Histopathology

Histopathology 2020, **77**, 340–350. DOI: 10.1111/his.14120

REVIEW

Gastrointestinal tissue-based molecular biomarkers: a practical categorisation based on the 2019 World Health Organization classification of epithelial digestive tumours

Javier I Quezada-Marín,[1,2] Alfred K Lam,[3] Atsushi Ochiai,[4] Robert D Odze,[5] Kay M Washington,[6] Masashi Fukayama,[7] Massimo Rugge,[8] David S Klimstra,[9] Iris D Nagtegaal,[10] Puay-Hoon Tan,[11] Mark J Arends,[12] John R Goldblum,[13] Ian A Cree[14] & Manuel Salto-Tellez[1,15]

[1]*Precision Medicine Centre of Excellence, Centre for Cancer Research and Cell Biology, Queen's University Belfast, Belfast, UK,* [2]*Molecular Pathology Laboratory, Anatomical Pathology Service, Puerto Montt Hospital, Puerto Montt, Chile,* [3]*Cancer Molecular Pathology, School of Medicine, Griffith University, Gold Coast, Queensland, Australia,* [4]*Exploratory Oncology Research and Clinical Trial Centre, National Cancer Centre, Kashiwa, Japan,* [5]*Dr Robert Odze Pathology LLC, Boston, MA, USA,* [6]*Department of Pathology, Vanderbilt University Medical Center, Nashville, TN, USA,* [7]*Department of Pathology, Graduate School of Medicine, University of Tokyo, Tokyo, Japan,* [8]*Surgical Pathology and Cytopathology Unit, Department of Medicine (DIMED), University of Padova, Padova, Italy,* [9]*Department of Pathology, Memorial Sloan-Kettering Cancer Center, New York, NY, USA,* [10]*Department of Pathology, Radboud University Medical Centre, Nijmegen, The Netherlands,* [11]*Division of Pathology, Singapore General Hospital, Singapore, Singapore,* [12]*Division of Pathology, Cancer Research UK Edinburgh Centre, Western General Hospital, Edinburgh, UK,* [13]*Department of Anatomic Pathology, Cleveland Clinic Foundation, Cleveland, OH, USA,* [14]*International Agency for Research on Cancer,World Health Organization, Lyon, France, and* [15]*Cellular Pathology, Belfast Health and Social Care Trust, Belfast, UK*

Quezada-Marín J I, Lam A K, Ochiai A, Odze R D, Washington K M, Fukayama M, Rugge M, Klimstra D S, Nagtegaal I D, Tan P-H, Arends M J, Goldblum J R, Cree I A & Salto-Tellez M.
(2020) *Histopathology* **77**, 340–350. https://doi.org/10.1111/his.14120

Gastrointestinal tissue-based molecular biomarkers: a practical categorisation based on the 2019 World Health Organization classification of epithelial digestive tumours

Molecular biomarkers have come to constitute one of the cornerstones of oncological pathology. The method of classification not only directly affects the manner in which patients are diagnosed and treated, but also guides the development of drugs and of artificial intelligence tools. The aim of this article is to organise and update gastrointestinal molecular biomarkers in order to produce an easy-to-use guide for routine diagnostics. For this purpose, we have extracted and reorganised the molecular information on epithelial neoplasms included in the 2019 *World Health Organization classification of tumours. Digestive system tumours*, 5th edn.

Keywords: biomarkers, gastrointestinal neoplasms, genomics, molecular pathology, WHO

Introduction

In the era of molecular medicine, with the expansion of digital pathology and the revolution in artificial intelligence (AI), molecular biomarker classifications of cancer

Address for correspondence: Professor Manuel Salto-Tellez, Precision Medicine Centre of Excellence, Centre for Cancer Research and Cell Biology, Queen's University Belfast, 97 Lisburn Road, Belfast BT9 7AE, UK. e-mail: m.salto-tellez@qub.ac.uk

are more important than ever before.[1] A rational cancer taxonomy is necessary to standardise diagnoses, make decisions on biomarker/drug development, and generate an appropriate background for AI tools.[2] Molecular biomarkers have a prominent role in oncological pathology, with diagnostic, predictive and/or prognostic value. A single and unified classification for biomarkers is important in order to collect relevant information and keep it updated. This represents a challenge for modern (morphomolecular) pathologists.[3]

The value of biomarkers in routine tissue diagnostics

A 'biomarker' is defined as 'any substance, structure, or process that can be measured in the body or its products and influence or predict the incidence of outcome or disease'.[4] Every year, there are between 15 000 and 20 000 new scientific articles on cancer biomarkers.[5] Unfortunately, of every 100 such biomarkers, <1% make it into a form that is useful for patient diagnosis or stratification,[6] mostly for a variety of scientific and technical reasons.[7]

As a result, there is no clear-cut evidence in the literature regarding which biomarkers are essential for diagnostics and/or therapeutic decision-making. However, the World Health Organization (WHO) classification of tumours consensus of international experts represents the best indication of how relevant these biomarkers are in routine diagnostic practice. Our goal is to summarise the use of these biomarkers in the gastrointestinal system (from the oesophagus to the anal canal), and obtain indications of the specific weight, form and relevance of biomarker analysis in disease taxonomy and clinical decision-making.

Current biomarker classification and proposed subcategorisation

A biomarker is a characteristic that is objectively measured and evaluated as an indicator of normal/pathogenic processes or pharmacological therapeutic responses.[8] The 2019 *World Health Organization classification of tumours. Digestive system tumours*, 5th edn. includes diagnostic, predictive and prognostic molecular biomarkers as the major categories.[9] Diagnostic biomarkers are intended to help pathologists establish a specific diagnosis; predictive markers indicate the probability of patients benefiting from a specific therapy; and prognostic markers determine the outcomes of patients, in the absence of specific treatments.[8]

The decision on which group a biomarker belongs to represents the first step of assessment. Some biomarkers may fulfil the criteria for more than one category. Additionally, it is possible to subdivide categories, improving their organisation and comprehension. In this manner, for a diagnostic biomarker, the second step is to determine whether it is useful in differential diagnosis or whether it contributes to cancer classification. For predictive biomarkers, the following question should be asked. Are there definitive randomised clinical trials or cohort studies that support their efficacy? If the answer is yes, then these would correspond to established predictive biomarkers. If the answer is negative, but they are currently under investigation, one may classify them as potentially predictive biomarkers. If they are not yet associated with any clinical trial, we propose to label them with the term 'preclinical predictive biomarkers'. In this circumstance, it is unlikely that they would be placed within one of the first-level groups within the WHO classification of tumours, although they often have relevance to the understanding of tumour pathogenesis, and may be included under this topic. For prognostic biomarkers, the main question is whether they are specific prognostic markers for a certain entity, or whether they are used to create risk stratification groups. Biomarkers that do not fit into any of these categories should be classified as 'others'. This classification is summarised in Figure 1.

