Chapter 1.1

Physiology and function of the mouth

Michael P. Escudier
King’s College London and Guy’s and St Thomas’ NHS Foundation Trust London, UK

1.1.1 Physiology

The mouth is an important organ as it is the entry point into the gastrointestinal (GI) tract and damage and disease can compromise dietary intake. Even very minor disorders can have a profound impact on nutritional status.

Anatomy

The oral cavity consists of a number of structures.

The lips surround the mouth and comprise skin externally and a mucous membrane (which has many minor salivary glands) internally, which together with saliva ensure adequate lubrication for the purposes of speech and mastication.

The cheeks make up the sides of the mouth and are similar in structure to the lips with which they are continuous but differ in containing a fat pad in the subcutaneous tissue. On the inner surface of each cheek, opposite the upper second molar tooth, is an elevation that denotes the opening of the parotid duct which leads back to the parotid gland located in front of the ear.

The palate (roof of the mouth) is concave and formed by the hard and soft palate. The hard palate is formed by the horizontal portions of the two palatine bones and the palatine portions of the maxillae (upper jaws). The hard palate is covered by thick mucous membrane that is continuous with that of the gingivae. The soft palate is continuous with the hard palate anteriorly and with the mucous membrane covering the floor of the nasal cavity posteriorly. The soft palate is made up of a fibrous sheet together with the glossopalatine and pharyngopalatine muscles and the uvula hangs freely from its posterior border.

The floor of the mouth can only be seen when the tongue is raised and is formed by the mucosa overlying the mylohyoid muscle. In the midline is the lingual frenum (a fold of mucous membrane), on either side of which is the opening of the submandibular duct from the associated submandibular gland.

The gingivae form a collar around the neck of the teeth and consist of mucous membranes connected by thick fibrous tissue to the periosteum surrounding the bones of the jaw. The gingivae are highly vascular and well innervated.

The teeth are important in mastication and in humans, who are omnivores, they enable both plant and animal tissue to be chewed effectively. Each tooth consists of a crown, which varies in shape dependent on the position in the mouth, and one or more roots. There are eight permanent teeth in each quadrant, consisting of two incisors, a canine, two premolars and three molars, resulting in a total of 32 permanent teeth.

The tongue is a highly mobile, muscular organ in the floor of the mouth which is important in speech, chewing and swallowing. In conjunction with the cheeks, it guides food between the upper and lower teeth until mastication is complete. The taste buds situated on the tongue are responsible for the sensation of taste (salt, bitter, sweet and sour).
**SECTION 1: Physiology and function of the gastrointestinal and hepatobiliary tract**

**Function**

The main role of the mouth is to prepare food for swallowing via the oesophagus and its subsequent passage to the stomach. The first phase of this process is mastication (chewing) which requires activity in the muscles of mastication (masseter, temporalis, medial and lateral pterygoids and buccinator). Chewing helps digestion by reducing food to small particles and mixing it with the saliva secreted by the salivary glands. The saliva lubricates and moistens dry food whilst the movement of the tongue against the hard palate produces a rounded mass (bolus) of food which can be swallowed.

The saliva required for this process is produced by the three paired major salivary glands (parotid, submandibular and sublingual), together with the many minor salivary glands throughout the oropharynx. The total daily production of saliva is around 500 mL, with the rate of production around 0.35 mL/min at rest which increases to 2.0 mL/min during eating and falls to 0.1 mL/min during sleep. The contribution of the various glands varies at rest and during eating (Table 1.1.1).

In addition to its role in digestion and taste, saliva produces a film which coats the teeth and mucosa and helps to cleanse and lubricate the oral cavity. It also prevents dessication of the oral mucosa and acts as a barrier to oral microbiota [1], both physically and through its antimicrobial activity. The buffers within it also help to maintain optimal pH for the action of the salivary amylase and maintain the structure of the teeth.

**Role in digestion**

Very little digestion of food occurs in the oral cavity. However, saliva does contain the enzyme amylase which begins the chemical process of digestion by catalysing the breakdown of starch into sugars.

**Box 1.1.1 Challacombe dry mouth scale**

One point for each feature to a maximum of 10
- Mirror sticks to one buccal mucosa
- Mirror sticks to both buccal mucosa
- Mirror sticks to tongue
- Saliva frothy
- No saliva pooling in floor of mouth
- Tongue shows loss of papillae
- Altered (smooth) gingival architecture
- Glassy appearance to oral mucosa
- Cervical caries (more than two teeth)
- Tongue highly fissured
- Tongue lobulated
- Debris on palate

**Table 1.1.1 Contribution of groups of salivary glands to overall saliva production at rest and during eating**

<table>
<thead>
<tr>
<th></th>
<th>Resting %</th>
<th>Stimulated %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parotid</td>
<td>20</td>
<td>50</td>
</tr>
<tr>
<td>Submandibular</td>
<td>65</td>
<td>49</td>
</tr>
<tr>
<td>Sublingual</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>Minor</td>
<td>7</td>
<td>1</td>
</tr>
</tbody>
</table>

**1.1.2 Measurement and assessment of function**

Salivary function is the most commonly assessed measure of oral function and can be achieved clinically by using the Challacombe dry mouth scale (Box 1.1.1).

A reasonable indication of salivary function may be obtained by measuring the resting (unstimulated) salivary flow over a period of 10 min. In health, the rate will normally be around 0.35 mL/min with a range of 0.2–0.5 mL/min. However, this will be reduced in the presence of xerostomic medications or underlying conditions such as Sjögren’s syndrome and a value below 0.2 mL/min requires further investigation and below 0.1 mL/min is indicative of an underlying condition or disease process. Whilst the stimulated parotid flow rate may also be determined, neither is particularly reliable and hence both should only be viewed as indicative rather than diagnostic.

