Recent advances in blood-based and artificial intelligence-enhanced approaches for gastrointestinal cancer diagnosis

Li-Shi Li, Xiang-Yu Guo, Kun Sun

Abstract

Gastrointestinal (GI) cancers are among the most common cancer types and leading causes of cancer-related deaths worldwide. There is a tremendous clinical need for effective early diagnosis for better healthcare of GI cancer patients. In this article, we provide a short overview of the recent advances in GI cancer diagnosis. In the first part, we discuss the applications of blood-based biomarkers, such as plasma circulating cell-free DNA, circulating tumor cells, extracellular vesicles, and circulating cell-free RNA, for cancer liquid biopsies. In the second part, we review the current trends of artificial intelligence (AI) for pathology image and tissue biopsy analysis for GI cancer, as well as deep learning-based approaches for purity assessment of tissue biopsies. We further provide our opinions on the future directions in blood-based and AI-enhanced approaches for GI cancer diagnosis, and we think that these fields will have more intensive integrations with clinical needs in the near future.

Key Words: Liquid biopsy; Plasma cell-free DNA; Circulating tumor cell; Extracellular vesicle; Exosome; Artificial intelligence

©The Author(s) 2021. Published by Baishideng Publishing Group Inc. All rights reserved.

Core Tip: Recent studies have discovered a variety of blood-based biomarkers with great potential in improving the diagnosis and surveillance of gastrointestinal (GI) cancers. In this article, we review the latest advances in the diagnosis of various GI
Gastrointestinal (GI) cancers, which include tumors from the colon, rectum, stomach, pancreas, esophagus, anus, gallbladder, liver, and bile duct, are among the most common cancer types and leading causes of cancer-related deaths worldwide.[1] Colorectal cancer (CRC), in particular, is the third most common cancer in men (9%) and women (8%) in the United States.[1-3] Most patients with GI cancers are in the advanced stage upon diagnosis and have relatively poor prognosis outcome. For example, the overall 5-year survival rate for patients with hepatocellular carcinoma (HCC) is around 20%[3], while the survival rate for early stage patients is as high as 70%[1,3]. This is because various efficient treatment strategies, such as surgical resection and organ transplantation, are feasible for early-stage patients only. Hence, early diagnosis is of high clinical significance for better healthcare of cancer patients.

Pathological analysis of the tumor tissue is currently the “gold standard” for clinical diagnosis of cancers.[4,5] In particular, image-based computed tomography (CT) colonography, optical colonoscopy, and endoscopic biopsy are mostly widely used in the diagnosis of GI cancers. Recent studies have demonstrated that blood-based approaches, also known as liquid biopsy, offer substantial advantages over conventional tissue biopsy-based diagnostic methods, including minimally invasive nature, cost-efficiency, clinical convenience for real-time monitoring, and the potential to promote higher patient compliance. Major analytes in current liquid biopsy studies include circulating cell-free DNA (cfDNA), circulating tumor cells (CTCs), as well as extracellular vesicles (EVs) that transport the biomolecules from tumor cells, such as lipids, proteins, DNA fragment, and RNAs.[6-9] Moreover, advances and integration of artificial intelligence (AI) have brought revolutionary changes to the clinical and translational cancer studies[10-12], especially in cancer diagnosis. The identification and utilization of disease-specific biomarkers, and development of exquisite diagnostic algorithms are both highly valuable for early diagnosis of cancers.

In this article, we provide an overview of current approaches for GI cancer diagnosis, especially in CRC and HCC. We focus on blood-based biomarkers, as well as AI algorithms for image-based tissue biopsies. We review the identification of biomarkers and development of diagnostic methods, and provide our opinions on future directions in early diagnosis of GI cancers.

## BLOOD-BASED LIQUID BIOPSY APPROACHES

With the advantages of being highly efficient, minimally invasive, and cost-effective, liquid biopsy has become an emerging technique for early diagnosis of cancers. Tumor-derived cell-free DNA in peripheral blood provide a surrogate for researchers/clinicians to investigate the genomic and epigenomic landscape as well as dynamics during treatment of malignant tumors[13,14]. Conventionally, liquid biopsy approaches cover the analysis of various circulating components, including cfDNA, cfRNA, CTCs, and tumor-associated EVs (Figure 1)[15], and significant progress has been made in these areas in the past few years (Table 1).

### Circulating cell-free DNA

CfDNA molecules in plasma originate from various tissues in human body. In cancer
| Cancer type | Biomarker type | Biomarkers | No. of patients /controls | Accuracy | Ref. |
|-------------|----------------|------------|--------------------------|----------|------|
| GC          | cfDNA copy number | HER2 gene | 60/30 | Sn = 73.3%, Sp = 93.3% | [19] |
| GC          | cfDNA methylation | BARHL2 | 128/30 | Sn = 90%, Sp = 100% | [104] |
| GC          | CTC            | CTC level | 116/31 | Sn = 85.3%, Sp = 90.3% | [49] |
| GC          | Exosomal IncRNA | IncRNA-UEGCI | 10/5 | AUC = 0.876 | [105] |
| GC          | Exosomal IncRNA | IncRNA-psck2-2-1 | 63/29 | Sn = 84%, Sp = 86.5% | [77] |
| GC          | Exosomal IncRNA | IncRNA-GNAQ6:1 | 43/27 | AUC = 0.732 | [71] |
| GC          | Exosomal IncRNA | IncRNA-HOTTIP | 126/120 | AUC = 0.827 | [106] |
| GC          | Exosomal miRNA  | miR-1246 | 117/82 | AUC = 0.843 | [107] |
| GC          | Serum cfRNA | miR-10b-5p, miR-195-5p, miR-185-5p, miR-296-5p, miR-132-3p, and miR-20a-3p | 441/233 | AUC = 0.702 | [68] |
| GC          | Serum cfRNA | miR-21 | 50/50 | Sn = 88.4%, Sp = 79.6% | [81] |
| GC          | Serum cfRNA | miR-30a-5p, miR-659-3p and miR-3917 | 354 | AUC = 0.82 | [82] |
| GC          | Serum cfRNA | BIGALT5-AS1 | 107/87 | AUC = 0.816 | [108] |
| CRC         | CTC mRNA | ECT2 gene | 90/151 | AUC = 0.821 | [61] |
| CRC         | Serum cfRNA | miR-21, miR-29a, miR-125b | 160/77 | AUC = 0.827 | [96] |
| CRC         | Serum cfRNA | miR-30a-5p | 138/60 | Sn = 77.5%, Sp = 78.3% | [109] |
| CRC         | cfDNA methylation | SFRP2 | 62/55 | Sn = 69.4%, Sp = 87.3% | [21] |
| CRC         | CTC protein | CD133+CD54+CD44+ Protein | 10/10 | Sn = 88.2%, Sp = 92.4% | [62] |
| CRC         | Exosomal miRNA | miR-27a, miR-130a | 40/40 | AUC = 0.773 | [110] |
| CRC         | Exosomal miRNA | miR-125a-3p | 50/50 | AUC = 0.685 | [111] |
| CRC         | Exosomal protein | CPNE3 | 92/32 | Sn = 67.5%, Sp = 84.4% | [112] |
| PC          | CTC protein | Vimentin | 100/30 | AUC = 0.968 | [58] |
| PC          | cfDNA methylation | BNC1, ADAMTS1 | 39/95 | Sn = 94.8%, Sp = 91.6% | [23] |
| PC          | cfDNA methylation | CDO1 | 160 | Sn = 95% | [24] |
| PC          | CTC | CTC level | 126 | Sn = 100%, Sp = 88.6% | [113] |
| PC          | Exosomal IncRNA | CLDN1, TIMP1, MAL2, MARCH2, ITIH2, HIST1H2BK, KRT19 and FGA | 284/117 | AUC = 0.931 | [79] |
| PC          | Exosomal miRNA | miR-21 | 27 | AUC = 0.9 | [76] |
| PC          | Exosomal miRNA | miR-155 | 27 | AUC = 0.89 | [76] |
| PC          | Exosomal miRNA | miR-451a | 56 | Sn = 69.2%, Sp = 70.8% | [114] |
| PC          | Exosomal miRNA | miR-133a | 110/64 | Sn = 90.6%, Sp = 87.2%, AUC = 0.893 | [115] |
| HCC         | cfDNA methylation | cfDNA 5hmC | 1204/958 | AUC = 0.846 | [30] |
| HCC         | Exosomal miRNA | miR-10b-5p | 90/26 | AUC = 0.934 | [75] |
| HCC         | Exosomal miRNA | hnRNPH1 | 88/68 | Sn = 85.2%, Sp = 76.5% | [80] |
| HCC         | Plasma cfRNA | hsa_circ_0000976, hsa_circ_0007750 and hsa_circ_0139897 | 158/53 | AUC = 0.863 | [85] |
| HCC         | Serum cfRNA | miR-143 | 131/122 | Sn = 80.3%, Sp = 82.4% | [87] |
| HCC         | Serum cfRNA | miR-132, miR-212 | 80/42 | Sn = 93.75%, Sp = 63.44% | [116] |

GC: Gastric cancer; CRC: Colorectal cancer; PC: Pancreatic cancer; HCC: Hepatocellular carcinoma; Sn: Sensitivity; Sp: Specificity; AUC: Area under the curve.
blood-based liquid biopsy in cancer diagnosis. In the peripheral blood of cancer patients, circulating cell-free nucleic acids, circulating tumor cells, and extracellular vesicles carry the tumor-associated information, which allow noninvasive diagnosis of cancer patients.

