CAS INSIGHTS™

BIOMARKERS

The key to early cancer detection
Introduction

Cancer is a leading cause of death worldwide.\(^1\) With its incidence expected to rise,\(^2\) early detection is critical to reducing morbidity and mortality with higher survival rates, the possibility for less invasive treatments with fewer side effects, and less emotional burden on patients and their families.\(^3,4\) It can also reduce costs and help fuel research into cancer detection and treatment.

Cancer biomarkers provide information about cancer and are, therefore, essential tools for early detection. A biomarker is any biomolecule, cellular structure, or bioactivity that can be measured and evaluated as an indicator of pathogenic processes, normal biological processes, or pharmacological responses to a treatment.\(^5,6\)

In this report, we examine data from the CAS Content Collection\(\textsuperscript{TM}\),\(^7\) the largest human-curated collection of published scientific information, to provide insights into advances in cancer biomarkers. We look at trends in specific cancers, geographic regions, types of biomarkers, as well as investment and clinical trials. We also explore Excelra’s Biomarker Insights dataset, which contains manually compiled biomarker information for selected disease indications. We have put a special emphasis on pancreatic and liver cancers because they are known to be difficult to detect early, with high morbidity and mortality.\(^8\)–\(^11\)
What makes a biomarker?

What are biomarkers?

Cancer biomarkers belong to a variety of biological molecule types and structures, including proteins (like enzymes, hormones, and tumor-associated antigens), nucleic acids, and small organic molecules (Figure 1).

![Figure 1. General types of cancer biomarkers](image)

Biomarkers also encompass imaging (e.g., techniques like MRI, CT, and PET scans), and they extend to tools and technologies applied in the prediction, diagnosis, and pharmacological responses to a therapeutic treatment. These include liquid biopsies, mass spectrometry, and immunochemistry techniques.

When are biomarkers used?

Clinical uses for cancer biomarkers include:

- **Susceptibility/risk** to assess an individual’s potential for developing a disease.
- **Diagnostic (screening)** to detect a specific cancer in an individual. These markers require high specificity and sensitivity.
- **Monitoring** to assess the status of a disease or condition.
- **Prognostic** to predict the course of the disease and help inform treatment.
- **Predictive** to anticipate a potential treatment response.
- **Safety** to indicate chances of toxicity as an adverse effect of a medication.

The choice of biomarker(s) depends on the type of cancer being investigated and the specific diagnostic needs. A combination of biomarkers and diagnostic methods is often used to increase accuracy.

How are biomarkers developed?

Since the early 20th century, scientific advancements in molecular biology and genetics have helped enable biomarker development and continue shaping the landscape to this day.

However, biomarkers must undergo rigorous validation, testing, and clinical trial phases before receiving regulatory approval for clinical use. The development and adoption of cancer biomarkers can vary by cancer type, as well as the availability of technologies and funding. Although many biomarker candidates were recently proposed in compendia, few have passed these validation requirements in clinical trials.
Thanks to recent sizeable developments in technology, there has been persistent growth in biomarker-related scientific publications – with over 30,000 (mainly journal articles and patents) in the CAS Content Collection related to early cancer detection/diagnosis (Figure 2A). Notably, we found a nearly 30% increase in journal articles in the last two years. The growth in patents, however, is slower, indicating that we are still at the knowledge accumulation stage, which precedes its subsequent transfer into patentable applications.

Figure 2B compares the yearly growth rate of biomarker publications for early cancer detection to those related to anti-tumor agents. Initially, the intense search for anti-tumor drugs resulted in a higher growth rate in the area. However, publications related to biomarkers for early cancer detection have significantly overtaken them within the last decade due to successful cancer treatment only achievable at an early, localized stage of the disease.
**Geographical distribution of research**

We found China, the U.S., South Korea, Japan, and Germany to have the most published journal articles and patents related to cancer biomarkers research, with China as an eminent leader. Over the years, there has been a gradual increase in the proportion of publications from China and South Korea compared to the U.S., Japan, and Germany (Figure 3).

**Distribution of biomarker research across different cancers**

Breast cancer, lung cancer, and liver cancer were explored in the highest number of cancer biomarker-related documents (Figure 4A). However, when normalized over all documents related to the specific cancer type, pancreatic cancer had the highest value (Figure 4A, orange line). All major cancer types saw substantial growth in the recent three-year period, but lymph node cancer, pancreatic cancer, and liver cancer are drawing attention with steady growth in recent publications (Figure 4B).
Leading biomarker types, techniques, and uses

Among biomarker types and structures, proteins were among the first and most used in cancer diagnostics. Most are based on cancer antibodies/immunoglobulins, enzymes, and hormones, as reflected by the number of published documents (Figure 5A).