**1.1.3 Dental disease**

The oral cavity is home to around 500 different microbial species. These bacteria together with saliva and other particles constantly form a sticky, colourless ‘plaque’ on the surface of teeth. Brushing and flossing help to remove this layer which is intimately involved in the development of dental caries and gingivitis. Plaque that is not removed can harden
and form calculus which requires professional cleaning by a dentist or dental hygienist to prevent the development of periodontal disease which can lead to the destruction of the dental support structures and eventually loss of the affected tooth or teeth.

Whilst both dental caries and periodontal disease have been common for many years, non-carious tooth surface loss, particularly in the form of erosion, is a more recent development and is associated with modern lifestyle and dietary intake.

**Dental caries**

Dental caries can occur at any stage throughout life and is one of the most common preventable diseases in childhood [2]. In developed countries there has been a fall in the lifetime experience of dental caries by at least 75% since the 1960s but it still remains a concern in children from low socioeconomic groups and immigrants from outside Western Europe.

The occurrence of decay requires the presence of teeth, oral microbiota, carbohydrate and time. Following a meal, oral microbiota in plaque on the tooth surface ferment carbohydrate to organic acids. This rapid acid production lowers the pH at the enamel surface below the level (the critical pH) at which enamel will dissolve. When the carbohydrate supply is exhausted, the pH within plaque rises, due to the outward diffusion of the acids and their metabolism and neutralisation, and remineralisation of enamel can occur. Dental caries only progresses when demineralisation is greater than remineralisation.

As a result, the risk of dental decay is greatly increased by the intake of fermentable carbohydrate, e.g. sugars, at a frequency which results in the pH remaining below the critical level (the highest pH at which there is a net loss of enamel from the teeth, which is generally accepted to be about 5.5 for enamel). This risk can be negated by the total avoidance of sugar or at least minimised by limiting the frequency of intake, e.g. no between-meals consumption.

**Periodontal disease**

The presence of bacteria on the gingiva causes inflammation (gingivitis), resulting in the gums becoming red and swollen and often bleeding easily. Gingivitis is a mild form of gum disease that can usually be reversed with regular tooth brushing and flossing. This form of gum disease does not include any loss of bone or support tissue.

If gingivitis is not treated, the inflammation can spread and result in the loss of attachment of the gum to the tooth and the development of ‘pockets’ that are colonised by bacteria. The body’s immune system fights these bacteria and as a by-product the body’s natural response and bacterial toxins break down the bone and connective tissue that support the teeth. If this condition remains untreated, the teeth may eventually become mobile and require removal.

While some people are more susceptible than others to periodontal disease, smoking is one of the most significant risk factors and also reduces the chances of successful treatment. Periodontal disease has been reported as a potential risk factor for cardiovascular disease, poorly controlled diabetes and preterm low birth weight [3].

**Non-carious tooth surface loss**

Regular consumption of acidic foods and drinks can reduce the pH below the critical level and the surface layer of enamel is then lost through a combination of erosion, attrition (action of teeth on teeth) and abrasion (by foodstuffs). Over time, the full thickness of the enamel may be lost in this way, leaving exposed dentine which is often associated with sensitivity to temperature changes. This situation may be avoided by limiting the intake of acidic food and drink, e.g. carbonated drinks.

**1.1.4 Oral manifestations of gastrointestinal disease**

Oral manifestations can arise either as a direct presentation of the condition itself or secondary to the effects of the condition or its treatment.

Malabsorption may lead to iron, vitamin B12 or folate deficiency whilst blood loss is most commonly associated with iron deficiency. In all cases, a deficiency state may occur, resulting in anaemia. This can present with depapillation of the tongue (glossitis), a burning sensation affecting the oral mucosa, angular cheilitis or oral ulceration. Correction of the underlying deficiency state will
therefore be associated with their improvement and resolution.

Medical therapy commonly involves the use of corticosteroids or other immunosuppressive medications. Both of these increase the risk of opportunistic infections and hence oral candidosis [4] is frequently seen in the form of angular cheilitis (redness, crusting and splitting of the corners of the mouth), denture stomatitis (erythema of the mucosa in contact with the fit surface of a denture), acute pseudomembranous candidosis or oral soreness/burning affecting the tongue or oral mucosa. Some medications, e.g. methotrexate, may also cause oral ulceration which will only resolve on cessation of the treatment.

In contrast, disease-specific presentations vary and are discussed below.

**Gastro-oesophageal reflux disease**

Due to the high acidity of the gastric contents (pH 1), chronic gastro-oesophageal reflux disease may result in erosion of the teeth [5]. This classically affects the palatal aspect of the upper anterior teeth but may extend further to affect the upper premolar and molar teeth.

**Coeliac disease**

Coeliac disease may present with oral ulceration or dental enamel defects and, less commonly, atrophic glossitis. In addition, whilst the caries indexes are often lower than in unaffected individuals, they may experience delay in tooth eruption [6].

**Crohn’s disease and orofacial granulomatosis**

The precise relationship between Crohn’s disease and orofacial granulomatosis remains unclear [7]. They share many orofacial manifestations including cervical lymphadenopathy, lip swelling, angular cheilitis, mucosal tags, full-thickness gingivitis, submandibular duct ‘staghorning’, fibrous banding and oral ulceration [8].

The oral ulceration seen may arise in relation to an associated deficiency state or medical therapy when it takes a linear form and occurs in the sulci, it is suggestive of underlying GI involvement requiring further investigation [8].

Crohn’s disease may also rarely present with pyostomatitis gangrenosum (chronic ulceration) affecting the tongue or oral mucosa [9].