However, in order to understand, adequately categorise and subcategorise these biomarkers, it is necessary to methodically evaluate other attributes associated with them.

Variables to consider in the categorisation of biomarkers

CONTEXT: SYSTEM, ORGAN, AND ENTITY

Context is a relevant aspect to consider in any biomarker assessment. A specific biomarker can have different attributes according to the location (system/organ) and the disease (entity) in question. For example, the presence of epidermal growth factor receptor (*EGFR*) gene activating mutations in oesophageal squamous cell carcinoma (SCC) is an adverse prognostic factor, and EGFR-targeted therapies have failed to improve survival.[10-12] The same molecular alterations in non-small-cell lung carcinoma (NSCLC) confer a better prognosis and also provide the patient with an opportunity to receive tyrosine kinase inhibitor therapy with a significant chance of improved survival.[13] Activating mutations of v-raf murine

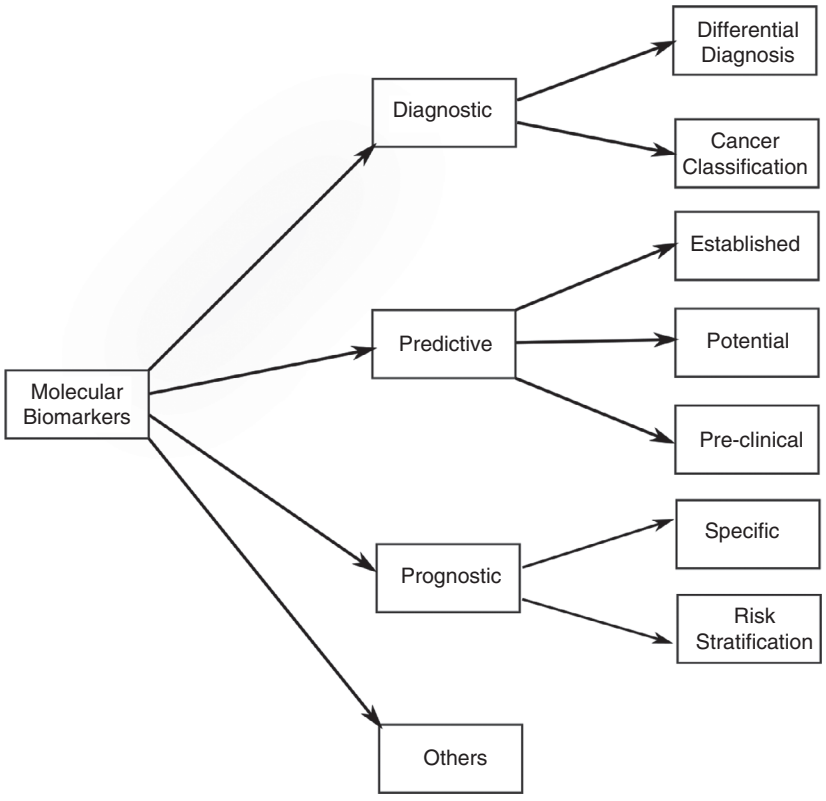

Figure 1. Proposed categorisation of molecular biomarkers. The subdivision into specific categories and subcategories improves the organisation, facilitates their comprehension, and allows an adequate update of molecular biomarkers.

sarcoma viral oncogene homologue B1 (*BRAF*) in colorectal carcinoma (CRC) may confer resistance to anti-EGFR therapy and help to establish a worse prognosis,[14] whereas the same alterations (*BRAF* V600E activating mutation) in melanoma predict response to treatment with BRAF inhibitors, such as vemurafenib.[15] Sometimes, the existence of certain molecular alterations is known, but there are no clinical trials available to support their routine use. On other occasions, support exists for a specific organ, but not for others. An example of this is human epidermal growth factor receptor 2 gene (*HER2*; *ERBB2*) mutation in small-intestine adenocarcinoma, which can be detected. In theory, a patient with cancer harbouring the mutation will obtain benefit from anti-ERBB2 therapy, but it does not yet have an established predictive value.[16] On the other hand, alterations in the same biomarker in oesophagus/oesophagogastric junction adenocarcinoma are predictive of response to this targeted therapy.[17]

STATUS: SPECIFIC MOLECULAR ALTERATIONS

Evaluation of the status of a biomarker implies specification of the molecular alterations that give it clinical utility. Determination of the specific alteration (activating mutation, translocation, overexpression, etc.) conceptually corresponds to the exact molecular phenomenon involved. In this manner, the lack of *KRAS*/*NRAS* activating mutations in CRC predicts a favourable response to anti-EGFR therapy.[14] In contrast, the presence (but not the lack) of activating *KIT* mutations in other malignant neoplasms, such as melanoma or gastrointestinal stromal tumours, predicts response to imatinib therapy.[18]

LEVEL OF DETECTION

The level at which the alteration is detected is also crucial in the evaluation of biomarker status. The clinical utility of the biomarker can be detected at the genetic, the transcriptomic and/or at the protein level; specific mutations of *KRAS*, *NRAS* and *BRAF* in CRC are good examples of genetic-level detection.[19] Some alterations detected at the transcriptomic level are oncotype Dx in breast cancer[20] and consensus molecular subtypes in CRC.[21] At the protein level, examples are mesenchymal–epithelial transition factor (c-MET) in CRC and anaplastic lymphoma kinase fusion in NSCLC; both can be detected by immunohistochemistry.[22,23] To add more complexity, a biomarker can be detected at different levels with different degrees of

clinical significance, independently of the system/organ where it occurs. An example of this is *ERBB2* in NSCLC, in which mutation is not associated with *ERBB2* amplification or ERBB2 overexpression, suggesting a distinct entity and a potential different therapeutic target.[24] Conversely, evaluation of *ERBB2* in gastric/gastro-oesophageal junction adenocarcinomas and breast carcinomas shows that gene amplification and protein overexpression are both useful in the prediction of targeted therapies.[17,25]