Figure 1 Blood-based liquid biopsy in cancer diagnosis. In the peripheral blood of cancer patients, circulating cell-free nucleic acids, circulating tumor cells, and extracellular vesicles carry the tumor-associated information, which allow noninvasive diagnosis of cancer patients.

patients, the dying tumor cells release their DNA, which contains the tumor-specific signatures, including somatic mutations, copy number aberrations (CNAs), and altered DNA methylation profiles, into the circulation (which is known as ctDNA). However, tumor-derived cfDNA usually account for a small fraction of the total cfDNA, and the concentration is even lower in early-stage cancer patients, which largely limits the accuracy of most cfDNA-based liquid biopsy assays. To this end, in recent studies, researcher have explored the application of extra-deep sequencing and discovered a broad range of tumor-related signals in cfDNA to further improve their performance in both diagnosis and prognosis monitoring.

Conventional genomic and epigenomic alterations: Somatic mutations are one of the most widely studied tumor-associated genomic alterations. In a proof-of-concept study, extra-high sequencing of cfDNA in HCC patients, researchers are able to non-invasively resolve the mutation landscape of the tumor with high accuracy[16], which would be informative in guiding the therapeutic strategies for cancer patients. The somatic mutations in cfDNA are widely used as biomarkers in cancer diagnosis and prognosis monitoring. For example, in pancreatic cancer (PC), the overall survival of patients with detectable KRAS mutation in cfDNA was significantly shorter than that of patients without such mutations[17]. Notably, the somatic mutations detected in cfDNA may not solely come from the tumors. In fact, recent studies have proven that clonal hematopoiesis also introduces somatic mutations which are detectable in cfDNA. Hence, it is highly recommended to analyze the blood cells in parallel to cfDNA in somatic mutation-oriented studies.

DNA modifications, mostly 5-methylcytosine (5mC) and 5-hydroxymethylcytosine (5hmC), are abundant epigenetic regulators in the human genome. These modifications are known to affect gene expression, and suffer from gross abnormalities in almost all kinds of tumors. In cancer patients, such modifications are detectable in cfDNA and could be used to point to the tissue origin of the tumor. For example, Chan et al[20] analyzed the CNA and 5mC profiles of cfDNA simultaneously using whole-
Emerging cfDNA fragmentomic features: Plasma cell-free DNA is known as naturally fragmented, small molecules. Recent studies have revealed that the tumor-derived DNA possesses a remarkable difference in fragmentation characteristics compared to the background DNA (which mostly comes from the hematopoietic system) [31], and demonstrated the clinical significance of these features in the diagnosis of various GI cancers.

cfDNA molecules show a dominant size peak at 166 bp and most of them are shorter than 200 bp, and this phenomenon is believed to correlate with the nucleosome structure [32]. In cancer patients, tumor-derived cfDNA is even shorter than the background DNA [33]. Hence, it is doable to enrich the tumor-derived cfDNA through selection of shorter cfDNA molecules in principle. Indeed, Mouliere et al. [34] demonstrated the enrichment of tumor-derived cfDNA in fragment sizes between 90 and 150 bp in a cohort of 344 plasma samples from various cancer patients. They also showed that using an integrated analysis of cfDNA at different size ranges, they could classify CRC patients from healthy individuals with a high specificity of 94% [34].

Besides size pattern, recent studies further showed that the relationship between the fragmentation pattern and nucleosome structure could point to the tissue origin of cfDNA [35, 36]. In particular, cfDNA of different origins tends to have different preferences on fragment ends. Jiang et al. [16] compared the genome-wide fragment end patterns of cfDNA in one HCC patient and one non-cancerous control; they reported that many genomic loci show significantly higher chances to serve as fragment ends (which they called “preferred-ends”) in the HCC patient, and the usage of such genomic loci could differentiate HCC patients from non-cancerous controls with a favorable accuracy. Moreover, the same team further showed that the nucleotide frequencies (or motifs) at the cfDNA end, which are believed to correlate with nuclease cutting preference, also change in various cancers [37, 38]. In fact, they found a nuclease, DNASIL3, which plays key roles in the apoptotic fragmentation process and shows strong sequence preferences when cutting the DNA [39]. Interestingly, this gene is significantly down-regulated in many cancers and thus affects the end motif pattern of the tumor-derived DNA [37]. The discovery of these novel biomarkers largely extends our knowledge base on the biology of cfDNA and cancer biomarkers; these novel biomarkers also promise the development of less complex and cheaper assays than those based on the conventional genomic/epigenomic biomarkers. However, most current studies are still in proof-of-principle stage, thus large-scale validation studies are essential to assure the performances before clinical investigations.
**Circulating tumor cells**

CTCs are released into the bloodstream of patients through the blood vessels and lymphatic vessels. When the primary cancer cells acquire an increased propensity for migration, invasiveness, and resistance, these tumor cells will invade through the tissue surrounding their site of origin and keep invading until they enter the blood or the lymphatic system, and become the CTCs[40,41]. Although most of CTCs will die within hours, only a small part of them can potentially form a new cancerous lesion at a distant site and grow into a new metastasis[42,43].

There are many commercial kits for CTC isolation and enrichments. CTC isolation and enrichment are mostly based on their physical and immunological properties. The physical features of CTCs, such as the size, density, and electric charge, are utilized to enrich CTCs from blood. For example, ISET (isolation by size of epithelial tumor cells) system and a microfluidic chip developed by Haber's lab filter CTCs from blood by the size of CTCs[44-46], CAP (centrifugal affinity plate) system isolates CTCs by the density gradient centrifugation[47], and ApoStream system enriches CTCs by dielectrophoretic field-flow fractionation[48,49]. Enriching CTCs based on the immunological properties of CTCs depends on the antigens expressed on the surface of epithelial cells, including epithelial cell adhesion molecule (EpCAM) and cytokeratin (CK). CellSearch is the first and only FDA-approved test for capturing CTCs that express EpCAM by using immunomagnetic beads[50,51]. There are also some methods that combine epithelial marker immunofluorescence staining with automated microscopy, such as NYONE[52]. At the protein level, fluorescence in situ hybridization (FISH) has been utilized for identifying chromosomal rearrangements in CTCs [53]. Immunocytochemistry (ICC) is the major technique that is used by several commercial kits for isolating CTCs, including ApoStream and OecoQuick. Flow cytometry is another protein analysis method employed for CTC isolation. The analysis of genetic information of CTCs is mainly based on PCR and next-generation sequencing (NGS), such as RNA-based digital PCR and single-cell sequencing[54,55]. Due to the low concentration and fragile property of CTCs and limitations of current methods, combination and further optimization may achieve better enrichment and detection for CTCs.

The clinical application of CTCs depends on the number of CTCs and the information about protein expression and gene mutation in CTCs. Kang et al[56] showed that for individuals with two or more CTCs per 7.5 mL of blood, 97.1% of them were GC patients in their cohort. The increased number of CTCs after treatment was also correlated with early recurrence[27]. Zhang et al[27] showed that in GC patients, the patients with five or more CTCs per 7.5 mL in postoperative blood had a significantly shorter overall survival and disease-free survival compared to those with less CTCs. The protein expression patterns in CTCs are also an important biomarker for cancer diagnosis. Around 50% GC patients with CK’CD44+ CTCs were diagnosed with distant metastases[57]. Wei et al[58] found that, in 76% of PDAC patients, vimentin was upregulated in CTCs, and a high level of vimentin in CTCs was associated with a shorter recurrence-free survival. Similarly, PDAC patients with CTCs that expressed high levels of mucin 1 (MUC-1) and cancer stem cell markers, such as ALDH, CD133, and CD44, were associated with a worse overall survival and tumor recurrence[59,60]. Chen et al[61] found up-regulated expression of ECT2 in CTCs from patients with CRC. Fang et al[62] found that CRC patients with liver metastasis showed a higher percentage of ’CD133’CD54’ CD44’ cell subpopulations. Moreover, the upregulated genes in CTCs, such as SHH, SMO, POU5F1B, and ALCAM, were demonstrated as potential markers for therapeutic response and prognosis prediction in PC patients[63]. In HCC patients, Zhou et al[64] discovered that the presence of EpCAM mRNA+ CTCs was associated with early recurrence when combined with Treg/CD4+ T cell ratio.

**Exosomes**

Exosomes are membrane-bound extracellular nanovesicles (EVs) released by many types of cells, having a size ranging from 30 to 150 nm. Exosomes have been found in various bodily fluids, including plasma, urine, and malignant ascites[65]. Exosomes are involved in mechanisms of cell-cell communication in physiological and pathological tissues through transporting biomolecules, such as lipids, proteins, mRNA, microRNAs, and DNA fragments, and the intercellular transfer of signal components[66,67]. Tumor-associated exosomes are larger than those derived from healthy cells and are enriched with different types of mediators of tumorigenesis, including proteins, growth factors, immunomodulatory molecules, and nucleic acids that arise from the cytoplasm of donor cancer cells[68].
In the last decade, a significant number of studies have focused their attention on the exosome-mediated cross-talk which can promote the activation of signaling pathways and reprogram the functions of recipient cells through their cargo transfer that occurs between cancer and normal cells, especially in the tumor microenvironment[65,69-71].