In recent years, metabolites and exosomes have emerged as promising new classes of markers and exhibited a fast and consistent yearly growth rate in published documents. According to yearly growth, RNAs, specifically mRNA and ncRNA, are also among the attractive candidates (Figure 5B).

Figure 4. (A) Number of publications in the CAS Content Collection related to biomarkers for the diagnosis of various cancers (columns); the line indicates the biomarker/cancer number of documents ratio; (B) Yearly growth of biomarker-related documents for the last five years (2018–2022)

Figure 5. (A) Number of publications in the CAS Content Collection related to various biomarkers for cancer diagnosis (B) and their yearly growth over the last five years (2018–2022)
In terms of techniques, we found immunoassays, polymerase chain reactions (PCR), proteomics/protein microarrays, and gene expression profiling to be the most used techniques for biomarker detection.

In terms of use, diagnostic biomarkers were the most used, accounting for almost half of the publications in our collection. All biomarker types marked substantial growth in the last two to three years, but susceptibility/risk and predictive biomarkers have the highest yearly growth rate.

The top biomarkers

Some biomarkers are specific to a particular type of cancer, while others are applied to a wider variety of malignancies. Unsurprisingly, we found the biomarkers that were explored in the highest number of documents usually applied to several cancer types (Figure 6).
The most prevalent biomarkers are:

- **p53 protein** — with p53 being one of the most frequently mutated genes in cancer, it is used across multiple cancers and clinical applications.44–50
- **Carcinoembryonic antigen (CEA)** is a serum marker associated with colorectal cancer as well as breast, liver, prostate, and other cancers.51

The overall picture

In **Figure 7**, we present a mind map of the cancer biomarkers research area, indicating the number of documents related to each subcategory. The type of molecule/structure applied as a biomarker and its functionality are the areas attracting the most attention.
Figure 7. Mind map of the cancer biomarker research area with an indication of the number of documents in each subcategory.
Further insights into pancreatic and liver cancer

Of over 30,000 publications in the CAS Content Collection related to biomarkers for early cancer detection/diagnosis, over 4,000 are related to liver cancer, and over 3,000 are related to pancreatic cancer. Despite the high interest, these areas still lack reliable biomarkers.

Pancreatic cancer

Some of the accepted biomarkers and risk factors for the detection of pancreatic cancer include serum cancer antigen 19-9 (CA 19-9), KRAS, p16/CDKN2A, TP53, and SMAD4 somatic mutations. However, these markers come with low sensitivity and specificity. The poor prognosis and high mortality associated with pancreatic cancer are largely attributed to the late detection of the disease.

Current pancreatic cancer biomarker needs are identifying those at high risk, improving screening for detection, and helping to differentiate better between benign and neoplastic lesions. In the last decade, several studies have explored the potential for novel biomarkers for early pancreatic cancer detection. These markers extend from metabolites to genetic mutations, standalone or combination panels, and use varied sample sources to move toward non-invasive testing. However, they lack stringent validation and are not included in treatment guidelines.

Liver cancer

Just like pancreatic cancer, the liver cancer field (hepatocellular carcinoma) lacks useful biomarkers for surveillance and early diagnosis. There are some established risk factors, and clinical practice guidelines include screening techniques through imaging. However, no molecular markers have been found to be specific or sensitive. While several biomarkers have been explored, further validation is needed.

Further insights from Excelra’s Biomarker Insights dataset

Excelra’s Biomarker Insights dataset (biomarkerinsights.com) allowed us to explore connections between disease, biomarker, drug, and clinical outcomes. Within the dataset, which spans the last seven years, there were 1,007 biomarkers associated exclusively with pancreatic cancer and 3,094 with liver cancer. An additional 425 were shared between the two indications (https://www.excelra.com/databases/custom-biomarker-knowledgebase/).