**Ulcerative colitis**

Oral features of ulcerative colitis are generally secondary to the underlying condition or its treatment. Rarely, pyostomatitis vegetans (a generalised ulceration of the oral mucosa) may be the initial presentation of previously occult ulcerative colitis [10].

**Irritable bowel syndrome (IBS)**

A significant number of patients with IBS also have orofacial pain such as facial arthromyalgia (16%, [11]) or persistent orofacial pain (atypical facial pain, atypical odontalgia) [12]. Conversely, IBS has been shown to be present in many (64%) patients diagnosed with facial arthromyalgia [11].

**References**

1.1 Physiology and function of the mouth


Chapter 1.2

Physiology and function of the oesophagus

Rami Sweis
University College London Hospital, London, UK

The oesophagus co-ordinates the transport of food and fluid from the mouth to the stomach. The oesophagogastric junction (OGJ) is a physiological barrier which reduces reflux of gastric contents. In harmony, these processes limit contact of the swallowed bolus, refluxed acid and other chemicals with oesophageal mucosa. Disruption of function can interrupt bolus delivery or induce gastro-oesophageal reflux. Symptoms produced may range in severity from heartburn and regurgitation to dysphagia and pain.

1.2.1 Anatomy

Oesophagus

The oesophagus is a muscular tube connecting the pharynx to the stomach. The cervical oesophagus extends distally from the cricopharyngeus and the thoracic oesophagus terminates at the hiatal canal before it flares into the gastric fundus. The muscularis propria consists of the outer longitudinal and inner circular muscle layers. The musculature is divided into the proximal striated and mid-distal smooth muscle. This proximal ‘transition zone’ is located one-third of the distance from the pharynx and is the site with the weakest force of peristaltic contractions [1].

Histologically, the oesophageal wall is composed of the mucosa, submucosa and muscularis mucosa. The oesophageal body is lined by non-keratinised stratified squamous epithelium which abruptly joins with the glandular gastric columnar epithelium at the squamocolumnar junction. This can be the site of mucosal change associated oesophagitis and Barrett’s oesophagus.

The antireflux barrier

The OGJ is not a clearly identifiable sphincter but its sphincter-like properties can be defined functionally as a high-pressure zone between the stomach and oesophagus. Sphincter competence is dependent on the integrity and overlap of the intrinsic lower oesophageal sphincter (LOS) and diaphragmatic crura. A separation, hiatus hernia, is associated with disruption of LOS integrity, loss of the intra-abdominal LOS segment and an increased susceptibility to gastro-oesophageal reflux.

1.2.2 Physiology and function

Voluntary swallowing initiates with ‘deglutitive inhibition’ of the smooth muscle oesophagus and LOS. This reflex relaxation is nitric oxide mediated and permits passage of the bolus with minimal resistance. The subsequent excitatory, predominantly cholinergic, activity produces a progressive wave of smooth muscle excitation. A co-ordinated peristalsis clears the bolus from the oesophagus.

The LOS exhibits a continuous resting (basal) tone which relaxes on stimulation of the intramural nerves such as during deglutitive inhibition (swallowing). Disruption of this physiological process may impact on bolus transport and induce symptoms
1.2 Physiology and function of the oesophagus

A representative normal swallow using high-resolution manometry is presented in Figure 1.2.1.

Spontaneous LOS relaxations normally occur as a response to gastric postprandial distension and bloating: ‘transient lower oesophageal sphincter relaxation’ (TLOSR). LOS relaxation can also follow peristaltic activity: ‘swallow-induced lower oesophageal sphincter relaxation’ (SLOSR). Gastro-oesophageal reflux and belch occur when there is equalisation of pressure between the stomach and oesophagus (common cavity) (Figure 1.2.2). Patients with gastro-oesophageal reflux disease (GORD) do not have an increased frequency of TLOSRs; rather, the tendency of reflux to occur during these events is greater [2]. The effectiveness of oesophageal clearance of refluxed material is an important contributor to the severity of GORD [3–5]. Other determinants of GORD include the presence and size of a hiatus hernia, increasing age and obesity as well as the calorie and fat content of the diet [6,7].

Measurement and assessment of function

In the absence of disease on endoscopy and failure to respond to empirical therapy, guidelines recommend manometry and ambulatory reflux testing [8,9]. Recent advances in technology provide better insight into the assessment of oesophageal function and disease.

Box 1.2.1 Co-ordinated peristaltic activity

Co-ordinated peristaltic activity is a multistep process which usually requires:

• a pharyngeal ‘pump’ – to push food and fluid through the oesophagus
• gravity – whereby bolus weight contributes to its aboral progress
• appropriate relaxation and opening of the oesophagogastric junction
• effective oesophageal motor function – deglutitive inhibition followed by co-ordinated peristaltic contraction
• a positive oesophagogastric pressure drop.

Manometry

Peristalsis and OJG activity can be measured with manometry. Conventional manometry (4–8 sensors) measures the circumferential contraction, pressure wave duration and peristaltic velocity of single water swallows. High-resolution manometry (HRM; 21–36 sensors) is an advance on conventional systems as it provides a compact, spatiotemporal representation of oesophageal pressure activity. In addition, it can measure the forces that drive movement of food and fluid through the oesophagus and OJG [10]. An uninterrupted well-co-ordinated peristalsis defines oesophageal motility while the presence of a positive pressure gradient in the absence of obstruction describes whether this motility is effective and likely to clear the bolus [11] (see Figure 1.2.1). Thus HRM improves diagnostic sensitivity to peristaltic dysfunction as symptoms and mucosal damage are more likely to occur as a result of disturbed bolus transport and poor clearance [5]. Furthermore, recent advances in methodology have shown how HRM can also facilitate the assessment of swallowing behaviour (eating and drinking) when symptoms are more likely to be triggered [5,12,13] (Box 1.2.2).