Several examples of exosome biomarkers, such as proteins, microRNAs, and long non-coding RNAs (lncRNAs), have been shown to have the potential for the diagnosis and prognosis of specific types of cancer. For example, GPC-1 was identified as a diagnostic exosome marker for PC and CRC. CD9 and CD147 were found to be upregulated in exosomes derived from serum of CRC patients[72]. Previously, seven miRNAs including let-7a, miR-21, miR-23a, miR-155, miR-223, miR-1246, and miR-1229 were reported to be highly expressed in the serum exosomes of CRC patients and were significantly decreased after surgical resection[73]. MiR-103 was reported to be delivered from hepatoma cell-derived exosomes into endothelial cells, and induced metastasis[74]. In HCC patients, serum exosomal miR-10b-5p was found to be upregulated and could be used to distinguish HCC patients from healthy controls with an AUC of 0.934[75]. Nakamura et al[76] found that exosomal miR-21 and miR-155 were significantly upregulated in PDAC patients compared to chronic pancreatitis patients. In GC patients, exosomal IncRNA pck2-2:1 and IncRNA GNAQ-6:1 were significantly downregulated and could be served as biomarkers for early diagnosis for GC[77,78]. Yu et al[79] developed a panel of eight exosomal RNAs, which could distinguish PDAC patients from chronic pancreatitis patients with an AUC of 0.931. The exosomal mRNA ITM2BP1 was found to have a high expression level in HCC patients and could be used to distinguish HCC patients from chronic hepatitis B patients with an AUC of 0.865[80]. Furthermore, IncRNAs, such as LncRNA-ARSR, Lnc-son26t, and LncRNA-h19, were showed to be associated with tumor progression[72,81].

**Circulating cell-free tumor RNAs**

Circulating cell-free RNAs (cfRNAs) are also present in the blood circulation and have the potential to serve as cancer biomarkers for cancer diagnosis. CfRNAs may be released into the blood through mechanisms of cell death or exosome-mediated signaling by living cells.

Historically, the study of cfRNA has focused on a small group of known cancer-related mRNA, miRNAs, and lncRNAs[76-78,82]. For instance, miR-21 showed an increased expression level in GC patients, which could serve as a classification biomarker with an 88.4% sensitivity and 79.6% specificity[83]. Huang et al[84] designed a panel of six miRNAs that were upregulated in the serum of GC patients, which showed an AUC of 0.702 in distinguishing GC patients from healthy controls. Shimura et al[85] found that three miRNAs (miR-30a-5p, miR-659-3p, and miR-3917) were overexpressed in GC patients who had peritoneal metastasis and could be used to distinguish patients with or without peritoneal metastasis with an AUC of 0.82. Yamada et al[86] showed that the expression of three serum miRNAs, including miR-21, miR-29a, and miR-125b, could discriminate CRC patients from healthy ones with an AUC of 0.827. In HCC patients, serum miR-143 was downregulated and could distinguish HCC patients from healthy controls with an AUC of 0.831[87]. Another study showed that a panel of three plasma circRNAs were found to distinguish HCC patients from healthy ones with an AUC of 0.863[88].

**AI-ENHANCED ALGORITHMS FOR IMAGE-BASED AND TISSUE-BASED PATHOLOGY ANALYSIS**

The early diagnosis and prognosis of a cancer type have become a necessity in cancer research. Image-based and tissue biopsy-based pathology analyses are considered to be the “gold standard” for their accuracy and effectiveness for cancer screening.

In recent years, AI has become an emerging field with the advances of computational power and massive amounts of learning data. Over the last two decades, different machine learning (ML) algorithms, such as support vector machines (SVM), random forest (RF), and artificial neural networks (ANNs), have been widely used in the field of medicine[89-91]. More recently, convolutional neural networks (CNN), a novel algorithm, show its great potential in processing and interpreting radiological and pathological images (Figure 2)[92,93]. Previous studies demonstrated that a multi-layer CNN with pre-trained GoogleNet Inception v3 architecture on skin cancer level prediction could achieve comparable results to dermatologists[94]. Similarly, another image-based deep learning model proposed by Kermann et al[10] in analyzing optical
coherence tomography (OCT) images of the retina is on par with the ophthalmologist, with an accuracy of 96.6%, sensitivity of 97.8%, and specificity of 97.4% on the four-class classification problem. Most importantly, both studies utilized transfer learning when training the model to save the training time, which is also a widely used technique in training an ML model in medical images due to the lack of training data in most cases[95].

Traditionally, the most effective and direct method of screening colorectal carcinoma is endoscopy, particularly colonoscopy[96]. Abdominal computerized tomography (CT) or abdominal magnetic resonance imaging (MRI) are employed when more precise and additional information is needed about the lesion to help assess cancer and look for any signs of spread. However, these methods are usually labor-intensive and time-consuming and require professional knowledge. In the era of big data, the number of colonoscopy examinations and radiology screening is enormously increasing[96]. To this end, the deep learning-based radiology and digital pathology workflow could assist the radiologists and pathologists to diagnose diseases in various aspects, such as tumor classification, gland segment, tumor microenvironment analysis, and prognostic survival prediction, to ease their ever-growing workload and improve their efficiency[92]. More recently, Luo et al[97] built a real-time Gastrointestinal Artificial Intelligence Diagnostic System (GRAIDS) for the diagnosis of upper GI cancer through analysis of imaging data from clinical endoscopies. The GRAIDS achieved a high diagnostic accuracy in detecting upper GI cancers and had a similar predicted sensitivity as expert endoscopies and outperformed the non-expert endoscopies. Besides, they performed a multi-center study and further validated the system from the other five primary care hospitals, thus indicating the general applicability of GRAIDS.
Apart from AI application in colonoscopy, CT, and MRI images, the digital pathology also drew much attention from researchers, in particular adopting deep learning on the whole slide images (WSIs) [92,98-100]. However, the WSIs are multi-gigabyte images, present high morphological variance, and contain various types of artifacts. Such circumstances prevent CNN directly from applying to the high-resolution images. On the other hand, typically small labeled datasets in pathological images impede the generality of the deep learning model [101]. Kermany et al. [102] developed a multiple-instance learning-based deep learning system that only used the reported annotated datasets to tackle the limited number of labeled images and the AUCs were all above 0.98 when testing on multiple cancer types. Joseph et al. [96] introduced a novel machine learning-based system on prediction of CRC outcome for whole digitized haematoxylin and eosin (HE) stained histopathology slides in learning heterogeneous and discriminant contents. They demonstrated the effectiveness of the method and presented a detailed analysis of its different elements which corroborate its ability to extract and learn salient, discriminative, and clinically meaningful content.

At present, pathological analysis of tissue biopsies is still the most widely approach in cancer diagnosis. Recent studies have shown that combination of molecular data and deep learning approach could predict the cancerous status and tissue origin of pan-cancer biopsies with high accuracy. For instance, Sun et al. [103] developed GeneCT, which showed an overall accuracy of 98.2% in differentiating tumor samples from adjacent normal samples, and an overall accuracy of 98.6% in predicting the tissue origins of the biopsies. In addition, as tissue biopsy is one of the most widely materials in cancer studies, and the purity of the biopsies is crucial for appropriate project design and correct interpretation of the data. To this end, Fan et al. [12] showed that the low prediction accuracy of GeneCT in some datasets suggested impurity of the samples. The performance of GeneCT was further evaluated using metastatic tumor samples, which is a common scenario in GI cancers. As a result, GeneCT shows potential as an accurate classification tool for tissue biopsy-based diagnosis, as well as an easy-to-use quality-control tool for cancer studies.

CONCLUSION

GI malignancies, especially CRC and HCC, are common cancers that lead to death of patients with cancers worldwide. Liquid biopsy plays an important role in early-stage GI cancer diagnosis as a noninvasive tumor detection strategy. Biomarkers including cfDNA, CTCs, and exosomes of liquid biopsy are widely used for the detection of tumor-associated molecules. The advantages of cfDNA, such as high amount, being relatively stable and easy for extraction and preservation, mostly representing the information from the major tumor clone, and the significant features including fragment sizes, the proportion of fragment, the end-motif, and genomic alterations, make it the preferential and accurate biomarker for the cancer diagnosis. CTCs are also important liquid biomarkers because they not only contain direct molecular information from the tumor clone but also can be cultured in vitro. However, CTCs are very rare and require fresh blood for isolation.

The potential of liquid biopsy in clinical practice is enormous in the future. However, there are several challenges that need to be resolved for the utility of liquid biopsy in clinical diagnosis. First, standardization should be established for the sample collection and storage conditions, and biomarker molecules and the suitable detection and analysis strategy depend on sample type. Second, tumor-associated molecule detection techniques require high sensitivity for early tumor detection and high specificity for early-stage tumor screening. At the same time, the low content of tumor biomarkers and the varied interference factors in the circulation make the high requirements for the high sensitivity and specificity of detection strategies. Reducing purification steps could shorten the time of sample treatment and reduce the cost, and optimization of the throughput platform could help to increase both the sensitivity and specificity of the biomarkers.

AI based on machine learning and deep learning is widely used in all walks of life. Recently, rapid developments have been made in the utility of AI in medicine areas and showed promising results in terms of accuracy for cancer diagnosis. However, the scale and quality of the training and validation datasets of most of the studies were relatively limited to apply this technique in clinical practice. Moreover, it also requires the external cross-validation, especially in tumor classification. Therefore, future studies with a larger number of datasets with high-quality annotations and external
cross-validation are required for routine practice-level validation.