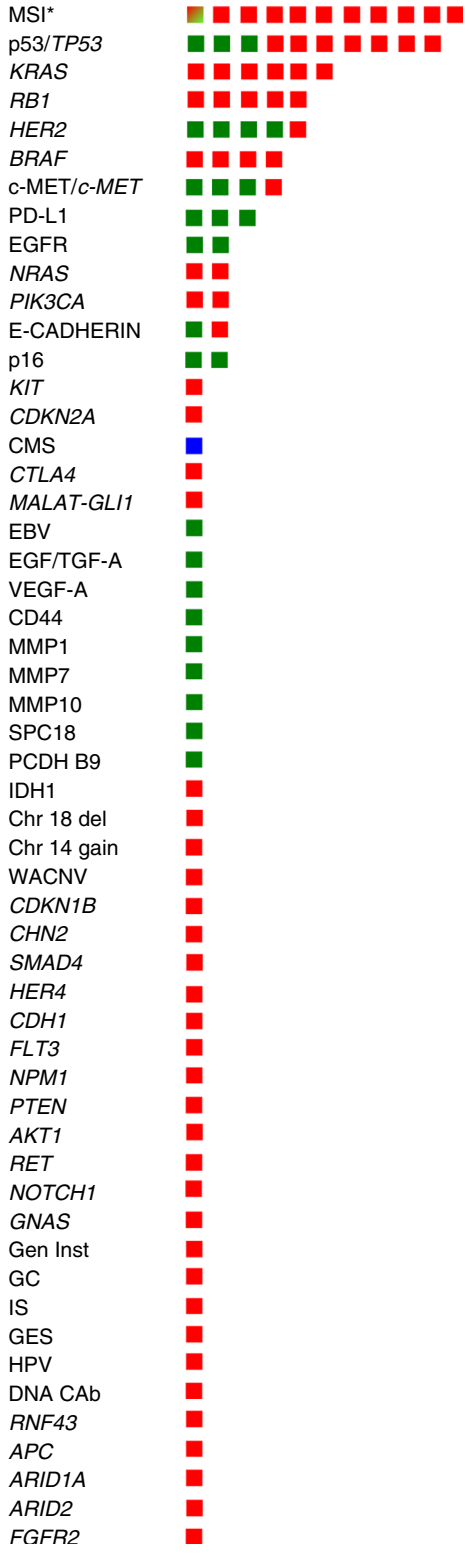

Figure 2. Gastrointestinal molecular biomarker frequency by detection technology. Each square represents an individual count of a molecular biomarker in the gastrointestinal system. The technologies of detection are: immunohistochemical (IHC) test (green); transcriptomic tool (blue); and DNA-based mutational assays (red). *In most cases, microsatellite instability (MSI) is studied with DNA-based mutational assays. However, in some organs, such as the large intestine, IHC testing is useful as a diagnostic biomarker of Lynch syndrome, whereas DNA-based mutational assays are used for predictive and prognostic analysis (red/green square). AAdC, ampullary adenocarcinoma; AdC, adenocarcinoma; *AKT1*, AKT serine-threonine kinase 1 gene; ApAC, appendiceal adenocarcinoma; *APC*, adenomatous polyposis coli gene; *ARID1A*, AT-rich interactive domain-containing protein 1A gene; *ARID2*, AT-rich interactive domain-containing protein 2 gene; ASD, anal squamous dysplasia; BD, Barret dysplasia; *BRAF*, serine/threonine protein kinase B-raf gene; CD44, CD44 antigen; *CDH1*, cadherin 1 gene; *CDK1B*, cyclin-dependent kinase inhibitor 1B gene; *CDKN2A*, cyclin-dependent kinase inhibitor 2A gene; *CHN2*, chimerin 2 gene; Chr 14 gain, chromosome 14 gain; Chr 18 del, chromosome 18 deletion; c-MET, tyrosine-protein kinase Met receptor; CMS, consensus molecular group; CTLA4, cytotoxic T-lymphocyte antigen 4; *CTNNB1*, catenin β1 gene; DNA CAb, deoxyribonucleic acid content abnormalities; EBV, Epstein–Barr virus; EGF, epidermal growth factor; EGFR, epidermal growth factor receptor; ERBB2, human epidermal growth factor receptor 2; *FGFR2*, fibroblast growth factor receptor 2 gene; *FLT3*, fms-like tyrosine kinase 3 gene; GB, gastroblastoma; GC, genomic classification; GD, gastric dysplasia; Gen Inst, genomic instability; GES, gene expression signature; *GNAS*, guanine nucleotide-binding protein G gene; *HER4*, human epidermal growth factor receptor 4 gene; HPV, human papilloma virus; *IDH1*, isocitrate dehydrogenase 1 gene; IS, Immunoscore; *KIT*, KIT proto-oncogene, receptor tyrosine kinase; KRAS, Kirsten rat sarcoma viral oncogene homologue; *MALAT1–GLI1*, MALAT1–GLI1 fusion gene; *MET*, tyrosine-protein kinase Met gene; *MLH1*, MutL homologue 1 gene; MMP1, matrix metalloproteinase-1; MMP7, matrix metalloproteinase-7; MMP10, matrix metalloproteinase-10; NAAdC, non-ampullary adenocarcinoma; NEN, neuroendocrine neoplasm; *NOTCH1*, Notch homologue 1 translocation-associated gene; *NPM1*, nucleophosmin 1 gene; *NRAS*, NRAS proto-oncogene GTPase; p16, protein p16; p53, tumour suppressor protein p53; PCDH B9, protocadherin B9; PD-L1, programmed death-ligand 1; *PIK3CA*, phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit-α gene; *PTEN*, phosphatase and tensin homologue gene; *RB1*, retinoblastoma 1 gene; *RET*, ret proto-oncogene; *RNF43*, ring finger protein 43 gene; SCC, squamous cell carcinoma; *SMAD4*, SMAD family member 4 gene; *SPC18*, septal pore cap protein 18 gene; TGF-α, transforming growth factor-α; *TP53*, tumour suppressor protein 53 gene; UC, undifferentiated carcinoma; VEGF-A, vascular endothelial growth factor A; WACNV, whole arm copy number variation.

Gastrointestinal system biomarkers update

GENERAL

A total of 54 different biomarkers are mentioned 98 times across the gastrointestinal tract chapters of the WHO blue book. Figure 2 summarises them, showing whether the technology used corresponds to immunohistochemical (IHC) tests, transcriptomic tools, or DNA-based mutational assays. Microsatellite instability (MSI) is described 11 times: one for diagnostic, four for prognostic and six for therapeutic purposes. p53/*TP53* is mentioned 10 times: three times as an IHC test, and seven as a DNA-based mutational analysis tool. *KRAS* study is indicated on six occasions, all of which correspond to DNA mutational analysis with predictive utility, with one use as a potential prognostic biomarker. In contrast, overexpression of ERBB2 is indicated five times, all involving detection by the use of IHC analysis, but with the consideration that some cases will require confirmation with fluorescence *in-situ* hybridisation. The organ with the most biomarkers mentioned was the small intestine/ampulla, with 31 different markers being mentioned, the vast majority being specific prognostic markers with potential for future use. Finally, the large intestine has more established biomarkers for current routine use than any other anatomical site, with 15 markers of diagnostic, established predictive and prognostic use.