Ambulatory reflux studies

Gastro-oesophageal reflux disease (GORD) occurs when gastric contents pass into the oesophagus at an increased frequency, are not effectively cleared or are perceived in an exaggerated manner. This can lead to mucosal damage and/or symptoms with varying degrees of severity. Presenting symptoms alone are an unreliable guide to identifying oesophageal dysfunction [14,15]. Objective testing is required to avoid inappropriate medical and surgical therapy. Ambulatory pH monitoring provides an assessment of oesophageal acid exposure and symptoms. Standard testing is performed using a 24-hour nasopharyngeal pH catheter (with or without impedance, see next section). Diagnosis is made based on measurements of oesophageal acid exposure (e.g. total number of reflux events and percent time reflux events cause a pH drop below a threshold of 4) as well as the association of reflux events with symptoms. Measurements can
be further subdivided into upright and supine. However, intolerance to the nasal catheter can influence the result.

Multiple intraluminal impedance with pH monitoring (MII-pH)

Oesophageal symptoms are often related to disturbed bolus transport rather than acid reflux [16]. Also symptoms may persist despite effective acid suppression as acid-reducing medications do not influence the frequency or volume of non-acid reflux episodes [17,18]. Multiple intraluminal impedance (MII) can determine the direction of bolus movement, the success or failure of bolus transit and the proximal extent of the refluxate. Furthermore, it can discriminate between liquid and gas reflux. When combined with a pH sensor (MII-pH), it can differentiate between acid (pH <4), weakly acid (pH 4–7) or weakly alkaline (pH >7) reflux [19]. Therefore, MII-pH is considered to be more sensitive than standard pH testing, with up
Transient lower oesophageal sphincter relaxation followed shortly afterwards by a common cavity during which there is equalisation of pressure between the stomach and oesophagus when reflux is most likely to occur. The event is terminated and the oesophagus is cleared of refluxed contents with the arrival of a well-coordinated primary peristalsis. Oesophageal and lower oesophageal sphincter pressures return to baseline levels following completion of peristalsis. TLOSR, transient lower oesophageal sphincter relaxation.

**Box 1.2.2 Hierarchical analysis of high-resolution manometry**

Hierarchical analysis of high-resolution manometry studies according to the Chicago Classification whereby pathology in the OGJ is considered first. Major motility disorders (achalasia, absent peristalsis, diffuse oesophageal spasm and extreme hypertensive disorders) are never found in healthy individuals, are commonly associated with impaired bolus transport and, in turn, often lead to symptoms. The significance of peristalsis abnormalities described in ‘Other motility disorders’ is not clear as these can also be found in asymptomatic individuals [20].

**I. OGJ obstruction**

**Achalasia**
- Classic (non-relaxing LOS + aperistalsis + dilated oesophagus)
- Compression (non-relaxing LOS + aperistalsis + oesophageal pressurisation)
- Vigorous (non-relaxing LOS + oesophageal spasm)

**Other obstruction**
- Eosinophilic oesophagitis
- Benign or malignant stricture
- Post surgery (e.g. antireflux procedure)

**II. Major motility disorder**
- Absent peristalsis
- Diffuse spasm
- Jackhammer oesophagus (nutcracker with extreme pressures)

**III. Other motility disorders**
- Weak peristalsis
- Frequent failed peristalsis
- Hypertensive peristalsis
- Rapid contractility
to 20% improvement in diagnostic yield [21]. Indications for its use are the same as for standard ambulatory pH studies. In those with established GORD but ongoing symptoms despite optimal medical therapy, MII-pH can be performed while on acid reducing medication in order to identify if (non-acid) reflux is the culprit or to exclude breakthrough acid reflux. In addition, in the assessment of atypical disease (e.g. laryngopharyngeal reflux, aerophagia, supragastric belching, cough).

Wireless pH monitoring (Bravo®)

Wireless pH monitoring (Bravo®, Given Imaging) is an endoscopically placed, catheter-free, ambulatory pH monitoring system (Figure 1.2.3). Bravo® is a viable option for those who are intolerant to the nasal catheter [6]. It can measure for prolonged periods (at least 48 h) [22,23] and is especially suitable for patients with intermittent symptoms [22,24] or those with persistent typical symptoms whose catheter-based study was inconclusive [25]. However, Bravo® cannot discriminate between liquid and gas reflux nor can it differentiate between acid and nonacid reflux.

1.2.3 Pathology

Motility

An important advance of the modern HRM-based classification (the Chicago Classification) [26–28] is that it is hierarchical; the OGJ is considered first because pathology within the OGJ will influence oesophageal function above [20]
1.2 Physiology and function of the oesophagus

(see Box 1.2.2). In addition, the Chicago Classification makes a clear distinction between dysmotility that is ‘never seen in normal individuals’ (Major motility disorders) and that which may be merely ‘outside the normal range’. In the former, treatment is usually directed at correcting the underlying pathology whereas in the latter, therapy often targets symptoms [29,30].

Achalasia, a ‘Major motility disorder’, is characterised by a non-relaxing LOS and the absence of oesophageal peristalsis. The Chicago Classification further categorises achalasia into three subtypes, each with its own response to medical (pneumatic dilation and botulinum toxin) and surgical (Heller myotomy) therapy [31,32] (see Box 1.2.2). Left untreated, the compression subtype (an HRM diagnosis) is thought to ‘decompensate’ and lead to classic achalasia. Furthermore, this compression subtype has the best response to all forms of therapy (botulinum toxin, dilatation, myotomy) classic achalasia [33,34]. On the other hand, many hypertensive oesophageal disorders can also be found in asymptomatic individuals and have shown varying degrees of success with therapy. Nitrates, calcium channel blockers and sildenafil can influence function in some but often tricyclic antidepressants and selective serotonin receptor inhibitors are required to target symptoms [35,36].