ACKNOWLEDGEMENTS

We would like to thank Ms. Qi Wang from Shenzhen Bay Laboratory for her technical assistance.

REFERENCES

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2020. CA Cancer J Clin 2020; 70: 7-30 [PMID: 31912902 DOI: 10.3322/caac.21590]

2. Siegel RL, Miller KD, Goding Sauer A, Fedewa SA, Buttery LF, Anderson JC, Cercak A, Smith RA, Jemal A. Colorectal cancer statistics, 2020. CA Cancer J Clin 2020; 70: 145-164 [PMID: 32133645 DOI: 10.3322/caac.21601]

3. Society AC. Cancer Facts & Figures 2021: Atlanta: American Cancer Society, 2021

4. Moore H, Dodd N. Computed tomographic colonography (CTC); colorectal cancer diagnosis with CTC in an Auckland population. J Med Imaging Radiat Oncol 2013; 57: 572-575 [PMID: 24119271 DOI: 10.1111/1754-9885.12062]

5. Greenhill C. Colorectal cancer: CTC for diagnosing symptomatic colorectal cancer. Nat Rev Gastroenterol Hepatol 2013; 10: 198 [PMID: 23458893 DOI: 10.1038/argastro.2013.37]

6. Fici P. Cell-Free DNA in the Liquid Biopsy Context: Role and Differences Between ctDNA and CTC Marker in Cancer Management. Methods Mol Biol 2019; 1909: 47-73 [PMID: 30580422 DOI: 10.1007/978-1-4939-8973-7-4]

7. Keup C, Storbeck M, Hauch S, Hahn P, Sprenger-Haussels M, Tewes M, Mach P, Hoffmann O, Kimmig R, Kasimir-Bauer S. Cell-Free DNA Variant Sequencing Using CTC-Depleted Blood for Comprehensive Liquid Biopsy Testing in Metastatic Breast Cancer. Cancers (Basel) 2019; 11 [PMID: 30781720 DOI: 10.3390/cancers11020238]

8. Nanduri LK, Hissa B, Weitz J, Schölch S, Bork U. The prognostic role of circulating tumor cells in colorectal cancer. Expert Rev Anticancer Ther 2019; 19: 1077-1088 [PMID: 31778322 DOI: 10.1080/14737140.2019.1699065]

9. Maltoni R, Paitleschi M, Ravaioili S, Tumedei MM, Roeca A, Melegari E, Altini M, Puccetti M, Manunta S, Bravaccini S. Cell-Free DNA Variant Sequencing Using CTC-Depleted Blood for Comprehensive Liquid Biopsy Testing in Metastatic Breast Cancer. Cell Transplant 2020; 29: 963689720925057 [PMID: 32584148 DOI: 10.1177/0963689720925057]

10. Kermany DS, Goldbaum M, Cai W, Valentim CCS, Liang H, Baxter SL, McKeown A, Yang G, Wu X, Yan F, Dong J, Prasadha MK, Pei J, Ting MYL, Zhu J, Li C, Hewett S, Ziyar I, Shi A, Zhang R, Zheng L, Hou R, Shi W, Xu X, Yuan D, Wan C, Zhang ED, Chang CL, Li O, Wang X, Singer MA, Sun X, Xu J, Tafreshi A, Lewis MA, Xia H, Zhang K. Identifying Medical Diagnoses and Treatable Diseases by Image-Based Deep Learning. Cell 2018; 172: 1122-1131.e9 [PMID: 29474911 DOI: 10.1016/j.cell.2018.02.010]

11. He B, Lu Q, Lang J, Yu H, Peng C, Bing P, Li S, Zhou Q, Liang Y, Tian G. A New Method for CTC Images Recognition Based on Machine Learning. Front Bioeng Biotechnol 2020; 8: 897 [PMID: 32850745 DOI: 10.3389/fbioe.2020.00897]

12. Fan F, Chen D, Zhao Y, Wang H, Sun H, Sun K. Rapid preliminary purity evaluation of tumor biopsies using deep learning approach. Comput Struct Biotechnol J 2020; 18: 1746-1753 [PMID: 32695267 DOI: 10.1016/j.csbj.2020.06.007]

13. Yeh P, Hunter T, Sinha D, Fioni S, Wallach E, Jiang D, Chan YC, Wong SQ, Silva MJ, Vedururu R, Doig K, Lann E, Arnau GM, Sempel T, Wall M, Zivanovic A, Agarwal R, Petrone P, Jones K, Westerman D, Blombery P, Seymour JP, Papenfuss AT, Dawson MA, Tam CS, Dawson SJ. Circulating tumour DNA reflects treatment response and clonal evolution in chronic lymphocytic leukaemia. Nat Commun 2017; 8: 14756 [PMID: 28303895 DOI: 10.1038/ncomms14756]

14. Hao X, Luo H, Hraveczyk M, Wei W, Wang W, Wang J, Flagg K, Hou J, Zhang H, Yi S, Safari M, Lin D, Chang C, Caughhey BA, Li G, Dhar D, Shi W, Zheng L, Hou R, Zhu J, Zhao L, Fu X, Zhang E, Zhang C, Zhu JK, Karim M, Xu RH, Zhang K. DNA methylation markers for diagnosis and prognosis of common cancers. Proc Natl Acad Sci USA 2017; 114: 7414-7419 [PMID: 28652331 DOI: 10.1073/pnas.1703577114]

15. Buechler C, Aslanidis C. Role of lipids in pathophysiology, diagnosis and therapy of hepatocellular carcinoma. Biochim Biophys Acta Mol Cell Biol Lipids 2020; 1865: 158658 [PMID: 32058031 DOI: 10.1016/j.bbalip.2020.158658]

16. Jiang P, Sun K, Tong YK, Cheng SH, Cheng THT, Heung MMS, Wong J, Wong VWS, Chan HLY, Chan KCA, Lo YMD, Chiu RWK. Preferred end coordinates and somatic variants as signatures of circulating tumor DNA associated with hepatocellular carcinoma. Proc Natl Acad Sci USA 2018; 115: E10925-E10933 [PMID: 30373822 DOI: 10.1073/pnas.1814616115]

17. Zhang Q, Peng J, Ye M, Weng W, Tan C, Ni S, Huang D, Sheng W, Wang L. KRAS Mutation Predicted More Metastases and Closer Resection Margins in Patients with Colorectal Cancer
Liver Metastases. Ann Surg Oncol 2020; 27: 1164-1173 [PMID: 31720935 DOI: 10.1245/s10434-019-08065-2]

Shoda K, Ishikawa D, Fujita Y, Masuda K, Hiramoto H, Hamada J, Arita T, Konishi H, Komatsu S, Shiozaki A, Kakihara N, Okamoto K, Taniguchi H, Imoto I, Otsuji E. Monitoring the HER2 copy number status in circulating tumor DNA by droplet digital PCR in patients with gastric cancer. Gastric Cancer 2017; 20: 126-135 [PMID: 28579451 DOI: 10.1007/s10120-016-0599-z]

Chen Z, Zhang C, Zhang M, Li B, Niu Y, Chen L, Yang J, Lu S, Gao J, Shen L. Chromosomal instability of circulating tumor DNA reflect therapeutic responses in advanced gastric cancer. Cell Death Dis 2019; 10: 697 [PMID: 31541076 DOI: 10.1038/s41419-019-1907-4]

Chen KC, Jiang P, Chan CW, Sun K, Wong J, Hui EP, Chan SL, Chan WC, Hui DS, Ng SS, Chan HL, Wong CS, Ma BB, Chan AT, Lai PB, Sun H, Chiu RW, Lo YM. Noninvasive detection of cancer-associated genome-wide hypomethylation and copy number aberrations by plasma DNA bisulfite sequencing. Proc Natl Acad Sci USA 2013; 110: 18761-18768 [PMID: 24191000 DOI: 10.1073/pnas.1313959110]

Li H, Wang Z, Zhao G, Ma Y, Chen Y, Xue Q, Zheng M, Fei S. Performance of a MethyLight assay for methylated SFRP2 DNA detection in colorectal cancer tissue and serum. Int J Biol Markers 2019; 34: 54-59 [PMID: 30852954 DOI: 10.17272/ijbmr.2019.003056]

Gai W, Ji L, Lam WKJ, Sun K, Jiang P, Chan AWH, Wong J, Lai PBS, Ng SSM, Ma BBY, Wong GLH, Wong VWS, Chan HLY, Chiu RWK, Lo YMD, Chan KCA. Liver- and Colon-Specific DNA Methylation Markers in Plasma for Investigation of Colorectal Cancers with or without Liver Metastases. Clin Chem 2018; 64: 1239-1249 [PMID: 29903781 DOI: 10.1373/clinchem.2018.290304]

Eissa MAL, Lerner L, Abdelfatah E, Shankar N, Carner JK, Hasan NM, Yaghoubi H, Huang B, Kerner Z, Takaesu F, Wolfgang C, Kwak R, Ruiz M, Tam M, Pisnic TR 2nd, Iacobuzio-Donahue CA, Hruban RH, He J, Wang TH, Wood LD, Sharma A, Ahuja N. Promoter methylation of ADAMTS1 and BNC1 as potential biomarkers for early detection of pancreatic cancer in blood. Clin Epigenetics 2019; 11: 59 [PMID: 30953539 DOI: 10.1186/s13148-019-0650-0]