The biomarkers discussed below are summarised by category, subcategory and organ in Tables 1 and 2.

OESOPHAGUS

The routinely used molecular biomarkers in oesophageal lesions include the presence of aberrant IHC expression of p53, which may be associated with better diagnostic reproducibility for dysplasia (differential diagnosis biomarker) and an increased risk of neoplastic progression (prognostic risk stratification biomarker), in the context of Barrett's oesophagus.[26,27] In addition, ERBB2 overexpression and/or *ERBB2* amplification in lower oesophagus/oesophagogastric junction adenocarcinoma carries a predictive value for response to ERBB2-targeted therapy (established predictive biomarker).[17]

There are other markers that are not yet used in routine pathological analysis, but may be important in the near future. EGFR overexpression in oesophageal SCC is considered to be an adverse specific prognostic factor,[10] because targeted therapies have not been successful in improving survival.[11,12] The loss of mismatch repair protein expression (with the consequent MSI) and overexpression of programmed death-ligand 1 (PD-L1) in oesophagus/oesophagogastric junction adenocarcinoma are linked to the potential use of immune checkpoint inhibitors. These checkpoint inhibitors are undergoing clinical trial evaluation for immunotherapy,[28] as is cytotoxic T-lymphocyte antigen 4 overexpression (potentially predictive biomarkers).[28,29]

Other possible biomarkers are methylation of the cyclin-dependent kinase inhibitor 2A gene (*CDKN2A*) (*p16*) promoter, which inhibits its gene expression, and genomic instability (specifically copy number alterations). Both biomarkers have potential value as prognostic risk stratification biomarkers in Barrett dysplasia,[27] but this is not yet supported by strong retrospective or prospective studies.

STOMACH

The molecular biomarkers with current use in gastric tumours include ERBB2 overexpression and/or *ERBB2* amplification in gastric adenocarcinoma and gastric undifferentiated carcinoma, with established predictive value for response to ERBB2-targeted therapy,[17,30,31] and the presence of the *MALAT1–GLI1* fusion gene for diagnostic confirmation of gastroblastoma, a rare gastric biphasic tumour that has recently been described (differential diagnosis biomarker).[32] *TP53* and retinoblastoma gene 1 (*RB1*) mutations also act as differential diagnosis biomarkers, helping to distinguish gastric neuroendocrine carcinomas (NECs) from G3 neuroendocrine tumours (NETs), in which these genes are more frequently wild type; this is also applicable for the remainder of the digestive organs.[33]

A plethora of markers are used in gastric cancer biology, specifically with regard to specific prognostic markers, with little direct routine application. These include EGFR and c-MET overexpression,[30] MSI,[34] Epstein–Barr virus detection,[35] high expression levels of epidermal growth factor/transforming growth factor-α, vascular endothelial growth factor-A and CD44, reduced expression of E-cadherin, expression of matrix metalloproteinase (MMP)-1, MMP7, and MMP10, up-regulation of SPC18 (*SEC11A*), and protocadherin B9 (*PCDHB9*) overexpression.[36-38] MSI and PD-L1 expression are potential predictive molecular biomarkers under investigation in clinical trials.[34,39,40]

Finally, for gastric dysplasia, there are biomarkers of disease progression that are seldom used routinely today. These are: DNA content abnormalities (aneuploidy or elevated 4N fraction)[41]; aberrant p53/*TP53*; mutations of *RNF43*, the adenomatous polyposis coli

Table 1. Preinvasive molecular biomarkers by category, subcategory, and evaluated organ of the gastrointestinal tract

Category	Subcategory	Oesophagus	Stomach	Small intestine/ampulla	Appendix	Coloreactal	Anal canal
Diagnostic	Differential diagnosis	p53 (BD)					
	Cancer classification						
Predictive	Established						
	Potential						
	Preclinical						
Prognostic	Specific						
	Risk stratification	p53 (BD)	DNA CAb (GD)				p16 (ASD)
		CDKN2A (BD)	p53/*TP53* (GD)				HPV (ASD)
		Gen Inst (BD)	*RNF43* (GD)				
			p16 (GD)				
			APC (GD)				
			ARID1A (GD)				
			ARID2 (GD)				
			MSI (GD)				

APC, adenomatous polyposis coli gene; *ARID1A*, AT-rich interactive domain-containing protein 1A gene; *ARID2*, AT-rich interactive domain-containing protein 2 gene; ASD, anal squamous dysplasia; BD, Barret dysplasia; *CDKN2A*, cyclin-dependent kinase inhibitor 2A gene; DNA CAb, deoxyribonucleic acid content abnormalities; GD, gastric dysplasia; Gen Inst, genomic instability; HPV, human papilloma virus; MSI, microsatellite instability; *RNF43*, ring finger protein 43 gene; *TP53*, tumour suppressor protein 53 gene.

gene (*APC*), the AT-rich interactive domain-containing protein 1A gene (*ARID1A*), and the AT-rich interactive domain-containing protein 2 gene (*ARID2*)[42,43]; and inactivation by promoter methylation of *p16* and *MLH1* (with consequent MSI).[44,45]

SMALL INTESTINE AND AMPULLA

In ampullary and non-ampullary adenocarcinomas, only MSI is a regularly used molecular biomarker. Its indications include immunotherapy selection (established predictive biomarker), determination of a possible hereditary origin (differential diagnosis biomarker), and as a specific prognostic parameter (early results show that MSI may improve overall survival).[46,47] Markers with potential specific prognostic value, but without regular pathological use, include *KRAS* activating mutations in ampullary adenocarcinoma,[19] chromosome 18 deletion, chromosome 14 gain, whole arm copy number variations, and cyclin-dependent kinase inhibitor 1B gene (*CDKN1B*) mutations in small intestine and ampullary neuroendocrine neoplasms.[48-50]

Other biomarkers with potential specific prognostic or predictive preclinical value in non-ampullary adenocarcinoma are *TP53*, the isocitrate dehydrogenase gene (*IDH*), the cadherin-1 gene (*CDH1*), the fibroblast growth factor receptor 2 gene (*FGFR2*), the fms-like tyrosine kinase 3 gene (*FLT3*), the nucleophosmin gene (*NPM1*), the phosphatase and tensin homologue gene (*PTEN*), *c-MET*, *AKT1*, *RET*, the Notch homologue 1 translocation-associated gene (*NOTCH1*), the human epidermal growth factor receptor 4 gene (*ERBB4*), the chimerin 2 gene (*CHN2*), *KRAS*, the SMAD family member 4 gene (*SMAD4*), *ERBB2*, and *CTNNB1*/E-cadherin.[16,51-54]

