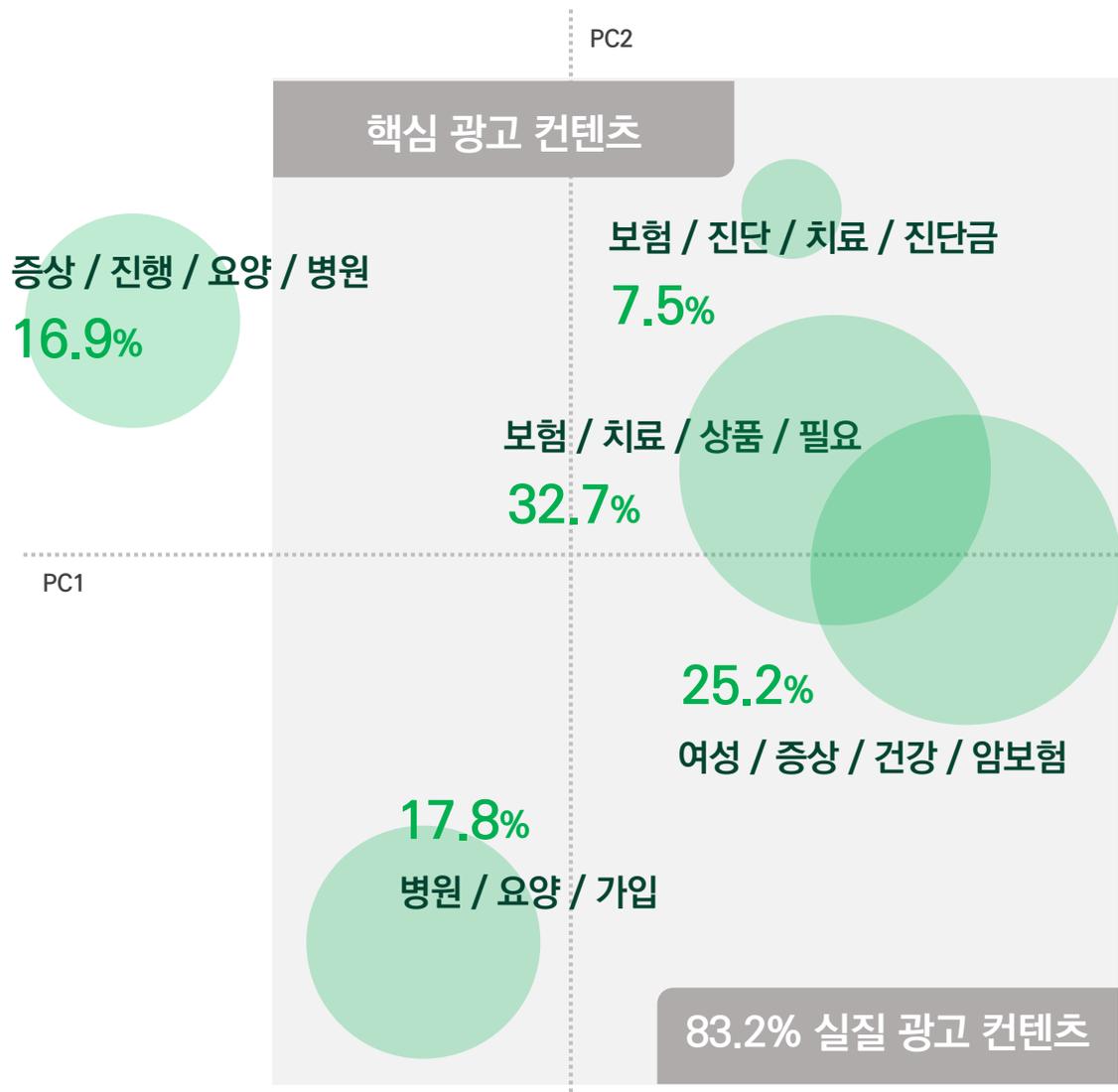
## 광고 콘텐츠 : 유방암



## 비광고 콘텐츠 : 유방암



# 광고 콘텐츠 : 유방암



## 광고 형식 예시

확진이 되었지만, 처음에 발견을 하여 입원하여 치료를 제대로 받으면 완치될 수 있다고 말했어요. ... 친구한테 질문해서 유방암 진단금 보험을 가지고 있어서 종양치료비 걱정을 하지 않고 보험금으로 치료를 하면서 집중해 볼 수 있다고 했어요. ... 상품을 하나쯤 가져야 겠다고 마음도 있었습니다.