Gastro-oesophageal reflux disease

Gastro-oesophageal reflux disease is subclassified into erosive oesophagitis, endoscopy-negative reflux disease (positive oesophageal acid exposure and/or reflux-symptom association with normal endoscopy) and functional heartburn (negative oesophageal acid exposure, negative reflux-symptom association, poor response to acid-reducing medication with normal endoscopy but ongoing symptoms) [37,38]. Differentiating between erosive oesophagitis, endoscopy-negative reflux disease and functional heartburn is essential to target appropriate therapy and oesophageal physiology studies are required to secure a diagnosis. In addition, an assessment of GORD should also be sought in patients presenting with dysphagia as oesophageal dysfunction can be exacerbated by or be a consequence of reflux disease.

1.2.4 Conclusion

In conclusion, GORD and dysphagia are common in the community and can be associated with significant morbidity and reduced quality of life. Furthermore, chronic reflux is related to the rising incidence of oesophageal adenocarcinoma, especially in those with Barrett’s oesophagus [39]. Such concerns emphasize the importance of appropriate and early investigation and management. In the absence of disease on endoscopy and failure to respond to empirical therapy, guidelines recommend manometry and ambulatory reflux testing. Advances in technology and methodology have revolutionised the way the oesophagus is investigated and provide a more ‘realistic’ assessment of function which can help guide therapy.

References

Chapter 1.3

Physiology and function of the stomach

Luca Marciani and Mark Fox
University of Nottingham, Nottingham, UK

1.3.1 Physiology, anatomy and function

The human stomach is a J-shaped organ of the gastrointestinal (GI) tract, located between the oesophagus and the duodenum, and it has a key role in digestion and absorption. The main anatomical regions are shown in Figure 1.3.1. The stomach’s main functions are to store and break down food and deliver digesta to the small intestine.

The stomach receives boluses of food via the lower oesophageal sphincter. It is able to reduce gastric wall tone via a vagally mediated reflex (‘accommodation’) which allows the reservoir to expand and accommodate increasing amounts of food without important increases in intragastric pressure [1]. In addition to ‘receptive’ accommodation mediated by mechanoreceptors in the gastric wall, once nutrients pass into the small intestine the gastric response is modulated by chemoreceptors and osmoreceptors to ensure that gastric emptying through the pylorus is controlled and optimized for efficient digestion [1,2].

During intragastric food processing, the stomach secretes hydrochloric acid, lipase and pepsin. This process is regulated by the central and enteric nervous system and neuroendocrine cell networks [3]. These secretions together with salivary enzymes active within the bolus start the chemical breakdown of food. At the same time, highly co-ordinated antropyloroduodenal contractions effect mechanical breakdown (trituration) of solid food. Gastric emptying is ultimately the result of these co-ordinated actions, controlled opening of the pylorus and antroduodenal differences in pressure which drive gastric emptying [4,5]. Liquids empty faster than solids, which are first triturated to small particles, usually less than 3 mm in size, to promote chemical digestion and absorption after delivery to the duodenum and small intestine [6]. Other physical factors such as meal viscosity, the density and breaking strength of food particles also affect the rate of gastric emptying [6–8].

![Figure 1.3.1 Schematic diagram of the human stomach.](image-url)
1.3.2 Measurement and assessment of gastric function

Measurement of gastric function has improved understanding of the physiological response to food in health and disease and in response to dietary or pharmacological intervention. A number of tests are available and are briefly described in the following sections [9].

Gastric accommodation and sensation

Gastric accommodation can be evaluated using the barostat test. This involves intubating the subject orally using a double-lumen catheter with a plastic bag on the tip. The balloon is commonly placed in the proximal stomach. An electronic barostat device is then used to control expansions of the bag to assess, for example, volume expansion during pressure-guided distension or after delivery of a test meal [10]. This is the ‘standard test’ of gastric accommodation though availability is limited, the method is invasive and the presence of a balloon in the stomach affects gastric relaxation. Gastric sensation elicited by barostat distension paradigms leads to brain cortical activations that can be assessed using functional brain magnetic resonance imaging (MRI) and positron emission tomography (PET) methods [11,12].

A simple and inexpensive alternative to the barostat is the drink test [13]. This involves ingesting water or a nutrient drink at a given rate until the maximum tolerated volume is reached. Subjective scores of sensation are collected during and after the test. The results are not easy to interpret due to variation in gastric capacity and the merits of this test are debated.

Conventional ultrasound has been used to measure the area of the proximal stomach after a meal in a sagittal section and the maximal diameter in an oblique frontal section [14]. Three-dimensional reconstruction of ultrasound images integrates this information and gives volume measurements; however, the technique is user dependent and can be used only with liquid meals.

The distribution of gastric contents within the stomach on scintigraphy provides some impression of gastric accommodation [15]. Another nuclear medicine test that can measure change in gastric volumes is single photon emission computed tomography (SPECT). This method involves injecting intravenously a $^{99m}$Tc-labelled compound which is taken up in the mucosa. A dual-headed gamma camera is used to measure the radiation emitted and reconstruct axial images of the stomach. A three-dimensional image can be reconstructed later; however, the temporal and spatial resolution are limited compared to MRI.

Magnetic resonance imaging is an emerging technique used to assess fasting and postprandial gastric volumes [16] due to the lack of ionising radiation, multiplanar imaging, speed and excellent contrast between different organs and intragastric meal components. It has been used to evaluate the effects of the barostat balloon in the stomach [17], finding that the bag increased postprandial gastric volumes. Cross-sections of the fundus [18] and maximum antral diameters following model meals [7] have also been measured using MRI and changes in these variables correlate closely with sensation of fullness and other symptoms in health and disease [8,19].