APPENDIX

Multiple studies have been conducted in appendiceal adenocarcinoma, but there is insufficient evidence to make firm recommendations regarding potentially predictive and preclinical biomarkers [*KRAS*, MSI, and the guanine nucleotide-binding protein G gene (*GNAS*)].[55-57] As in the rest of the digestive system, the presence of *TP53* and *RB1* mutations can help to

Table 2. Invasive molecular biomarkers by category, subcategory, and evaluated organ of the gastrointestinal tract

Category	Subcategory	Oesophagus	Stomach	Small intestine/ ampulla	Appendix	Colorectal	Anal canal
Diagnostic	Differential diagnosis		*MALAT1–GLI1* (GB)	MSI (AAdC–NAAdC)	*TP53* (NEN)	*BRAF* (AdC)	*TP53* (NEN)
			TP53 (NEN)	*TP53* (NEN)	*RB1* (NEN)	*TP53* (NEN)	*RB1* (NEN)
			RB1 (NEN)	*RB1* (NEN)		*RB1* (NEN)	
						MSI (CRC)	
	Cancer classification					CMS (AdC)	
						GC (AdC)	
Predictive	Established	ERBB2 (AdC)	ERBB2 (AdC–UC)	MSI (AAdC–NAAdC)		*KRAS* (AdC)	
						NRAS (AdC)	
						BRAF (AdC)	
						MSI (AdC)	
						PIK3CA (AdC)	
	Potential	MSI (AdC)	MSI (AdC)			*c-MET* (AdC)	PD-L1 (SCC)
		PD-L1 (AdC)	PD-L1 (AdC)				
		CTLA4 (AdC)					
	Preclinical	EGFR (SCC)		*BRAF* (NAAdC)	*KRAS* (ApAC)		*KRAS* (AdC)
				KRAS (NAAdC)	MSI (ApAC)		*NRAS* (AdC)
				IDH1 (NAAdC)	*GNAS* (ApAC)		MSI (AdC)
				ERBB2 (NAAdC)			
Prognostic	Specific		EGFR (AdC)	MSI (AAdC–NAAdC)		*BRAF* (AdC)	
			c-MET (AdC)	Chr 18 del (NEN)		MSI (AdC)	
		..	ERBB2 (AdC)	Chr 14 gain (NEN)		*PIK3CA* (AdC)	
			MSI (AdC)	WACNV (NEN)			
			EBV (AdC)	*CDKN1B* (NEN)			
			EGF/TGF-α (AdC)	*KRAS* (AAdC)			
			VEGF-A (AdC)	*TP53* (NAAdC)			
			CD44 (AdC)	*KRAS* (NAAdC)			
			E-cadherin (AdC)	*CHN2* (NAAdC)			
			MMP1 (AdC)	*SMAD4* (NAAdC)			
			MMP7 (AdC)	*KIT* (NAAdC)			
			MMP10 (AdC)	*HER4* (NAAdC)			
			SPC18 (AdC)	*CDH1* (NAAdC)			
			PCDH B9 (AdC)	*FGFR2* (NAAdC)			

Table 2. *(Continued)*

Category	Subcategory	Oesophagus	Stomach	Small intestine/ ampulla	Appendix	Colorectal	Anal canal
				FLT3 (NAAdC)			
				NPM1 (NAAdC)			
				PTEN (NAAdC)			
				c-MET (NAAdC)			
				AKT1 (NAAdC)			
				RET (NAAdC)			
				NOTCH1 (NAAdC)			
				CTNNB1/E-cadherin (NAAdC)			
				ERBB2 (NAAdC)			
	Risk stratification					IS (AdC)	
						GES (AdC)	

AAdC, ampullary adenocarcinoma; AdC, adenocarcinoma; *AKT1*, AKT serine-threonine kinase 1 gene; ApAC, appendiceal adenocarcinoma; *BRAF*, serine/threonine protein kinase B-raf gene; CD44, CD44 antigen; *CDH1*, cadherin 1 gene; *CDK1B*, cyclin-dependent kinase inhibitor 1B gene; *CHN2*, chimerin 2 gene; Chr 14 gain, chromosome 14 gain; Chr 18 del, chromosome 18 deletion; 0 c-MET, tyrosine-protein kinase Met receptor; *c-MET*, tyrosine-protein kinase Met receptor gene; CMS, consensus molecular group; CTLA4, cytotoxic T-lymphocyte antigen 4; EBV, Epstein–Barr virus; EGF, epidermal growth factor; EGFR, epidermal growth factor receptor; ERBB2, human epidermal growth factor receptor 2; *ERBB2*, human epidermal growth factor receptor 2 gene; *FGFR2*, fibroblast growth factor receptor 2 gene; *FLT3*, fms-like tyrosine kinase 3 gene; GB, gastroblastoma; GC, genomic classification; GES, gene expression signature; *GNAS*, guanine nucleotide-binding protein G gene; *HER4*, human epidermal growth factor receptor 4 gene; *IDH1*, isocitrate dehydrogenase 1 gene; IS, Immunoscore; *KIT*, KIT proto-oncogene, receptor tyrosine kinase; *KRAS*, Kirsten rat sarcoma viral oncogene homologue; *MALAT1–GLI1*, MALAT1–GLI1 fusion gene; MMP1, matrix metalloproteinase-1; MMP7, matrix metalloproteinase-7; MMP10, matrix metalloproteinase-10; MSI, microsatellite instability; NAAdC, non-ampullary adenocarcinoma; NEN, neuroendocrine neoplasm; *NOTCH1*, Notch homologue 1 translocation-associated gene; *NPM1*, nucleophosmin 1 gene; *NRAS*, NRAS proto-oncogene GTPase; p16, protein p16; p53, tumour suppressor protein p53; PCDH B9, protocadherin B9; PD-L1, programmed death-ligand 1; *PIK3CA*, phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit-α gene; *PTEN*, phosphatase and tensin homologue gene; *RB1*, retinoblastoma 1 gene; *RET*, ret proto-oncogene; SCC, squamous cell carcinoma; *SMAD4*, SMAD family member 4 gene; *SPC18*, septal pore cap protein 18 gene; TGF-α, transforming growth factor-α; *TP53*, tumour suppressor protein 53 gene; UC, undifferentiated carcinoma; VEGF-A, vascular endothelial growth factor A; WACNV, whole arm copy number variation.

distinguish appendiceal NECs from G3 NETs, in which these genes are more frequently wild type.[33]

COLON AND RECTUM

The molecular markers routinely used in CRC comprise the lack of activating mutations of *KRAS*/*NRAS* and *BRAF* (extended *RAS* testing), both of which have established predictive value for effective response to anti-EGFR therapy.[14,58] *BRAF* activating mutations have differential diagnostic utility in the exclusion of Lynch syndrome, and they are associated with an adverse specific prognosis.[59,60] MSI is an established predictive marker in colorectal adenocarcinoma: it is associated with a significant response to PD-L1 inhibitors for patients in whom conventional therapy has failed, confers a good prognosis to *BRAF* wild-type patients (specific prognostic marker), and is useful in Lynch syndrome diagnosis.[61]

In addition, two different methods are being used for colorectal adenocarcinoma molecular diagnostic classification: genomic-scale analysis (hypermutated or non-hypermutated colorectal cancers)[62] and transcriptomic profiling (consensus molecular subtypes for colorectal cancer).[21] Both have potential for subtype-based targeted therapies.