여러분도 필요한 사람은 활용해 보시죠.  
필요한 보험이 무엇인지 지금 바로 여기서 알아보세요.

실제 정보 설명없는 상품 광고 형식

## 광고 콘텐츠 : 폐암

### 핵심 광고 콘텐츠

치료 / 관리 / **면역력** / 진행  
**1** 39.5%

**2** 예방 / 수술 / **면역** / **서비스**  
 23.4%

치료 / 증상 / **병행** / **통합** / 악성  
 18.1%

**3**

81.0% 실질 광고 콘텐츠

PC2

발생 / 산재 / 세포 / 치료  
 12.2%

**4**

**5**  
 6.8%

세포 / 치료 / 진단 / 치유

PC1

## 비광고 콘텐츠 : 폐암

흡연 / 위험 / 발생 / 담배

11.7%

**4**

**1**  
 세포 / 치료 / 연구 / 종류  
 36.0%

PC1

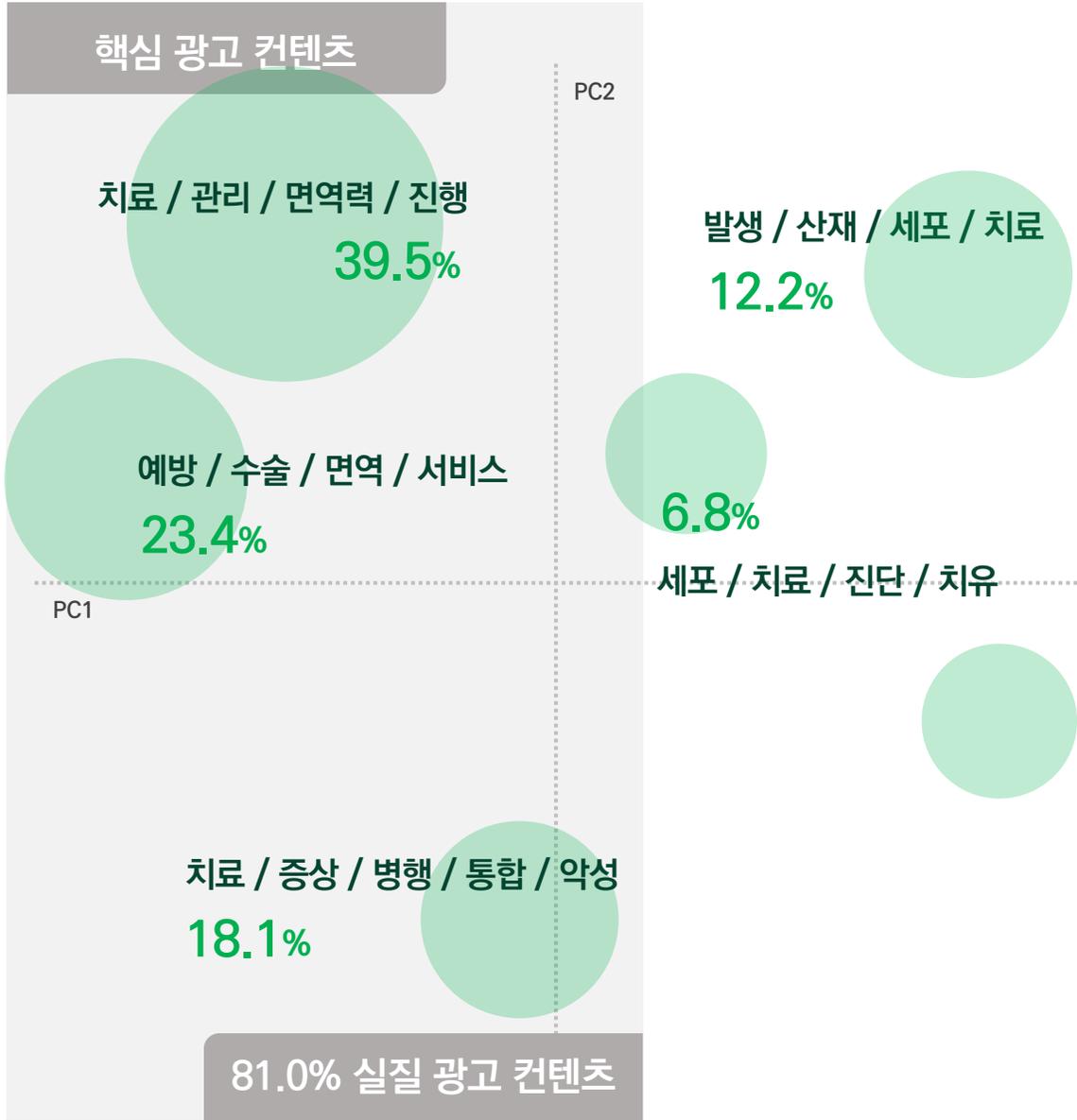
PC2

29.4%  