Gastric contractility

Antroduodenal motility can be measured using intraluminal manometry by passing a catheter nasogastrically through the pylorus and into the proximal duodenum. The catheter has a varying number of water-perfused or solid-state sensors. These detect the periodical stomach wall contractions and the pressure amplitude profiles with time can be displayed and analysed [20].

The high-resolution and high-speed capabilities of MRI allow imaging of the stomach serially at intervals of a few seconds. These images can be played as motility ‘movies’ and subsequently post-processed to measure motility in terms of antral contractions, frequency, speed and percentage occlusion [21–24]. An interesting finding from MRI studies is the lack of correlation between meal volumes and antral contractility that suggests these contractions are highly stereotyped after a meal and...
1.3 Physiology and function of the stomach

Gastric emptying

Gastric emptying can be measured by labelling test meals with $^{13}$C stable isotopes such as octanoic acid. The label is absorbed in the small intestine during digestion, metabolised to $^{13}$CO$_2$ and then expelled with the breath. As such, serial breath samples are taken at baseline and postprandially to calculate the increase of $^{13}$CO$_2$ with time, which is then assumed to be proportional to gastric emptying [26]. This is an advance on the oral paracetamol absorption under the assumption that the appearance in the blood is directly related to gastric emptying [27].

Using imaging, the simple radiopaque marker test involves the subject ingesting a number (about 20) of small radiopaque pellets with a test meal and following their emptying with fluoroscopy [28]. Results depend on the size and density of the pellets and test meal composition.

Gastric scintigraphy involves the patient eating a radiolabelled meal and measuring the gamma radiation emitted from the ‘region’ of the stomach using a gamma camera. This is carried out at various time points to measure the postprandial gastric emptying curve. The normal range of results depends on the test meal, though simplified protocols have been reported [29] and standardised scrambled egg substitute test meals have been validated in multicentre studies [30]. It is a widely used test and so far considered the ‘gold standard’ although it involves a radiation dose to the subject and results correlate only poorly with patient symptoms [31].

Wireless capsule pills that can measure pH, pressure and temperature have recently appeared on the market. Subjects swallow the pills with a test meal and a receiver worn on the belt records data continuously. The time at which the pill detects a step change up in pH is taken as the time at which the pill is emptied from the stomach [32]. However, given their large size and indigestibility, the emptying of a pill from the stomach is due to strong phase III contractions and not the fed pattern of meal emptying, making interpretation of the data difficult.

A different approach that uses pills to measure gastric emptying is based on magnetically marked solid pills that are ingested by the subjects with a meal and their spatial location monitored over time using non-invasive magnetic source imaging methods [33]. This method is elegant, but requires the use of superconducting quantum interference device (SQUID) magnetometers and has limited applications, mostly to monitor the dissolution of dosage forms for pharmaceutical use.

As described, ultrasound, SPECT and MRI can all measure cross-sections or entire volumes of the stomach. As such, they have all been employed to measure gastric emptying. MRI in particular can measure serially intragastric gas and meal volumes from which one can assess the gastric emptying curves [34,35]. Of particular interest is MRI’s ability to observe the intragastric fate of many food materials and their mixing and dilution [8,36–39].

1.3.3 Pathology

Reflux

Gastro-oesophageal reflux disease (GORD) is a very common disorder caused by the return of gastric contents (‘reflux’) back to the oesophagus, causing inflammation (e.g. oesophagitis) or symptoms (e.g. heartburn, acid regurgitation). Changes in gastric structure have been reported in patients with GORD that compromise the putative ‘flap-valve’ mechanism of the gastro-oesophageal reflux barrier. Additionally, delayed gastric emptying is common in patients with severe disease, prolonging the period after the meal during which reflux can occur.

Disorders of gastric emptying (gastroparesis)

Gastroparesis is a condition in which gastric emptying is delayed. It is classically found in diabetic patients but can be linked to connective tissue diseases, related to previous gastric surgery or have no clear cause (idiopathic). In diabetes, abnormal gastric emptying impairs glucose control and intake.
and digestion of nutrients and medications. Symptoms include prolonged fullness, nausea and vomiting after meals; however, a clear link between delayed emptying and symptoms is observed only in very severe cases. Rather, typical symptoms are associated more closely with impaired gastric accommodation and psychosocial factors as seen in functional dyspepsia.

Rapid gastric emptying can cause symptoms due to ‘dumping’ of nutrients into the small intestine which leads to a powerful neurohormonal response that can cause nausea but also faintness and other symptoms related to insulin-induced hypoglycaemia. In addition, rapid emptying can impair digestion and tolerance of certain nutrients (e.g. fat).

**Functional dyspepsia**

Functional dyspepsia is thought to be a heterogeneous condition characterised by specific gastric motor and sensory abnormalities. Symptoms include fullness, nausea, bloating and epigastric pain. Impaired gastric accommodation is linked to early satiety and weight loss, delayed gastric emptying to prolonged fullness and nausea, and visceral hypersensitivity to epigastric pain. It may be that breakdown of the dynamic, neurohormonal and functional response to food underlies all these abnormalities.

**Rumination**

Rumination is a behavioural disorder in which, responding to dyspeptic or reflux symptoms, patients subconsciously contract their abdominal muscles, forcing gastric contents back to the mouth repeatedly after meals. At this point, the patient often swallows the food again (hence ‘rumination’) or spits out the food, which can lead to undernutrition. This condition is often mistaken for vomiting or reflux disease; however, it does not respond to antiemetics or antacid medication and requires behavioural therapy.