Nishizawa N, Harada H, Kumamoto Y, Kaizu T, Katoh H, Tajima H, Ushiki H, Yokoi K, Igarashi K, Fujiyama Y, Okuwaki K, Iwai T, Watanabe M, Yamashita K. Diagnostic potential of hypermethylation of the cysteine dioxygenase 1 gene (CDO1) promoter DNA in pancreatic cancer. Ann Surg Oncol 2011; 18: 2868-2855 [PMID: 31325200 DOI: 10.1111/j.1445-2960.2011.04511.x]

Kim Y, Wen X, Jeong S, Cho NY, Kim WH, Kang GH. Combinatory low methylation statuses of SAT-α and L1 are associated with shortened survival time in patients with advanced gastric cancer. Gastric Cancer 2019; 22: 37-47 [PMID: 29926315 DOI: 10.1007/s10120-018-0852-8]

Harada H, Hosoda K, Moriya H, Mieno H, Ema A, Ushiki H, Washio M, Nishizawa N, Ishii S, Yokota K, Tanaka Y, Kaidai T, Soeno T, Kosaka Y, Watanabe M, Yamashita K. Cancer-specific promoter DNA methylation of Cysteine dioxygenase type 1 (CDO1) gene as an important prognostic biomarker of gastric cancer. PLoS One 2019; 14: e0214872 [PMID: 30934021 DOI: 10.1371/journal.pone.0214872]

Zhang Q, Shan F, Li Z, Gao J, Li Y, Shen L, Ji J, Lu M. A prospective study on the changes and clinical significance of pre-operative and post-operative circulating tumor cells in resectable gastric cancer. J Transl Med 2018; 16: 171 [PMID: 29925382 DOI: 10.1186/s12967-018-1544-1]

Li W, Zhang X, Lu X, You L, Song Y, Luo Z, Zhang J, Nie J, Zheng W, Xu D, Wang Y, Dong Y, Yu S, Hong J, Shi J, Hao H, Luo F, Hua L, Wang P, Qian X, Yuan F, Wei L, Cui M, Zhang T, Liao Q, Dai M, Liu Z, Chen G, Meckel K, Adhihari S, Jia G, Bissonnette MB, Zhao Y, Zhang W, He C, Liu J. 5-Hydroxymethylcytosine signatures in circulating cell-free DNA as diagnostic biomarkers for human cancers. Cell Res 2017; 27: 1243-1257 [PMID: 28925386 DOI: 10.1038/cr.2017.121]

Cai J, Chen L, Zhang Z, Zhang X, Lu X, Liu W, Shi G, Ge Y, Gao P, Yang Y, Ke A, Xiao L, Dong R, Zhu Y, Yang X, Wang J, Zhu T, Yang D, Huang X, Sui C, Qiu S, Shen F, Sun H, Zhou W, Zhou J, Nie J, Zeng C, Stroup FK, Chuic BC, Lau WY, He C, Wang H, Zhang W, Fan J. Genome-wide mapping of 5-hydroxymethylcytosines in circulating cell-free DNA as a non-invasive approach for early detection of hepatocellular carcinoma. Gut 2019; 68: 2195-2205 [PMID: 31335876 DOI: 10.1136/gutjnl-2019-318882]

Xiao Z, Wu W, Wu C, Li M, Sun F, Zheng L, Liu G, Li X, Yun Z, Tang J, Yu Y, Luo S, Sun W, Feng X, Cheng Q, Tao X, Wu S, Tao J. 5-Hydroxymethylcytosine signature in circulating cell-free DNA as a potential diagnostic factor for early-stage colorectal cancer and precancerous adenoma. Mol Oncol 2021; 15: 138-150 [PMID: 33107199 DOI: 10.1002/1876-0261.12833]

Sun K, Jiang P, Chan KC, Wong J, Cheng YK, Liang RH, Chan WK, Ma ES, Chan SL, Cheng SH, Chan RW, Tong YK, Ng SS, Wong RS, Hui DS, Leung TN, Leung TY, Lai PB, Chiu RW, Lo YM. Plasma DNA tissue mapping by genome-wide methylation sequencing for noninvasive prenatal, cancer, and transplantation assessments. Proc Natl Acad Sci USA 2015; 112: E5503-E5512 [PMID: 26392541 DOI: 10.1073/pnas.1508736112]

Sun K, Jiang P, Wong AIC, Cheng YKY, Cheng SH, Zhang H, Chan KCA, Leung TY, Chiu RWK, Lo YMD. Size-tagged preferred ends in maternal plasma DNA shed light on the production mechanism and show utility in noninvasive prenatal testing. Proc Natl Acad Sci USA 2018; 115: E5106-E5114 [PMID: 29760053 DOI: 10.1073/pnas.1804134115]

Underhill HR, Kitzman JO, Hellwig S, Welker NC, Daza R, Baker DN, Gligorich KM, Rostomily RC, Bromer MP, Shendure J. Fragment Length of Circulating Tumor DNA. PLoS Genet 2016; 12: e1006162 [PMID: 27428049 DOI: 10.1371/journal.pgen.1006162]
34 Mouliere F, Chandrananda D, Piskorz AM, Moore EK, Morris J, Ahlborn LB, Lair R, Goranova T, Marass F, Heider K, Wan JCM, Supernat A, Hudecova I, Gounaris I, Ros S, Jimenez-Linan M, Garcia-Corbacho J, Patel K, Ostrup O, Murphy S, Eldridge MD, Gale D, Stewart GD, Burge J, Cooper WN, van der Heijden MS, Massie CE, Watts C, Corrie P, Pacey S, Brindle KM, Baird RD, Mao-Sørensen M, Parkinson CA, Smith CG, Brenton JD, Rosenfeld N. Enhanced detection of circulating tumor DNA by fragment size analysis. *Sci Transl Med* 2018; 10 [PMID: 30404863 DOI: 10.1126/scitransmed.aar021]

35 Snyder MW, Kircher M, Hill AJ, Daza RM, Shendure J. Cell-free DNA Comprises an In Vivo Nucleosome Footprint that Infoms Its Tissues-Of-Origin. *Cell* 2016; 164: 57-68 [PMID: 26771485 DOI: 10.1016/j.cell.2015.11.050]

36 Ulz P, Thallinger GG, Auer M, Graf R, Kaschofer K, Jahn SW, Abete L, Pritzau G, Petru E, Geigl JB, Heitzer E, Speicher MR. Inferring expressed genes by whole-genome sequencing of plasma DNA. *Nat Genet* 2016; 48: 1273-1278 [PMID: 27571261 DOI: 10.1038/ng.3648]

37 Jiang P, Sun K, Peng W, Cheng SH, Ni M, Yeung PC, Heung MMS, Xie T, Shang H, Zhou Z, Chan RWY, Wong J, Wong VWS, Poon LC, Leung TY, Lam WKJ, Chan JYK, Chan HLY, Chan KCA, Chiu RWK, Lo YMD. Plasma DNA End-Motif Profiling as a Fragmentomic Marker in Cancer, Pregnancy, and Transplantation. *Cancer Discov* 2020; 10: 664-673 [PMID: 32116020 DOI: 10.1158/2159-8290.CD-19-0622]

38 Sun K, Jiang P, Cheng SH, Cheng THT, Wong J, Wong VWS, Ng SS, Ma BBY, Leung TY, Chan SL, Mok TSK, Lai PBS, Chan HLY, Sun H, Chan KCA, Chiu RWK, Lo YMD. Orientation-aware plasma cell-free DNA fragmentation analysis in open chromatin regions infoms tissue of origin. *Genome Res* 2019; 29: 418-427 [PMID: 30808726 DOI: 10.1101/gr.247319.119]

39 Serpas L, Chan RWY, Jiang P, Ni M, Sun K, Rashidfarrokhi A, Soni C, Sisirak V, Lee WS, Cheng SH, Peng W, Chan KCA, Chiu RWK, Reizis B, Lo YMD. Dnaase1 L3 deletion causes aberrations in length and end-motif frequencies in plasma DNA. *Proc Natl Acad Sci USA* 2019; 116: 641-649 [PMID: 30593563 DOI: 10.1073/pnas.1815031116]

40 Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell* 2011; 144: 646-674 [PMID: 21376230 DOI: 10.1016/j.cell.2011.02.013]

41 Plaks V, Koopman CD, Werb Z. Cancer. Circulating tumor cells. *Science* 2013; 341: 1186-1188 [PMID: 24031008 DOI: 10.1126/science.1235226]

42 Heitzer E, Auer M, Gasch C, Pichler M, Ulz P, Hoffmann EM, Lax S, Waldispuehl-Geigl J, Mauermann O, Lackner C, Höfler G, Eissner F, Sill H, Sonamigg H, Pantel K, Riethdorf S, Bauernhofer T, Geigl JB, Speicher MR. Complex tumor genomes inferred from single circulating tumor cells by array-CGH and next-generation sequencing. *Cancer Res* 2013; 73: 2965-2975 [PMID: 23471846 DOI: 10.1158/0008-5472.CAN-12-4140]

43 Vignot S, Frampton GM, Sara JC, Yelensky R, Commo F, Brambilla C, Palmer G, Moro-Sibilot D, Ross JS, Cronin MT, André F, Stephens JP, Lazar V, Miller VA, Brambilla E. Next-generation sequencing reveals high concordance of recurrent somatic alternations between primary tumor and metastases from patients with non-small-cell lung cancer. *J Clin Oncol* 2013; 31: 2167-2172 [PMID: 23630207 DOI: 10.1200/JCO.2012.47.7737]