**2**  
 발견 / 수술 / 세포 / 항암 / 생존율

건강 / 증상 / 종양 / 검사 / 음식  
 22.8%

**3**

# 광고 콘텐츠 : 폐암



## 광고 형식 예시

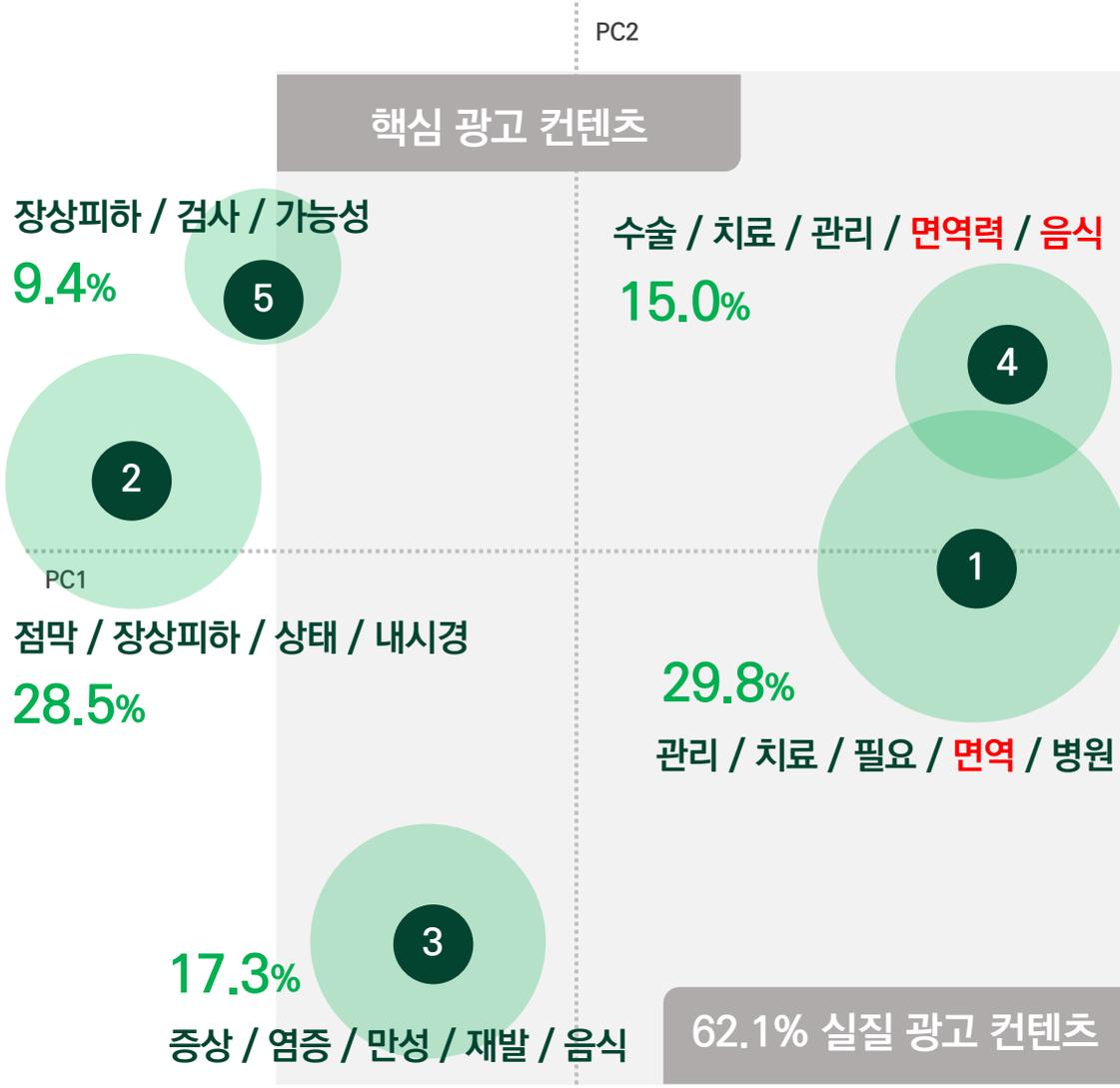
폐암은 폐에서 발생하는 암 종류를 가리키며, 주로 흡연과 관련이 있었어요. 폐암은 조기에 발견되지 않을 경우 주변 조직이나 림프절로 확산되어 치료가 어려워지는 경우가 많았어요. 폐암 1기는 폐암 조직이 작아 조기 발견 및 수술로 종양을 완전히 제거할 수 있는 경우, 5년 생존율이 약 60%에서 80%로 보고되었어요. 면역력이 떨어진 폐암 환우에게 면역력을 높여주는 다양한 주사치료를 써볼 수 있었어요. 미슬토, 싸이모신알파1, 셀레늄, 이문셀, 고용량비타민C 주사 등 이러한 면역치료제를 사용함으로써 암환우의 면역력을 높여주고 폐암 증상 완화를 줄여주는데 효과를 볼 수 있었어요.