**Cyclic vomiting**

Cyclic vomiting syndrome is a rare condition characterised by paroxysmal bouts of severe nausea and vomiting lasting several days separated by periods of normal health. It may be triggered by cannabis use; however, most cases are idiopathic and are thought to be linked to autonomic nerve dysfunction.

**Acute gastroenteritis**

Gastric infection is unusual except for *Helicobacter pylori* (see next section). However, ingestion of contaminated food can cause nausea and vomiting either directly due to toxins or indirectly due to infection and dysfunction of the small or large intestine.

**Helicobacter pylori**

*Helicobacter pylori*, a spiral-shaped bacterium located in the mucous layer of the stomach, may inhibit or promote acid secretion and causes different diseases depending on how the infection affects the stomach. Distal (antral) gastritis increases the production of gastric acid and increases the risk of duodenal ulceration. Conversely, generalised atrophic gastritis decreases the production of gastric acid with an increased risk of gastric cancer.

**Gastric cancer**

Gastric cancer usually arises in the glandular epithelium (‘adenocarcinoma’) although rare cancers of the smooth muscle (‘leiomyosarcoma’) and immune cells (‘lymphoma’) can also occur. The risk of adenocarcinoma is increased by smoking, alcohol abuse, certain factors in the diet (e.g. nitrates derived from preservatives) and, most importantly, atrophic gastritis induced by *Helicobacter pylori* infection. These cancers usually present in an advanced stage due to obstruction of food passage through the stomach with pain and vomiting or progressive anaemia. Treatment options are often limited and less than one in five patients survives more than 5 years.

**References**


Chapter 1.4

Physiology and function of the small intestine

Paul A. Blaker and Peter Irving
Guy’s and St Thomas’ NHS Foundation Trust, London, UK

The main functions of the small intestine are to complete the digestion of food through co-ordinated motility and secretion and to facilitate the absorption of water, electrolytes and nutrients. Approximately 9 L of fluid derived from oral intake (1.5 L) and exocrine secretions (7.5 L) enter the small intestine each day. Ninety per cent of this is reabsorbed in the small intestine with a further 8% absorbed in the colon. As such, only 100–150 mL of fluid is lost in faeces each day. The average length of the small intestine is 6.9 m but structural adaptations including mucosal folds, villi and microvilli mean that its surface area is 200–500 m². The first 100 cm of the small intestine are highly adapted to the absorption of nutrients, whereas the more distal portions are involved in reclaiming fluid and electrolytes. The small intestine is able to absorb far in excess of the body’s requirements and as such, large portions of this organ can be removed without deleterious effects. However, changes in absorption and secretion homeostasis can rapidly lead to diarrhoea, dehydration, electrolyte disturbance and malnutrition.

1.4.1 Anatomy and histology

The small intestine includes three substructures termed the duodenum, jejunum and ileum, which extend sequentially from the gastric pylorus to the ileocaecal valve. The wall comprises an outer serous coat (tunica serosa), a layer of smooth muscle fibres (muscularis externa), submucosa consisting of dense connective tissue, a thin layer of smooth muscle (muscularis mucosa) and a mucosal layer (tunica mucosa) covered by epithelial cells (Figure 1.4.1). The tunica mucosa is thrown into numerous subfolds, creating the intestinal villi, which contain a dense blood capillary and lymphatic network that supplies the epithelial cells. Enterocytes are the most abundant epithelial cells (80%) and are characterised by the presence of enterocytic microvilli (brush border) that further increases the small intestinal surface area. Goblet cells are interspersed between enterocytes and secrete mucus that acts as a protective coat and lubricant. Tubular intestinal glands are found at the base of the villi (crypts of Lieberkuhn), which contain cells that differentiate into enterocytes, goblet cells, endocrine, paracrine and immune cells (Paneth cells). Changes in the cellular structure between sections of the small intestine allow for functional subspecialisation (Table 1.4.1).

Duodenum

The duodenum is approximately 25–35 cm in length and is split into four parts. It starts as the duodenal bulb, which arises from the gastric pylorus, and ends at the ligament of Treitz, where it joins the jejunum at the duodenojejunal flexure. The common bile duct enters the small intestine in the second part of the duodenum via the ampulla of Vater.
**Figure 1.4.1** Structure of the small intestine.

**Table 1.4.1** Differences in the ultrastructure and function of the small intestine

<table>
<thead>
<tr>
<th>Layer</th>
<th>Duodenum</th>
<th>Jejunum</th>
<th>Ileum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Serosa</td>
<td>No change</td>
<td>No change</td>
<td>No change</td>
</tr>
<tr>
<td>Muscularis externa</td>
<td>Longitudinal and circular smooth muscle supplied by Auerbach’s plexus</td>
<td>Similar to duodenum</td>
<td>Similar to duodenum</td>
</tr>
<tr>
<td>Submucosa</td>
<td>Brunner’s glands +++ Meissner’s plexus</td>
<td>Brunner’s glands +</td>
<td>Brunner’s glands +</td>
</tr>
<tr>
<td>Muscularis mucosae</td>
<td>No change</td>
<td>No change</td>
<td>No change</td>
</tr>
<tr>
<td>Lamina propria</td>
<td>No Peyer’s patches</td>
<td>No Peyer’s patches</td>
<td>Peyer’s patches +++</td>
</tr>
<tr>
<td>Intestinal epithelium</td>
<td>Simple columnar Goblet cells Endocrine cells Paracrine cells Paneth cells</td>
<td>Villi longer than duodenum</td>
<td>Villi shorter than duodenum</td>
</tr>
<tr>
<td>Sodium content</td>
<td>145 mmol/L</td>
<td>Iron and folate absorption</td>
<td>125 mmol/L Vitamin B12 and bile salt absorption in terminal ileum</td>
</tr>
<tr>
<td>Specialised functions</td>
<td>Iron and folate absorption in proximal jejunum</td>
<td>Absorption of vitamin B1 and B2</td>
<td>Absorption of vitamin C</td>
</tr>
</tbody>
</table>

Section 1: Physiology and function of the gastrointestinal and hepatobiliary tract
1.4 Physiology and function of the small intestine

The duodenum is distinguished from other parts of the small intestine by the presence of numerous Brunner’s glands which secrete urogastrone (human epidermal growth factor), which is required for epithelial cell proliferation [1]. Consequently, the tips of the villi are continuously shed into the lumen and replaced by new cells from the crypts of Lieberkuhn. As such, the entire small intestine epithelium is renewed every 2–6 days.