To assess promising markers within the database, we scored the diagnostic and risk markers to prioritize the ones that can be considered for further validation and diagnostic application. The selected markers are landscaped for prioritization and selection, as shown in Figure 8A, with a zoomed-in view in Figure 8B. Some key biomarkers that are already on the market and being used as established markers are highlighted, as well as several others that could be promising.
Figure 8. Diagnostic and risk biomarkers of pancreatic and liver cancer. A biomarker landscape score for validity (Y-axis) and disease diagnosis association (X-axis). (A) The top panel is the complete plot, and the lower panel (B) represents a zoomed-in view. The validity score consists of factors like pre-existing regulatory qualification, number of supporting articles, and number of contexts for use. A diagnosis and risk score is a summation of the number of studies that concluded a positive or negative association of the marker with the outcome. A positively associated marker is high (expression level or presence of variant) when the disease is present, and the negatively associated one is low.
Capital investment in early cancer detection

Capital investment data from Pitchbook, an online platform for investment data, reveals a steady increase in invested capital and financial deals over the last 20 years (Figure 9A). GRAIL, a company that works on mult-cancer early detection (MCED) testing, has raised the highest capital of around $2 billion in the last 20 years (Figure 9B).

In terms of geographical distribution, the U.S. leads in the amount of capital invested, followed by China and South Korea (by a large margin — investment in the U.S. is around five times greater than China and around seven times greater than South Korea). Growth in the number of deals made over the last two decades for the few leading countries or regions shows a steady increase (excluding a minor dip in 2022), indicating companies’ continued interest in early cancer detection.

Further research developments

Companies worldwide are researching and creating tests to detect pancreatic and liver cancers at early stages to save lives and improve patient outcomes. Examples include blood tests that utilize machine learning, AI-enabled liquid biopsies, and a genetic-based early screening test.

In terms of clinical trials, early cancer diagnostic testing for pancreatic and liver cancers is starting to see increased numbers, with a sharp increase for both in the past few years. Nearly 60% of these trials are focused on the diagnosis of pancreatic cancer, and over 40% on the diagnosis of liver cancer. However, the majority of trials for both indications are currently in recruiting status.
Future outlook

The development of biomarkers for early cancer diagnosis holds significant promise in improving patient outcomes and transforming cancer care. However, we are still facing challenges, including:

- **Cancer heterogeneity** — identifying biomarkers that accurately represent this and can be universally applicable.
- **A balance of specificity and sensitivity** — the lack of specificity or too much sensitivity can lead to false positives, while highly specific biomarkers may miss some cases of cancer.
- **Individual variation** makes it difficult to establish accurate cut-off values for diagnosis or risk assessment.
- **Time-consuming and costly validation** in clinical trials and during regulatory approval.
- **Standardizing** laboratory techniques across different laboratories and platforms, necessary for consistency and comparability of results, can be challenging.
- **Availability and cost** of tests can limit widespread adoption.

Despite this, cancer biomarker research continues to advance, with ongoing efforts to discover and validate new biomarkers. Overall, efforts for early cancer diagnosis are focused on leveraging advancements in genomics, molecular profiling, non-invasive techniques, computational analysis, and collaborative research to enhance early detection, improve diagnostic accuracy, and enable personalized treatment approaches.

As technology improves and our understanding of cancer biology deepens, the possibility for earlier and more accurate cancer diagnosis is becoming increasingly achievable. We have a path to reduce cancer-related morbidity and mortality in the future significantly, but continued investment, along with regulatory support, will be key to realizing the full potential of biomarkers in cancer care.
References


For more details on the research landscape of biomarkers, see our publication at cas.org/insights
CAS connects the world’s scientific knowledge to accelerate breakthroughs that improve lives. We empower global innovators to efficiently navigate today’s complex data landscape and make confident decisions in each phase of the innovation journey. As a specialist in scientific knowledge management, our team builds the largest authoritative collection of human-curated scientific data in the world and provides essential information solutions, services, and expertise. Scientists, patent professionals, and business leaders across industries rely on CAS to help them uncover opportunities, mitigate risks, and unlock shared knowledge so they can get from inspiration to innovation faster. CAS is a division of the American Chemical Society.

**About Excelra:**
Excelra is a pioneer in biopharma data and analytics driven by experts with scientific acumen and technological expertise. With over two decades of excellence, Excelra catalyzes groundbreaking discoveries in the world of life sciences. Comprising a team of over 1000 Life Science & Technology professionals, having a global presence with offices in India, USA, UK, Belgium, Netherlands, Germany, and Switzerland. Excelra collaborates with 16 of the top 20 Pharma companies. Excelra’s unique approach empowers revolutionary drug discovery and development, seamlessly transforming research data into actionable insights.

**Connect with us at cas.org**