There are other markers in CRC that are not yet used currently, but show some promising results. c-MET overexpression has potential predictive value for response to c-MET inhibitors.[23] Phosphatidylinositol-

4,5-bisphosphate 3-kinase catalytic subunit-α gene (*PIK3CA*) activating mutations are associated with a worse specific prognosis, a negative predicted response to anti-EGFR therapy, and a successful adjuvant response to acetylsalicylic acid.[63,64] Immune-related markers such as the Immunoscore represent a potential prognostic stratification tool,[65,66] and gene expression signatures have more restricted prognostic use, specifically for determining the risk of recurrence after surgery.[67,68]

ANAL CANAL

The molecular biomarkers with clinical utility are p16 expression and polymerase chain reaction determination of high-risk human papilloma virus (HPV) genotypes (usually 16) in anal squamous dysplasia. The risk of progression from a high-risk lesion to SCC is influenced by HPV genotype, immune status, and other factors.[69-71]

Other markers with potential and/or preclinical predictive utility include PD-L1 expression in SCC [72,73] and *KRAS* and *NRAS* and MSI in anal adenocarcinoma.[74]

Conclusion

Despite the significant knowledge of the molecular basis of cancers of the digestive tract, there are relatively few biomarkers with established clinical utility, and most target common tumour types. Our review follows the new WHO approach to molecular markers, which is easily identifiable and can also be readily revisited when new information becomes available in the future. This systematic approach to the characterisation of new molecular markers may be used for the future taxonomy of cancers, which are also likely to benefit from computational pathology, especially within the next 5-year cycle of the WHO classification of tumours.

Conflicts of interest

D. Klimstra reports receiving personal fees from Merck, and personal fees from Paige.AI, outside the submitted work.

>Where authors are identified as personnel of the International Agency for Research on Cancer/World Health Organization, the authors alone are responsible for the views expressed in this article and they do not necessarily represent the decisions, policy or views of the International Agency for Research on Cancer/World Health Organization. The rest of the authors have anything to disclose related to this work.

Funding

There are no funding sources associated with this article.

Author contributions

J. Quezada-Marín: conceptualisation of biomarker classification, biomarker content preparation, and co-lead for manuscript preparation. A. K. Lam: biomarker content preparation and manuscript preparation. A. Ochiai: biomarker content preparation and manuscript preparation. R. Odze: biomarker content preparation and manuscript preparation. M. K. Washington: biomarker content preparation and manuscript preparation. M. Fukuyama: biomarker content preparation and manuscript preparation. M. Rugge: biomarker content preparation and manuscript preparation. D. S. Klimstra: biomarker content preparation and manuscript preparation. I. D. Nagtegaal: biomarker content preparation and manuscript preparation. P. H. Tan: biomarker content preparation and manuscript preparation. M. J. Arends: biomarker content preparation and manuscript preparation. J. R. Goldblum: biomarker content preparation and manuscript preparation. I. A. Cree: biomarker content preparation and co-lead for manuscript preparation. M. Salto-Tellez: conceptualisation of the paper; conceptualisation of the biomarker classification, biomarker content preparation, lead for manuscript preparation, and final editing.

References

1. Salto-Tellez M, Maxwell P, Hamilton P. Artificial intelligence—the third revolution in pathology. *Histopathology* 2019; **74**; 372–376.
2. Salto-Tellez M, Cree IA. Cancer taxonomy: pathology beyond pathology. *Eur. J. Cancer* 2019; **115**; 57–60.
3. Salto-Tellez M. A case for integrated morphomolecular diagnostic pathologists. *Clin. Chem.* 2007; **53**; 1188–1190.
4. *Biomarkers in risk assessment: validity and validation*. International Programme on Chemical Safety, IPCS I. Geneva: World Health Organization, 2001.
5. Everett BM, Ridker PM. Biomarkers for cardiovascular screening: progress or passé? *Clin. Chem.* 2017; **63**; 248–251.
6. Kern SE. Why your new cancer biomarker may never work: recurrent patterns and remarkable diversity in biomarker failures. *Cancer Res.* 2012; **72**; 6097–6101.
7. Salto-Tellez M, Kennedy RD. Integrated molecular pathology: the Belfast model. *Drug Discov. Today* 2015; **20**; 1451–1454.
8. Atkinson A, Colburn W, Degruttola V *et al.* Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. *Clin. Pharmacol. Ther.* 2001; **69**; 89–95.