...

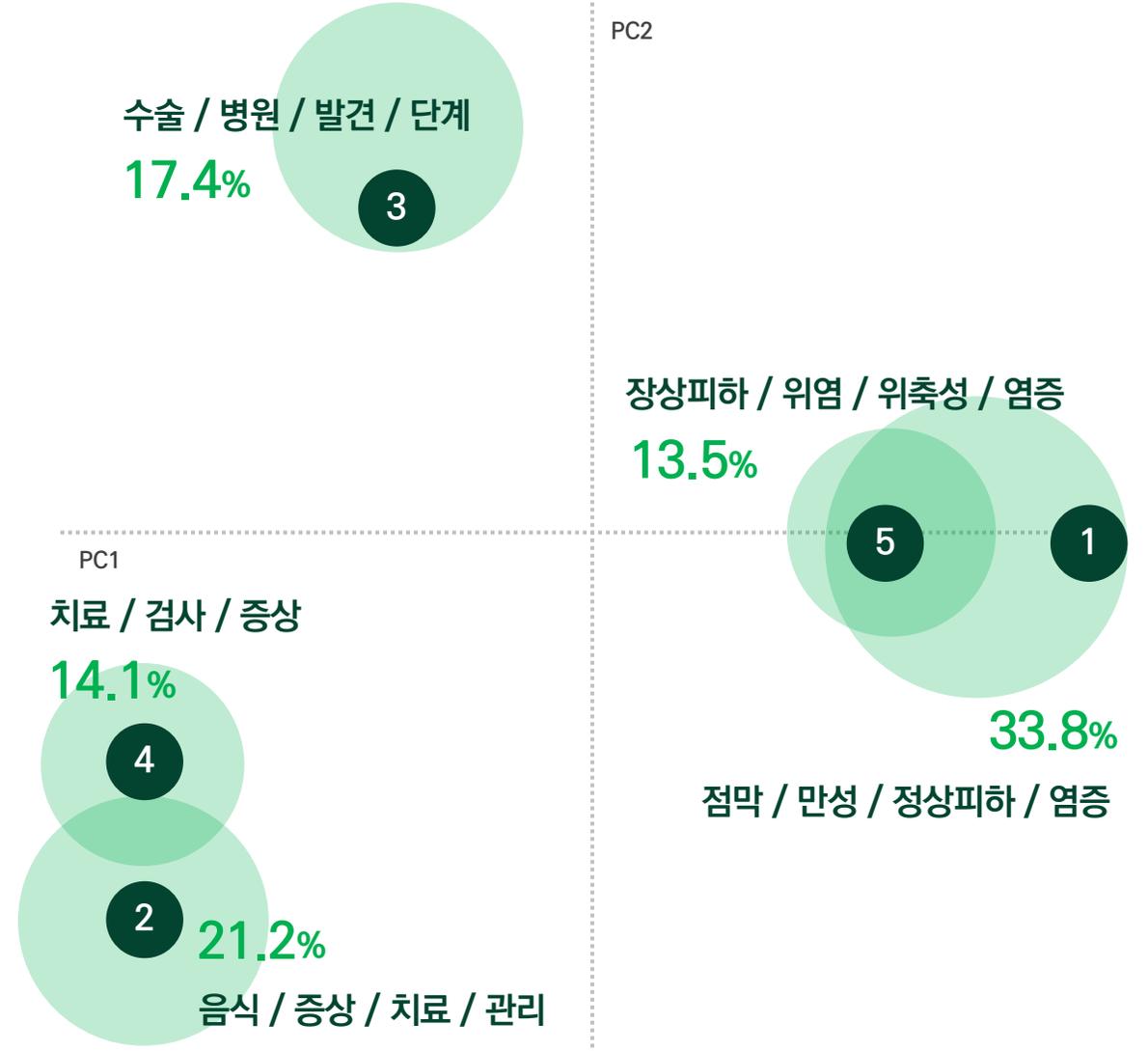
이 과정에서 면역력이 저하되면서 림프부종이나 탈모, 구토, 구역, 입맛의 변화와 같은 여러가지 부작용을 겪을 수 있었기 때문에 부작용 완화에 도움이 되는 면역치료 프로그램을 이용하여 면역력을 높이는데 도움을 받았어요.