44 Hofman V, Ilie ML, Long E, Selva E, Bonnetaud C, Molina T, Vénissac N, Muroux X, Vielh P, Hofman P. Detection of circulating tumor cells as a prognostic factor in patients undergoing radical surgery for non-small-cell lung carcinoma: comparison of the efficacy of the CellSearch Assay™ and the isolation by size of epithelial tumor cell method. *BMC Cancer* 2016; 16: 477 [PMID: 27417942 DOI: 10.1186/s12885-016-2461-4]

45 Edd JF, Mishra A, Dubash TD, Herrera S, Mohammad R, Williams EK, Hong X, Mutlu BR, Walsh JR, Machado de Carvalho F, Aldickacti B, Nieman LT, Stott SL, Kapur R, Maheswaran S, Haber DA, Toner M. Microfluidic concentration and separation of circulating tumor cell clusters from large blood volumes. *Lab Chip* 2020; 20: 558-567 [PMID: 31934715 DOI: 10.1039/c9lc01122f]

46 Lee SW, Hyun KA, Kim SI, Kang JY, Jung HI. Enrichment of circulating tumor cells using a centrifugal affinity plate system. *J Chromatogr A* 2014; 1373: 25-30 [PMID: 23453456 DOI: 10.1016/j.chroma.2014.11.027]

47 Balasubramanian P, Kinders RJ, Kummars S, Gupta V, Hasagawa D, Menachery A, Lawrence SM, Wang L, Ferry-Gaok C, Davis D, Pachman RE, Tomaszewski JE, Doroshov JH. Antibody-independent capture of circulating tumor cells of non-epithelial origin with the ApoStream® system. *PLoS One* 2017; 12: e0217541 [PMID: 28403214 DOI: 10.1371/journal.pone.0217541]

48 Le Du F, Fujii T, Kida K, Davis DW, Park M, Liu DD, Wu W, Chavez-MacGregor M, Barcenas CH, Valero V, Tripathy D, Reuben JM, Ueno NT. EpCAM-independent isolation of circulating tumor cells with epithelial-to-mesenchymal transition and cancer stem cell phenotypes using the ApoStream® in patients with breast cancer treated with primary systemic therapy. *PLoS One* 2020; 15: e0229903 [PMID: 32214335 DOI: 10.1371/journal.pone.0229903]

49 Swennenhuis JF, Reuners J, Thys K, Aerssens J, Terstappen LW. Efficiency of whole genome amplification of single circulating tumor cells enriched by CellSearch and sorted by FACS. *Genome Med* 2013; 5: 106 [PMID: 24286536 DOI: 10.1186/gm1510]

50 Huebner H, Fasching PA, Gumbrech W, Jud S, Rauch C, Matzas M, Paulicka P, Friedrich K, Lux...
Li LS et al. Recent advances in gastrointestinal cancer diagnosis

MP, Volz B, Gass P, Hähler L, Meier-Stiegen F, Hartkopf A, Neubauer H, Altmstedt K, Beckmann MW, Fehm TN, Ruebner M. Filtration based assessment of CTCs and CellSearch® based assessment are both powerful predictors of prognosis for metastatic breast cancer patients. *BMC Cancer* 2018; 18: 204 [PMID: 29463222 DOI: 10.1186/s12885-018-4115-1]

52 Hendricks A, Brandt B, Geisen R, Dall K, Röder C, Schafnauer C, Becker T, Hinz S, Sebens S. Isolation and Enumeration of CTC in Colorectal Cancer Patients: Introduction of a Novel Cell Imaging Approach and Comparison to Cellular and Molecular Detection Techniques. *Cancers (Basel)* 2020; 12 [PMID: 32947003 DOI: 10.3390/cancers12092643]

53 Wang A, Li Z, Wang Q, Bai Y, Ji X, Fu T, Ji K, Yue Y, Han T, Wu X, Zhang J, Yang Y, Xu G, Bu Z, Ji J. Diagnostic value of negative enrichment and immune fluorescence in situ hybridization for intraperitoneal free cancer cells of gastric cancer. *Chin J Cancer Res* 2019; 31: 945-954 [PMID: 31949396 DOI: 10.21147/j.issn.1000-9604.2019.06.10]

54 Miyamoto DT, Lee RJ, Kalinicin M, LiCausi JA, Zheng Y, Chen T, Milner JD, Emmons E, Ho U, Broderick K, Silva E, Javaid S, Kwan TT, Hong X, Dahl DM, McGovern EF, Fjestadhuu JA, Smith MR, Sequist LV, Kapur R, Wu CL, Stott SL, Ting DT, Giobbe-Hurder A, Toner M, Maheswaran S, Haber DA. An RNA-Based Digital Circulating Tumor Cell Signature Is Predictive of Drug Response and Early Dissemination in Prostate Cancer. *Cancer Discov* 2018; 8: 288-303 [PMID: 29301747 DOI: 10.1158/2159-8829.CD-16-1460]

55 Winter M, Cai Z, Winklek K, Georgiou K, Inglis D, Lavranos T, Rezaei M, Warkiani M, Thierry B. Circulating tumour cell RNA characterisation from colorectal cancer patient blood after inertial microfluidic enrichment. *MethodX* 2019; 6: 1512-1520 [PMID: 31304099 DOI: 10.1016/j.mex.2019.06.012]

56 Xiang HM, Kim GH, Jeon HK, Kim DH, Jeon TY, Park DY, Jeong H, Chun WJ, Kim MH, Park J, Lim M, Kim TH, Cho YK. Circulating tumor cells detected by lb-on-a-disc: Role in early diagnosis of gastric cancer. *PLoS One* 2017; 12: e0180251 [PMID: 28662130 DOI: 10.1371/journal.pone.0180251]

57 Szczepanik A, Sierzega M, Drabik G, Pituch-Noworolska A, Kołodziejczyk P, Zembala M. CD44+ cytokeratin-positive tumor cells in blood and bone marrow are associated with poor prognosis of patients with gastric cancer. *Gastric Cancer* 2019; 22: 264-272 [PMID: 30056567 DOI: 10.1007/s10120-018-0858-2]

58 Wei T, Zhang X, Zhang Q, Yang J, Chen Q, Wang J, Li X, Chen J, Ma T, Li G, Gao S, Lou J, Qae R, Wang Y, Dang X, Zheng L, Liang T, Bai X. Vimentin-positive circulating tumor cells as a biomarker for diagnosis and treatment monitoring in patients with pancreatic cancer. *Cancer Lett* 2019; 452: 237-243 [PMID: 30905814 DOI: 10.1016/j.canlet.2019.03.069]

59 Dotan E, Alpaugh RK, Ruth K, Negin BP, Denlinger CS, Hall MJ, Astsaturov I, McAleer C, Futtipalli P, Thrash-Bingham C, Meropol NJ, Cohen SJ. Prognostic Significance of MUC-1 in Circulating Tumor Cells in Patients With Metastatic Pancreatic Adenocarcinoma. *Pancreas* 2016; 45: 1131-1135 [PMID: 26967453 DOI: 10.1097/MPA.0000000000000619]

60 Poruk KE, Blackford AL, Weiss MJ, Cameron JL, He J, Goggins M, Rasheed ZA, Wolfgang CL, Wood LD. Circulating Tumor Cells Expressing Markers of Tumor-Initiating Cells Predict Poor Survival and Cancer Recurrence in Patients with Pancreatic Ductal Adenocarcinoma. *Clin Cancer Res* 2017; 23: 2681-2690 [PMID: 27789528 DOI: 10.1158/1078-0432.CCR-16-1467]

61 Chen CJ, Sung WW, Chen HC, Chen YJ, Hsu HT, Lin YM, Lin SH, Peck K, Yeh KT. Early Assessment of Colorectal Cancer by Quantifying Circulating Tumor Cells in Peripheral Blood: ECT2 in Diagnosis of Colorectal Cancer. *Int J Mol Sci* 2017; 18 [PMID: 28363231 DOI: 10.3390/ijms18040743]

62 Fang C, Fan C, Wang C, Huang Q, Meng W, Yu Y, Yang L, Peng Z, Hu J, Li Y, Mo X, Zhou Z. CD133+CD54+CD44+ circulating tumor cells as a biomarker of treatment selection and liver metastasis in patients with colorectal cancer. *Oncotarget* 2016; 7: 77389-77403 [PMID: 27764803 DOI: 10.18652/oncotarget.12575]

63 Amantini C, Morelli MB, Nabissi M, Piva F, Marinelli O, Maggi F, Bianchi F, Bittoni A, Berardi R, Giampieri R, Santoni G. Expression Profiling of Circulating Tumor Cells in Pancreatic Ductal Adenocarcinoma Patients: Biomarkers Predicting Overall Survival. *Front Oncol* 2019; 9: 874 [PMID: 31552188 DOI: 10.3389/fonc.2019.00874]

64 Zhou Y, Wang B, Wu J, Zhang C, Zhou Y, Yang X, Zhou J, Guo W, Fan J. Association of preoperative EpCAM Circulating Tumor Cells and peripheral Treg cell levels with early recurrence of hepatocellular carcinoma following radical hepatic resection. *BMC Cancer* 2016; 16: 506 [PMID: 27439521 DOI: 10.1186/s12885-016-2526-4]

65 Mittelbrunn M, Sánchez-Madrid F. Intercellular communication: diverse structures for exchange of genetic information. *Nat Rev Mol Cell Biol* 2012; 13: 328-335 [PMID: 22510790 DOI: 10.1038/nrm3335]