9. *World Health Organization classification of tumours. Digestive system tumours*, 5th edn. Available at: http://apps.who.int/bookorders/anglais/detart1.jsp?codlan=1&codcol=70&codcch=24003# (accessed 23 July 2019).
10. Zhang W, Zhu H, Liu X *et al*. Epidermal growth factor receptor is a prognosis predictor in patients with esophageal squamous cell carcinoma. *Ann. Thorac. Surg.* 2014; **98**; 513–519.
11. Ohashi S, Miyamoto S, Kikuchi O *et al*. Recent advances from basic and clinical studies of esophageal squamous cell carcinoma. *Gastroenterology* 2015; **149**; 1700–1715.
12. Crosby T, Hurt CN, Falk S *et al*. Chemoradiotherapy with or without cetuximab in patients with oesophageal cancer (SCOPE1): a multicentre, phase 2/3 randomised trial. *Lancet Oncol.* 2013; **14**; 627–637.
13. Devarakonda S, Morgensztern D, Govindan R. Genomic alterations in lung adenocarcinoma. *Lancet Oncol.* 2015; **16**; e342–e351.
14. Hsu H-C, Thiam TK, Lu Y-J *et al*. Mutations of *KRAS*/*NRAS*/*BRAF* predict cetuximab resistance in metastatic colorectal cancer patients. *Oncotarget* 2016; **7**; 22257–22270.
15. Larkin J, Del Vecchio M, Ascierto PA *et al*. Vemurafenib in patients with BRAF(V600) mutated metastatic melanoma: an open-label, multicentre, safety study. *Lancet Oncol.* 2014; **15**; 436–444.
16. Alvi MA, McArt DG, Kelly P *et al*. Comprehensive molecular pathology analysis of small bowel adenocarcinoma reveals novel targets with potential for clinical utility. *Oncotarget* 2015; **6**; 20863–20874.
17. Grillo F, Fassan M, Sarocchi F *et al*. HER2 heterogeneity in gastric/gastroesophageal cancers: from benchside to practice. *World J. Gastroenterol.* 2016; **22**; 5879–5887.
18. Wei X, Mao L, Chi Z *et al*. Efficacy evaluation of imatinib for the treatment of melanoma: evidence from a retrospective study. *Oncol. Res.* 2019; **27**; 495–501.
19. Kwon MJ, Kim JW, Jung JP *et al*. Low incidence of KRAS, BRAF, and PIK3CA mutations in adenocarcinomas of the ampulla of Vater and their prognostic value. *Hum. Pathol.* 2016; **50**; 90–100.
20. McVeigh TP, Kerin MJ. Clinical use of the Oncotype DX genomic test to guide treatment decisions for patients with invasive breast cancer. *Breast Cancer* 2017; **9**; 393–400.
21. Guinney J, Dienstmann R, Wang X *et al*. The consensus molecular subtypes of colorectal cancer. *Nat. Med.* 2015; **21**; 1350–1356.
22. Qin A, Gadgeel S. The current landscape of anaplastic lymphoma kinase (ALK) in non-small cell lung cancer: emerging treatment paradigms and future directions. *Target. Oncol.* 2017; **12**; 709–718.
23. Bradley CA, Salto-Tellez M, Laurent-Puig P *et al*. Targeting c-MET in gastrointestinal tumours: rationale, opportunities and challenges. *Nat. Rev. Clin. Oncol.* 2017; **14**; 562–576.
24. Li BT, Ross DS, Aisner DL *et al*. HER2 amplification and HER2 mutation are distinct molecular targets in lung cancers. *J. Thorac. Oncol.* 2016; **11**; 414–419.
25. Liu Z-H, Wang K, Lin D-Y *et al*. Impact of the updated 2018 ASCO/CAP guidelines on HER2 FISH testing in invasive breast cancer: a retrospective study of HER2 FISH results of 2233 cases. *Breast Cancer Res. Treat.* 2019; **175**; 51–57.
26. Kaye PV, Haider SA, Ilyas M *et al*. Barrett's dysplasia and the Vienna classification: reproducibility, prediction of progression and impact of consensus reporting and p53 immunohistochemistry. *Histopathology* 2009; **54**; 699–712.
27. Zeki S, Fitzgerald RC. The use of molecular markers in predicting dysplasia and guiding treatment. *Best Pract. Res. Clin. Gastroenterol.* 2015; **29**; 113–124.
28. Shaib WL, Nammour JPA, Gill H *et al*. The future prospects of immune therapy in gastric and esophageal adenocarcinoma. *J. Clin. Med.* 2016; **5**; 100.
29. Bang Y-J, Cho JY, Kim YH *et al*. Efficacy of sequential ipilimumab monotherapy versus best supportive care for unresectable locally advanced/metastatic gastric or gastroesophageal junction cancer. *Clin. Cancer Res.* 2017; **23**; 5671–5678.
30. Fuse N, Kuboki Y, Kuwata T *et al*. Prognostic impact of HER2, EGFR, and c-MET status on overall survival of advanced gastric cancer patients. *Gastric Cancer* 2016; **19**; 183–191.
31. Motoshima S, Yonemoto K, Kamei H *et al*. Prognostic implications of HER2 heterogeneity in gastric cancer. *Oncotarget* 2018; **9**; 9262–9272.
32. Graham RP, Nair AA, Davila JI *et al*. Gastroblastoma harbors a recurrent somatic MALAT1–GLI1 fusion gene. *Mod. Pathol.* 2017; **30**; 1443–1452.
33. Rindi G, Klimstra DS, Abedi-Ardekani B *et al*. A common classification framework for neuroendocrine neoplasms: an International Agency for Research on Cancer (IARC) and World Health Organization (WHO) expert consensus proposal. *Mod. Pathol.* 2018; **31**; 1770–1786.
34. Zhu L, Li Z, Wang Y *et al*. Microsatellite instability and survival in gastric cancer: a systematic review and meta-analysis. *Mol. Clin. Oncol.* 2015; **3**; 699–705.
35. Shinozaki-Ushiku A, Kunita A, Fukayama M. Update on Epstein-Barr virus and gastric cancer. *Int. J. Oncol.* 2015; **46**; 1421–1434.
36. Yasui W, Oue N, Aung PP *et al*. Molecular-pathological prognostic factors of gastric cancer: a review. *Gastric Cancer* 2005; **8**; 86–94.
37. Oue N, Naito Y, Hayashi T *et al*. Signal peptidase complex 18, encoded by SEC11A, contributes to progression via TGF-α secretion in gastric cancer. *Oncogene* 2014; **33**; 3918–3926.
38. Mukai S, Oue N, Oshima T *et al*. Overexpression of PCDHB9 promotes peritoneal metastasis and correlates with poor prognosis in patients with gastric cancer. *J. Pathol.* 2017; **243**; 100–110.
39. Polom K, Marano L, Marrelli D *et al*. Meta-analysis of microsatellite instability in relation to clinicopathological characteristics and overall survival in gastric cancer: microsatellite instability in gastric cancer. *Br. J. Surg.* 2018; **105**; 159–167.
40. Topalian SL, Taube JM, Anders RA *et al*. Mechanism-driven biomarkers to guide immune checkpoint blockade in cancer therapy. *Nat. Rev. Cancer* 2016; **16**; 275–287.
41. Wen KW, Rabinovitch PS, Huang D *et al*. Use of DNA flow cytometry in the diagnosis, risk stratification, and management of gastric epithelial dysplasia. *Mod. Pathol.* 2018; **31**; 1578–1587.
42. Fassan M, Simbolo M, Bria E *et al*. High-throughput mutation profiling identifies novel molecular dysregulation in high-grade intraepithelial neoplasia and early gastric cancers. *Gastric Cancer* 2014; **17**; 442–449.
43. Min B-H, Hwang J, Kim NK *et al*. Dysregulated Wnt signalling and recurrent mutations of the tumour suppressor RNF43 in early gastric carcinogenesis. *J. Pathol.* 2016; **240**; 304–314.
44. Hu G, Qin L, Zhang X *et al*. Epigenetic silencing of the MLH1 promoter in relation to the development of gastric cancer and its use as a biomarker for patients with microsatellite instability: a systematic analysis. *Cell. Physiol. Biochem.* 2018; **45**; 148–162.
45. Sun Y, Deng D, You W-C *et al*. Methylation of p16 CpG islands associated with malignant transformation of gastric dysplasia in a population-based study. *Clin. Cancer Res.* 2004; **10**; 5087–5093.