## 설명이후 광고 형식

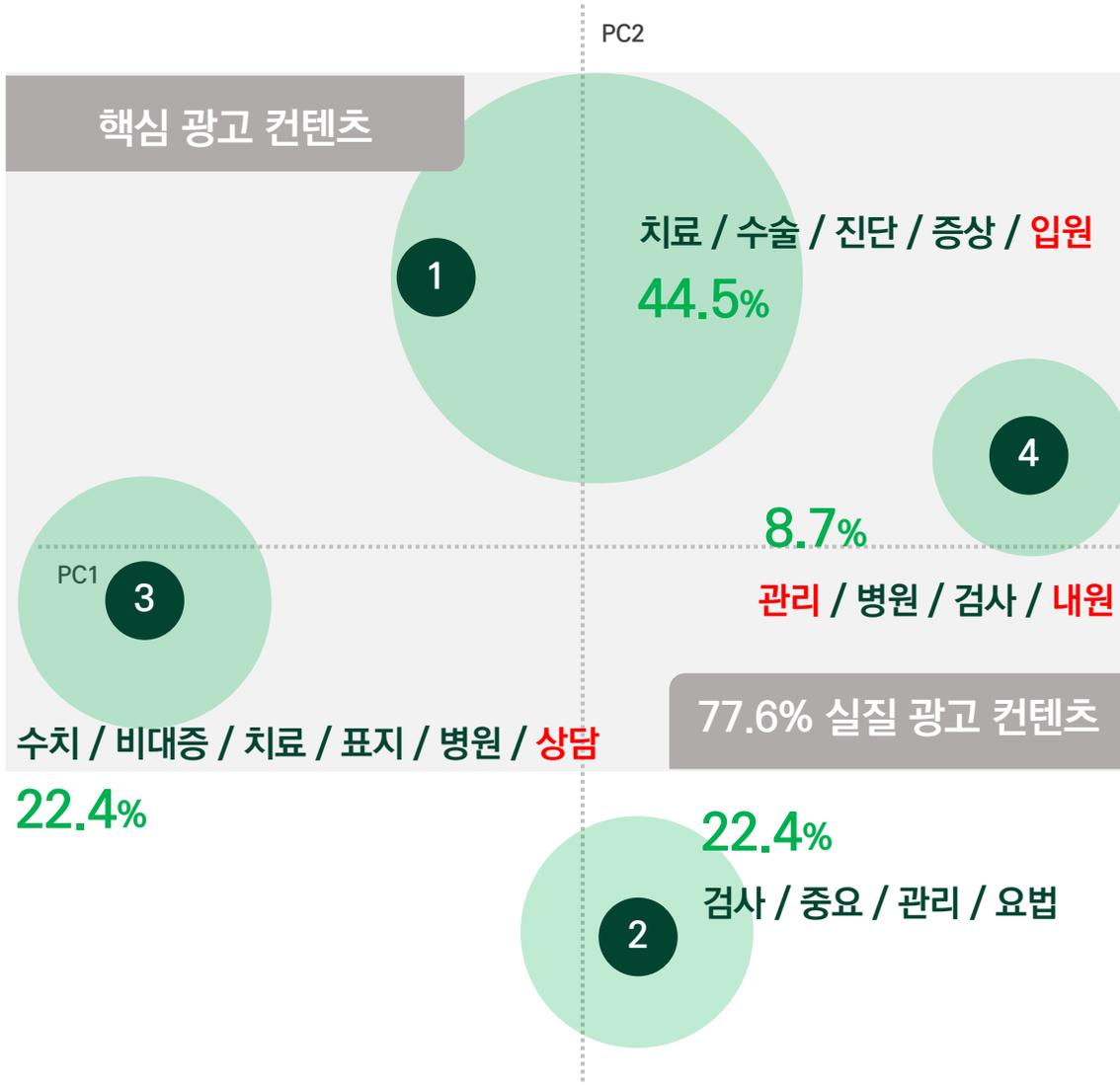
## 광고 콘텐츠 : 위암



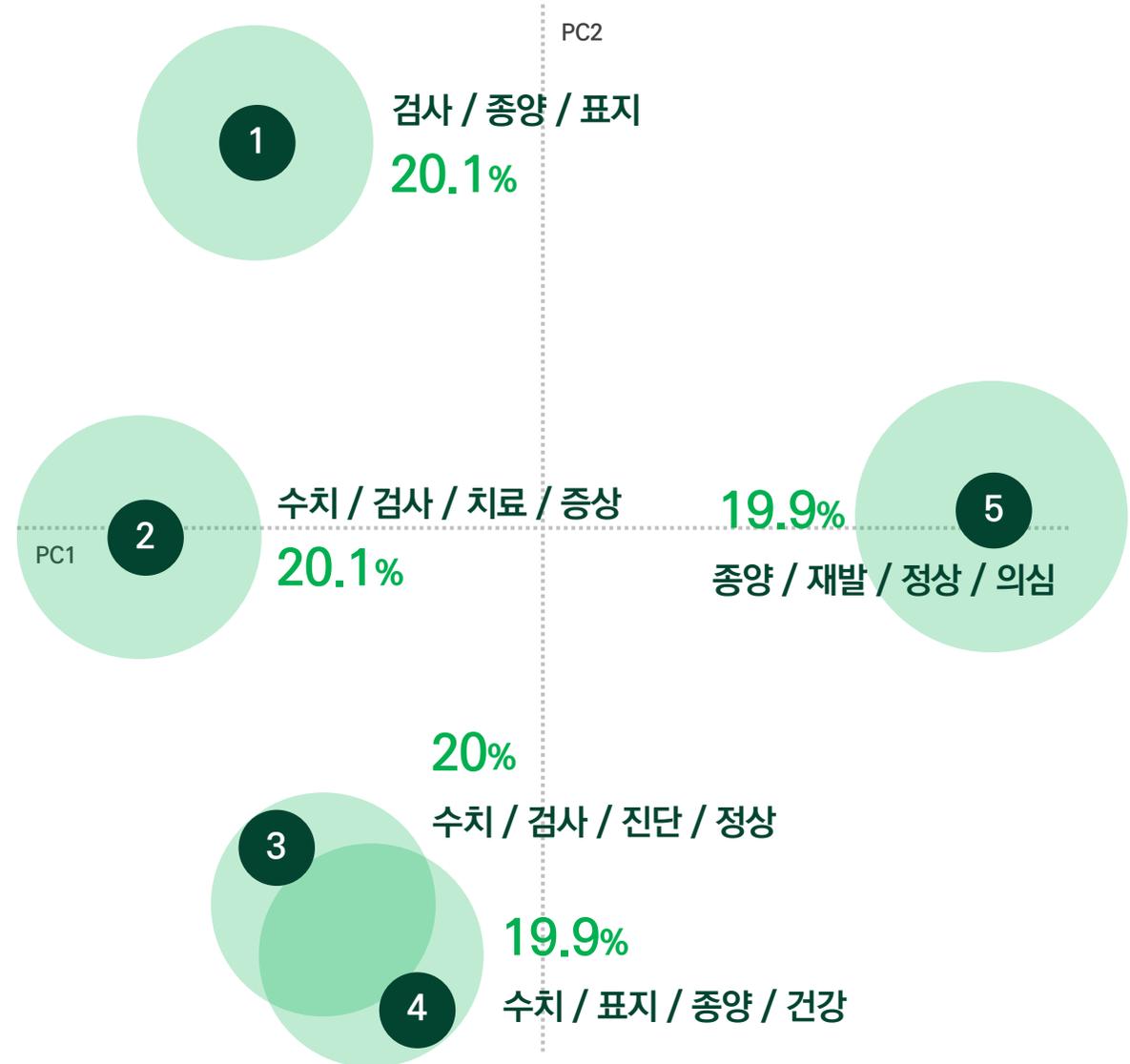
## 비광고 콘텐츠 : 위암



## 광고 콘텐츠 : 전립선암



## 비광고 콘텐츠 : 전립선암



## 광고 콘텐츠 : 간암

### 핵심 광고 콘텐츠

5

부작용 / 면역 / 항암  
6.5%