66 Théry C, Zitvogel L, Amigorena S. Exosomes: composition, biogenesis and function. *Nat Rev Immunol* 2002; 2: 569-579 [PMID: 12154376 DOI: 10.1038/ari855]

67 Scavo MP, Pepola N, Tutino V, De Nunzio V, Ingrosso C, Rizzi F, Notarnicola M, Curri ML, Giannelli G. Exosomes for Diagnosis and Therapy in Gastrointestinal Cancers. *Int J Mol Sci* 2020; 21 [PMID: 31935918 DOI: 10.3390/ijms21010367]

68 Bebelman MP, Smit MJ, Pegtel DM, Baglio SR. Biogenesis and function of extracellular vesicles in cancer. *Pharmacol Ther* 2018; 188: 1-11 [PMID: 29476772 DOI: 10.1016/j.pharmthera.2018.02.013]
Li LS et al. Recent advances in gastrointestinal cancer diagnosis

69 Carpelan-Holmström MA, Haglund CH, Roberts PJ. Differences in serum tumor markers between colon and rectal cancer. Comparison of CA 242 and carcinoembryonic antigen. Dis Colon Rectum 1996; 39: 799-805 [PMID: 8674374 DOI: 10.1007/BF02054447]

70 Brzozowski JS, Jankowski H, Bond DR, McCague SB, Munro BR, Predebon MJ, Scarlett CJ, Skelding KA, Weidenhofer J. Lipidomic profiling of extracellular vesicles derived from prostate and prostate cancer cell lines. Lipids Health Dis 2018; 17: 211 [PMID: 30193584 DOI: 10.1186/s12944-018-0854-x]

71 Alzahrani FA, El-Magd MA, Abdelfattah-Hassan A, Saleh AA, Saaddelin IM, El-Shetry ES, Badawy AA, Alkarmir S. Potential Effect of Exosomes Derived From Cancer Stem Cells and MSCs on Progression of DEN-Induced HCC in Rats. Stem Cells Int 2018; 2018: 8058979 [PMID: 30224923 DOI: 10.1155/2018/8058979]

72 Xiao Y, Zhong J, Zhong B, Huang J, Jiang L, Jiang Y, Yuan J, Sun J, Dai L, Yang C, Li Z, Wang J, Zhong T. Exosomes as potential sources of biomarkers in colorectal cancer. Cancer Lett 2020; 476: 13-22 [PMID: 32044357 DOI: 10.1016/j.canlet.2020.01.033]

73 Lane RE, Korbie D, Hill MM, Trau M. Extracellular vesicles as circulating cancer biomarkers: opportunities and challenges. Clin Transl Med 2018; 7: 14 [PMID: 29855735 DOI: 10.1186/s40169-018-0192-7]

74 Nedaeein R, Manian M, yazayeri MH, Ranjbar M, Salehi R, Sharifi M, Mohaghegh F, Goli M, Jahednia SH, Avan A, Ghayour-Mobarhan M. Circulating exosomes and exosomal microRNAs as biomarkers in gastrointestinal cancer. Cancer Gene Ther 2017; 24: 48-56 [PMID: 27982021 DOI: 10.1038/cgt.2016.77]

75 Cho HY, Eun JW, Baek GO, Seo CW, Ahn HR, Kim SS, Cho SW, Cheong JY. Serum Exosomal MicroRNA, miR-10b-5p, as a Potential Diagnostic Biomarker for Early-Stage Hepatocellular Carcinoma. J Clin Med 2020; 9 [PMID: 31968558 DOI: 10.3390/jcm9101281]

76 Nakamura S, Sadakari Y, Ohtsuka T, Okayama T, Nakashima Y, Gotoh Y, Saeki K, Mori Y, Nakata K, Miyasaka Y, Onishi H, Oda Y, Goggins M, Nakamura M. Pancreatic Juice Exosomal MicroRNAs as Biomarkers for Detection of Pancreatic Ductal Adenocarcinoma. Ann Surg Oncol 2019; 26: 2104-2111 [PMID: 30820789 DOI: 10.1245/s10434-019-07629-z]

77 Cai C, Zhang H, Zhu Y, Zheng P, Xu Y, Sun J, Zhang M, Lan T, Gu B, Li S, Ma P, Serum Exosomal Long Noncoding RNA pesk2-2:1 As A Potential Novel Diagnostic Biomarker For Gastric Cancer. Onco Targets Ther 2019; 12: 10035-10041 [PMID: 31819499 DOI: 10.2147/OTT.S290033]

78 Li S, Zhang M, Zhang H, Hu K, Cai C, Wang J, Shi L, Ma P, Xu Y, Zheng P. Exosomal long noncoding RNA Inc-NGA-Q6-1 may serve as a diagnostic marker for gastric cancer. Clin Chim Acta 2020; 501: 252-257 [PMID: 31730812 DOI: 10.1016/j.cca.2019.10.047]

79 Yu S, Li Y, Liao Z, Wang Z, Qian L, Zhao J, Zong H, Kang B, Zou WB, Chen K, He X, Meng Z, Chen Z, Huang S, Wang P. Plasma extracellular vesicle long RNA profiling identifies a diagnostic signature for the detection of pancreatic ductal adenocarcinoma. Gut 2020; 69: 540-550 [PMID: 31562239 DOI: 10.1136/gutjnl-2019-318860]

80 Xu H, Dong X, Chen Y, Wang X. Serum exosomal hnrNPH1 mRNA as a novel marker for hepatocellular carcinoma. Clin Chem Lab Med 2018; 56: 479-484 [PMID: 29252188 DOI: 10.1515/cclm-2017-0327]

81 Baaarsi A, Nassar F, Mukherji D, Shamseddine A, Nasr R, Temraz S. Exosomal Non Coding RNA in LIQUID Biopsies as a Promising Biomarker for Colorectal Cancer. Int J Mol Sci 2020; 21 [PMID: 32092975 DOI: 10.3390/ijms21041398]

82 Redis RS,Calin S, Yang Y, You MJ, Calin GA. Cell-to-cell miRNA transfer: from body homeostasis to therapy. Pharmacol Ther 2012; 136: 169-174 [PMID: 22903157 DOI: 10.1016/j.pharmthera.2012.08.003]

83 Wu J, Li G, Wang Z, Yao Y, Chen R, Pu X, Wang J. Circulating MicroRNA-21 Is a Potential Diagnostic Biomarker in Gastric Cancer. Dis Markers 2015; 2015: 435656 [PMID: 26063956 DOI: 10.1155/2015/435656]

84 Huang Z, Zhu D, Wu L, He M, Zhou X, Zhang L, Zhang H, Wang W, Zhu J, Cheng W, Chen Y, Fan Y, Qi L, Yin Y, Zhu W, Shu Y, Liu P. Six Serum-Based miRNAs as Potential Diagnostic Biomarkers for Gastric Cancer. Cancer Epidemiol Biomarkers Prev 2017; 26: 188-196 [PMID: 28756776 DOI: 10.1158/1055-9965.EPI-16-0607]

85 Shimura T, Toden S, Kandimmalla R, Toiyama Y, Okagawa Y, Kanda M, Baba H, Koderaya Y, Kusumoki M, Goel A. Genomewide Expression Profiling Identifies a Novel miRNA-based Signature for the Detection of Peritoneal Metastasis in Patients With Gastric Cancer. Ann Surg 2019 [PMID: 31663973 DOI: 10.1097/SLA.0000000000003647]

86 Yamada A, Horimatsu T, Okagawa Y, Nishida N, Honjo H, Ida H, Kout T, Kusakata T, Sasaki Y, Yagi M, Higurashi T, Yukawa N, Amanuma Y, Kikuchi O, Muto M, Ueno Y, Nakajima A, Chiba T, Boland CR, Goel A. Serum miR-21, miR-29a, and miR-125b Are Promising Biomarkers for the Early Detection of Colorectal Neoplasia. Clin Cancer Res 2015; 21: 4234-4242 [PMID: 26038573 DOI: 10.1158/1078-0432.CCR-14-2793]

87 Zhang J, Lin H, Wang XY, Zhang DQ, Chen JX, Zhaung Y, Zheng XL. Predictive value of microRNA-143 in evaluating the prognosis of patients with hepatocellular carcinoma. Cancer Biomark 2017; 19: 257-262 [PMID: 28436387 DOI: 10.3233/CBM-160357]

88 Yu J, Ding WB, Wang MC, Guo XG, Xu J, Xu QG, Yang Y, Sun SH, Liu JF, Qin LX, Liu H, Yang F, Zhou WP. Plasma circular RNA panel to diagnose hepatitis B virus-related hepatocellular carcinoma: A large-scale, multicenter study. Int J Cancer 2020; 146: 1754-1763 [PMID: 31456215]
Li LS et al. Recent advances in gastrointestinal cancer diagnosis

DOI: 10.1002/jic.32647

Kourou K, Exarchos TP, Exarchos KP, Karamouzis MV, Fotiadis DJ. Machine learning applications in cancer prognosis and prediction. Comput Struct Biotechnol J 2015; 13: 8-17 [PMID: 25750696 DOI: 10.1016/j.csbj.2014.11.005]

Jiménez Pérez M, Grande RG. Application of artificial intelligence in the diagnosis and treatment of hepatocellular carcinoma: A review. World J Gastroenterol 2020; 26: 5617-5628 [PMID: 33898156 DOI: 10.3748/wjg.v26.i27.5617]