46. Diaz LA, Le DT. PD-1 blockade in tumors with mismatch-repair deficiency. *N. Engl. J. Med.* 2015; **373**; 1979.
47. Le DT, Durham JN, Smith KN *et al*. Mismatch repair deficiency predicts response of solid tumors to PD-1 blockade. *Science* 2017; **357**; 409–413.
48. Wang GG, Yao JC, Worah S *et al*. Comparison of genetic alterations in neuroendocrine tumors: frequent loss of chromosome 18 in ileal carcinoid tumors. *Mod. Pathol.* 2005; **18**; 1079–1087.
49. Kytölä S, Höög A, Nord B *et al*. Comparative genomic hybridization identifies loss of 18q22-qter as an early and specific event in tumorigenesis of midgut carcinoids. *Am. J. Pathol.* 2001; **158**; 1803–1808.
50. Karpathakis A, Dibra H, Pipinikas C *et al*. Prognostic impact of novel molecular subtypes of small intestinal neuroendocrine tumor. *Clin. Cancer Res.* 2016; **22**; 250–258.
51. Wheeler JMD. An insight into the genetic pathway of adenocarcinoma of the small intestine. *Gut* 2002; **50**; 218–223.
52. Bläker H, Helmchen B, Bönisch A *et al*. Mutational activation of the RAS-RAF-MAPK and the wnt pathway in small intestinal adenocarcinomas. *Scand. J. Gastroenterol.* 2004; **39**; 748–753.
53. Arai M, Shimizu S, Imai Y *et al*. Mutations of the Ki-ras, p53 and APC genes in adenocarcinomas of the human small intestine. *Int. J. Cancer* 1997; **70**; 390–395.
54. Murata M, Nagasawa Y, Oda K *et al*. Molecular and biological analysis of carcinoma of the small intestine: β-catenin gene mutation by interstitial deletion involving exon 3 and replication error phenotype. *Am. J. Gastroenterol.* 2000; **95**; 1576–1580.
55. Davison JM, Choudry HA, Pingpank JF *et al*. Clinicopathologic and molecular analysis of disseminated appendiceal mucinous neoplasms: identification of factors predicting survival and proposed criteria for a three-tiered assessment of tumor grade. *Mod. Pathol.* 2014; **27**; 1521–1539.
56. Gleeson EM, Feldman R, Mapow BL *et al*. Appendix-derived pseudomyxoma peritonei (PMP): molecular profiling toward treatment of a rare malignancy. *Am. J. Clin. Oncol.* 2018; **41**; 777–783.
57. Liu X, Mody K, de Abreu FB *et al*. Molecular profiling of appendiceal epithelial tumors using massively parallel sequencing to identify somatic mutations. *Clin. Chem.* 2014; **60**; 1004–1011.
58. Wong NA, Gonzalez D, Salto-Tellez M *et al*. RAS testing of colorectal carcinoma—a guidance document from the Association of Clinical Pathologists Molecular Pathology and Diagnostics Group. *J. Clin. Pathol.* 2014; **67**; 751–757.
59. Capper D, Voigt A, Bozukova G *et al*. BRAF V600E-specific immunohistochemistry for the exclusion of Lynch syndrome in MSI-H colorectal cancer. *Int. J. Cancer* 2013; **133**; 1624–1630.
60. Taieb J, Le Malicot K, Shi Q *et al*. Prognostic value of BRAF and KRAS mutations in MSI and MSS stage III colon cancer. *J. Natl. Cancer Inst.* 2017; **109**; djw272.
61. Dudley JC, Lin M-T, Le DT *et al*. Microsatellite instability as a biomarker for PD-1 blockade. *Clin. Cancer Res.* 2016; **22**; 813–820.
62. Muzny D, Bainbridge M, Chang K *et al*. Comprehensive molecular characterization of human colon and rectal cancer. *Nature* 2012; **487**; 330–337.
63. Prenen H, De Schutter J, Jacobs B *et al*. PIK3CA mutations are not a major determinant of resistance to the epidermal growth factor receptor inhibitor cetuximab in metastatic colorectal cancer. *Clin. Cancer Res.* 2009; **15**; 3184–3188.
64. Sartore-Bianchi A, Martini M, Molinari F *et al*. *PIK3CA* mutations in colorectal cancer are associated with clinical resistance to EGFR-targeted monoclonal antibodies. *Cancer Res.* 2009; **69**; 1851–1857.
65. Pagès F, Mlecnik B, Marliot F *et al*. International validation of the consensus Immunoscore for the classification of colon cancer: a prognostic and accuracy study. *Lancet* 2018; **391**; 2128–2139.
66. Boland P, Ma W. Immunotherapy for colorectal cancer. *Cancers* 2017; **9**; E50.
67. Yamanaka T, Oki E, Yamazaki K *et al*. 12-Gene recurrence score assay stratifies the recurrence risk in stage II/III colon cancer with surgery alone: the SUNRISE Study. *J. Clin. Oncol.* 2016; **34**; 2906–2913.
68. Kennedy RD, Bylesjo M, Kerr P *et al*. Development and independent validation of a prognostic assay for stage II colon cancer using formalin-fixed paraffin-embedded tissue. *J. Clin. Oncol.* 2011; **29**; 4620–4626.
69. de Sanjosé S, Brotons M, Pavón MA. The natural history of human papillomavirus infection. *Best Pract. Res. Clin. Obstet. Gynaecol.* 2018; **47**; 2–13.
70. Wong AK, Chan RC, Aggarwal N *et al*. Human papillomavirus genotypes in anal intraepithelial neoplasia and anal carcinoma as detected in tissue biopsies. *Mod. Pathol.* 2010; **23**; 144–150.
71. Patel HS, Silver AR, Levine T *et al*. Human papillomavirus infection and anal dysplasia in renal transplant recipients. *Br. J. Surg.* 2010; **97**; 1716–1721.
72. Yanik EL, Kaunitz GJ, Cottrell TR *et al*. Association of HIV status with local immune response to anal squamous cell carcinoma: implications for immunotherapy. *JAMA Oncol.* 2017; **3**; 974–978.
73. Ott PA, Piha-Paul SA, Munster P *et al*. Safety and antitumor activity of the anti-PD-1 antibody pembrolizumab in patients with recurrent carcinoma of the anal canal. *Ann. Oncol.* 2017; **28**; 1036–1041.
74. Herfs M, Roncarati P, Koopmansch B *et al*. A dualistic model of primary anal canal adenocarcinoma with distinct cellular origins, etiologies, inflammatory microenvironments and mutational signatures: implications for personalised medicine. *Br. J. Cancer* 2018; **118**; 1302–1312.