PC2

지방 / 수술 / 진단 / 검사  
32.0%

1

항암 / 증상 / 관리 / 면역력  
22.0%

3

PC1

2

진행 / 치료 / 양방 / 고주파  
28.0%

56.6% 실질 광고 콘텐츠

25.2%  
여성 / 증상 / 건강 / 암보험

4

11.4%  
치료 / 증상 / 정기 / 검사

## 비광고 콘텐츠 : 간암

증상 / 간경화 / 치료 / 진행<sup>PC2</sup>

1

26.3%

증상 / 검사 / 치료 / 간염  
24.6%

2

진행 / 간경변 / 위험  
6.2%

5

검사 / 지방 / 증상 / 황달  
24.6%

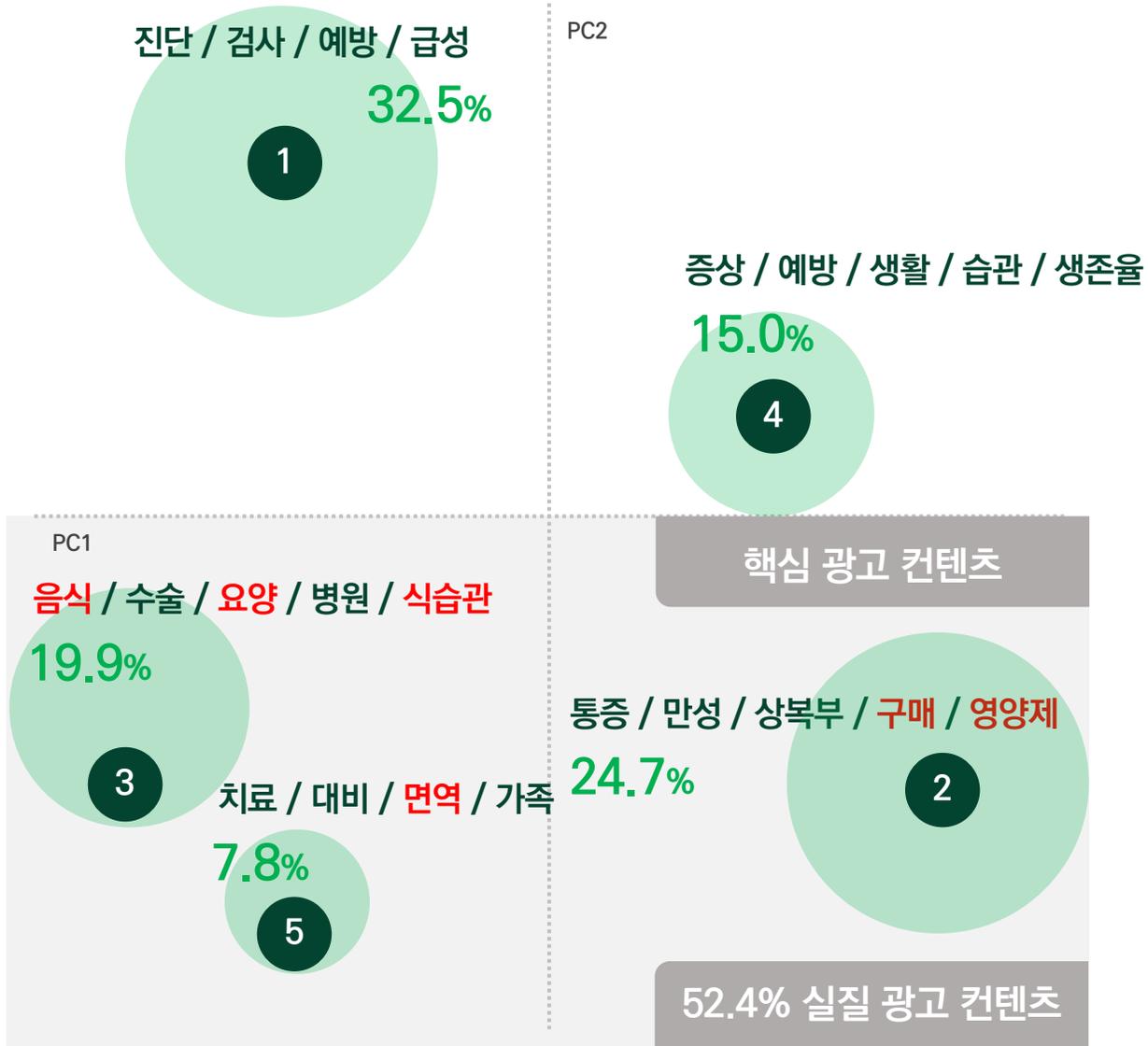
3

PC1

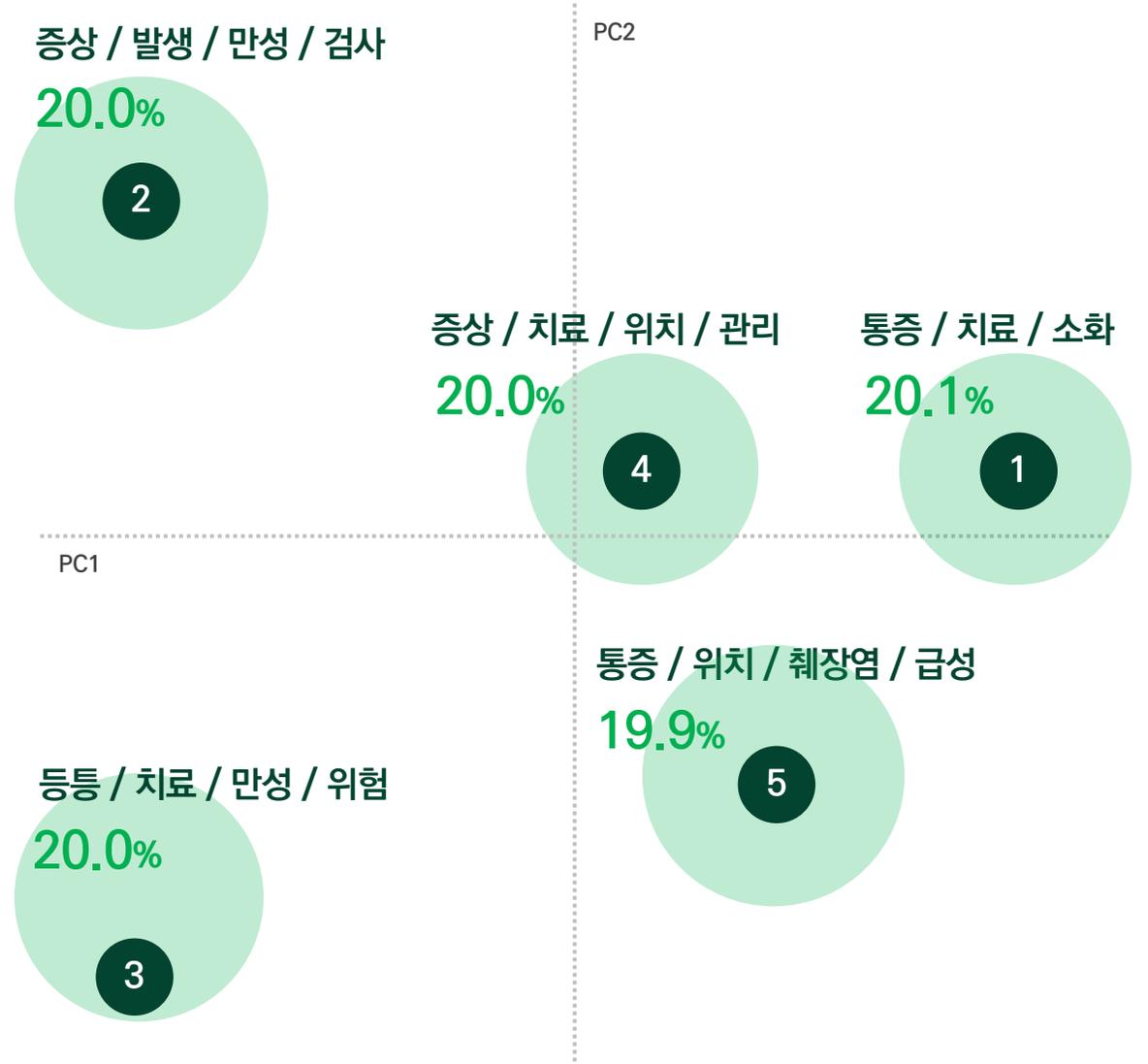
염증 / 치료 / 예방 / 기능  
18.3%

4

## 광고 콘텐츠 : 궤장암



## 비광고 콘텐츠 : 궤장암



## Summary

광고성 포스팅은 전체의 48.6%로 포털에서 검색을 통해 접할 수 있는 글 중 약 전반에 가까운 게시글이 광고성 포스팅인 것으로 조사됨.

**NAVER** 45.7%

**Google** 54.2%

광고성 게시글은 핵심 키워드를 통해 검색시 광고성 게시글이 노출되도록 하여 기초적이고 원론적인 정보를 제공하는 식의 유도를 통해 본래 광고 내용으로 이끄는 방식임