Jin P, Ji X, Kang W, Li Y, Liu H, Ma F, Ma S, Hu H, Li W, Tian Y. Artificial intelligence in gastric cancer: a systematic review. J Cancer Res Clin Oncol 2020; 146: 2339-2350 [PMID: 32613386 DOI: 10.1007/s00432-020-03304-9]

Thakur N, Yoon H, Chong Y. Current Trends of Artificial Intelligence for Colorectal Cancer Pathology Image Analysis: A Systematic Review. Cancers (Basel) 2020; 12 [PMID: 32668721 DOI: 10.3390/cancers12071884]

Wang Y, He X, Nie H, Zhou J, Cao P, Ou C. Application of artificial intelligence to the diagnosis and therapy of colorectal cancer. Am J Cancer Res 2020; 10: 3575-3598 [PMID: 33294256]

Esteva A, Kuprel B, Novos RA, Ko J, Swetter SM, Blau HM, Thrun S. Dermatologist-level classification of skin cancer with deep neural networks. Nature 2017; 542: 115-118 [PMID: 28117445 DOI: 10.1038/nature21056]

Cheplygina V, de Bruijne M, Pluim JPW. Not-so-supervised: A survey of semi-supervised, multi-instance, and transfer learning in medical image analysis. Med Image Anal 2019; 54: 280-296 [PMID: 30959445 DOI: 10.1016/j.media.2019.03.009]

Joseph DA, Meester RG, Zauber AG, Manninen DL, Winges L, Dong FB, Peaker B, van Ballegooijen M. Colorectal cancer screening: Estimated future colonoscopy need and current volume and capacity. Cancer 2016; 122: 2479-2486 [PMID: 27204081 DOI: 10.1002/cncr.30670]

Luo H, Xu G, Li C, He L, Luo L, Wang Z, Jing B, Deng Y, Jin Y, Li Y, Li B, Tan W, He C, Sescrojeni E, Wu Q, Huang J, Huang DW, Chen B, Lin SB, Chen QM, Yuan CM, Chen HX, Pu HY, Zhou F, He Y, Xu RH. Real-time artificial intelligence for detection of upper gastrointestinal cancer by endoscopy: a multicentre, case-control, diagnostic study. Lancet Oncol 2019; 20: 1645-1654 [PMID: 31591062 DOI: 10.1016/S1470-2045(19)30637-0]

Zhen SH, Cheng M, Tao YB, Wang YF, Juengpanich S, Jiang ZY, Jiang YK, Yan YY, Lu W, Lue JM, Qian JH, Wu ZY, Sun JH, Lin H, Cai XJ. Deep Learning for Accurate Diagnosis of Liver Tumor Based on Magnetic Resonance Imaging and Clinical Data. Front Oncol 2020; 10: 680 [PMID: 32547939 DOI: 10.3389/fonc.2020.00680]

Kiani A, Uyumnazturk B, Rajpurkar P, Wang A, Gao R, Jones E, Liu SY, Langloitz CP, Ball RL, Monteijn TJ, Martin BA, Berry GJ, Ozawa MG, Hazard FK, Brown RA, Chen SB, Wood M, Allard LS, Ylangan L, Ng AY, Shen J. Impact of a deep learning assistant on the histopathologic classification of liver cancer. NPJ Digit Med 2020; 3: 23 [PMID: 32140566 DOI: 10.1038/s41746-020-0232-8]

Chen M, Zhang B, Topatana W, Cao J, Zhu H, Juengpanich S, Mao Q, Yu H, Cai X. Classification and mutation prediction based on histopathology H&E images in liver cancer using deep learning. NPJ Precis Oncol 2020; 4: 14 [PMID: 32500270 DOI: 10.1038/s41698-020-0120-3]

Dimirou N, Arzanajlovic O, Caie PD. Deep Learning for Whole Slide Image Analysis: An Overview. Front Med (Lausanne) 2019; 6: 264 [PMID: 31824952 DOI: 10.3389/fmed.2019.00264]

Camppanella G, Hanna MG, Geneslaw L, Mirafiori A, Wernack Krauss Silva V, Busam KJ, Brogi E, Reuter VE, Kimstra DS, Fuchs TJ. Clinical-grade computational pathology using weakly supervised deep learning on whole slide images. Nat Med 2019; 15: 1301-1309 [PMID: 31380507 DOI: 10.1038/s41591-019-0508-1]

Sun K, Wang J, Wang H, Sun H. GeneCET: a generalizable cancerous status and tissue origin classifier for pan-cancer biopsies. Bioinformatics 2018; 34: 4129-4130 [PMID: 29947737 DOI: 10.1093/bioinformatics/bty524]

Yamamoto H, Watanabe Y, Oikawa R, Morita R, Yoshiya Y, Maehata T, Yasuda H, Inoh H. BARHL2 Methylation Using Gastric Wash DNA or Gastric Juice Exosomal DNA is a Useful Marker For Early Detection of Gastric Cancer in an H pylori-Independent Manner. Clin Transl Gastroenterol 2016; 7: e184 [PMID: 27441821 DOI: 10.1038/ctg.2016.40]

Lin LY, Yang L, Zeng Q, Wang L, Chen ML, Zhao ZH, Ye GD, Luo QC, Lv PY, Guo QW, Li BA, Cai JC, Cai WY. Tumor-originated exosomal Inc/UGEC1 as a circulating biomarker for early-stage gastric cancer. Mol Cancer 2018; 17: 84 [PMID: 29690888 DOI: 10.1186/s12943-018-0834-9]

Oehme F, Krahl S, Gyorffy B, Mueslin B, Rao V, Greif H, Ziegler N, Lin K, Thepklaysane ML, Polster H, Tonini T, Schneider M, Weitz J, Baenke F, Kahlert C. Low level of exosomal long non-coding RNA HOTTIP is a prognostic biomarker in colorectal cancer. RNA Biol 2019; 16: 1339-1345 [PMID: 31251124 DOI: 10.1080/15476286.2019.1637697]

Shi Y, Wang Z, Zhu X, Chen L, Ma Y, Wang J, Yang X, Liu Z. Exosomal miR-1246 in serum as a potential biomarker for early diagnosis of gastric cancer. Int J Clin Oncol 2020; 25: 89-99 [PMID: 31566750 DOI: 10.1007/s10147-019-01532-9]

Feng W, Zong W, Li Y, Shen X, Cui X, Ju S. Abnormally expressed long noncoding RNA B3GALT5-AS1 may serve as a biomarker for the diagnostic and prognostic of gastric cancer. J Cell Biochem 2020; 121: 557-565 [PMID: 31338903 DOI: 10.1002/jcb.29296]

Sun Y, Yang B, Lin M, Yu H, Chen H, Zhang Z. Identification of serum miR-30a-5p as a diagnostic and prognostic biomarker in colorectal cancer. Cancer Biomark 2019; 24: 299-305 [PMID: 31093126]
110 Liu X, Pan B, Sun L, Chen X, Zeng K, Hu X, Xu T, Xu M, Wang S. Circulating Exosomal miR-27a and miR-130a Act as Novel Diagnostic and Prognostic Biomarkers of Colorectal Cancer. Cancer Epidemiol Biomarkers Prev 2018; 27: 746-754 [PMID: 29739748 DOI: 10.1158/1055-9965.EPI-18-0067]

111 Wang J, Yan F, Zhao Q, Zhan F, Wang R, Wang L, Zhang Y, Huang X. Circulating exosomal miR-125a-3p as a novel biomarker for early-stage colon cancer. Sci Rep 2017; 7; 4150 [PMID: 28646161 DOI: 10.1038/s41598-017-04386-1]

112 Sun B, Li Y, Zhou Y, Ng TK, Zhao C, Gan Q, Gu X, Xiang J. Circulating exosomal CPNE3 as a diagnostic and prognostic biomarker for colorectal cancer. J Cell Physiol 2019; 234: 1416-1425 [PMID: 30078189 DOI: 10.1002/jcp.26936]

113 Court CM, Ankeny JS, Sho S, Winograd P, Hou S, Song M, Wainberg ZA, Girgis MD, Graeber TG, Agopian VG, Tseng HR, Tomlinson JS. Circulating Tumor Cells Predict Occult Metastatic Disease and Prognosis in Pancreatic Cancer. Ann Surg Oncol 2018; 25: 1000-1008 [PMID: 29442211 DOI: 10.1245/s10434-017-6290-8]

114 Takahasi K, Iinuma H, Wada K, Minezaki S, Kawamura S, Kainuma M, Ikeda Y, Shibuya M, Miura F, Sano K. Usefulness of exosome-encapsulated microRNA-451a as a minimally invasive biomarker for prediction of recurrence and prognosis in pancreatic ductal adenocarcinoma. J Hepatobiliary Pancreat Sci 2018; 25: 155-161 [PMID: 29130611 DOI: 10.1002/jhbp.524]

115 Wang Z. Diagnostic performance for declined microRNA-133a in pancreatic cancer. J Cell Biochem 2019 [PMID: 31773791 DOI: 10.1002/jcb.29547]

116 Wang F, Wang J, Ju L, Chen L, Cai W, Yang J. Diagnostic and prognostic potential of serum miR-132/212 cluster in patients with hepatocellular carcinoma. Ann Clin Biochem 2018; 55: 576-582 [PMID: 29357677 DOI: 10.1177/0004563218755815]