Jejunum and ileum

The jejunum is approximately 2.5 m in length, whereas the length of the ileum is more variable (average 2–4 m). Both are contained within the peritoneum and are suspended by a mesentery. Most of the jejunum lies in the left upper quadrant of the abdomen, whereas the ileum mainly occupies the right lower quadrant. The jejunal folds are larger than those found in the duodenum or ileum.

1.4.2 Physiology and function

The gastric antrum sieves liquid chyme through the remaining solid matter in the stomach and delivers a continuous slow rate of gastric contents into the duodenum. The presence of chyme in the small intestine leads to the release of the hormones cholecystokinin (CCK) and secretin, which stimulate secretion of bicarbonate and pancreatic enzymes, and cause contraction of the gallbladder, which releases bile [2–4]. Proteins and peptides are degraded into amino acids through the action of pancreatic trypsin, chymotrypsin and elastase and subsequently by enzymes on the brush border. Lipids are degraded into fatty acids and glycerol and following emulsification by bile salts, triglycerides are split into free fatty acids and monoglycerides by pancreatic lipase. Carbohydrates may be broken down by pancreatic amylase into oligosaccharides or may pass into the colon where they are metabolised by GI microbiota. Brush border enzymes including dextrinase, glycoamylase, maltase, sucrase and lactase further break down oligosaccharides into monosaccharides prior to absorption. It is estimated that up to 65% of the adult population demonstrate a deficiency in lactase activity.

Reflex peristaltic waves mediated by musculo-motor neurones propel the small intestinal contents at a rate of 1–2 cm/min, meaning that it takes an average of 2–6 h to reach the colon [5]. The intensity of the muscular contractions is influenced by the nature of the ingested food. Solid foods induce greater activity than liquid meals, and those that are high in glucose cause greater stimulation than ones high in fat.

Several mechanisms are involved in the absorption of nutrients by enterocytes, including passive diffusion, cytosis, active transfer and carrier-mediated transport [6]. Uptake of water is driven by the absorption of sodium (Na\(^+\)), potassium (K\(^+\)) and organic compounds and occurs through the formation of osmotic gradients. The absorption of Na\(^+\) is mediated by several different mechanisms including specific transmembrane carrier proteins.

1.4.3 Investigation of the small intestine

Correct diagnosis and management of small intestinal pathology are dependent on accurate history taking, clinical examination and specialist investigations. Non-bloody liquid stools greater than 1.5 L a day strongly suggest disease of the small intestine and weight loss may signify malabsorption. Here we summarise the key small intestinal investigations and describe their relevance to pathology.

Blood tests

Anaemia is detected on a full blood count (FBC). Iron deficiency anaemia is characterised by red cell microcytosis (low mean corpuscular volume, MCV), low serum ferritin and iron, low transferrin saturation and a high total iron-binding capacity. The most common causes of iron deficiency anaemia are a lack of dietary iron, gastrointestinal bleeding or proximal small intestinal pathology. Low serum folate may suggest disease in the duodenum or proximal jejunum, and is associated with a macrocytic anaemia (high MCV). Vitamin B12 deficiency also causes macrocytic anaemia and may be due to inadequate intake, autoimmune destruction of gastric parietal cells or antibodies to intrinsic factor (pernicious anaemia), pancreatic exocrine deficiency or disease in the terminal ileum.
Albumin is a protein synthesized in the liver that can be measured in blood tests. Hypoalbuminaemia may be due to a number of different causes but when associated with a history consistent with small intestinal disease, it may suggest protein-losing enteropathy or small intestinal inflammation [7].

Coeliac serology forms part of a screen for small intestinal pathology given the high prevalence of coeliac disease in northern Europeans (1 in 300–500). Antitissue transglutaminase (anti-tTG) is the best test, and is also useful to monitor response to gluten withdrawal.

Endoscopy

Upper gastrointestinal endoscopy (OGD) detects mucosal abnormalities in the oesophagus, stomach and duodenum. It also allows biopsies to be taken and is therefore the gold standard test in coeliac disease. Furthermore, OGD offers potential for therapeutic intervention in the management of upper gastrointestinal bleeding. Enteroscopes are longer than standard gastroscopes and allow deeper intubation of the small intestine.

Wireless capsule endoscopy involves swallowing a pill containing a small camera which transmits images to a receiver as it passes through the small intestine, allowing them to be viewed at a later date. This technique is particularly useful in the diagnosis of small intestinal vascular lesions and sites of inflammation that cannot be reached with conventional endoscopy [8] but is contraindicated in patients with small intestinal strictures and does not permit biopsies to be taken.

Radiology

A plain abdominal X-ray is useful in the diagnosis of intestinal obstruction. However, to provide intraluminal or mucosal detail, either enterography or enteroclysis is needed. Enterography, or small intestine follow-through, involves the ingestion of barium with X-rays being taken as it moves through the small intestine, thus potentially demonstrating small intestinal dilation, mucosal thickening, strictures, fistulae and tumours. Enhanced mucosal detail may be obtained through enteroclysis in which barium is rapidly infused through a nasoduodenal or nasojejunal tube. Computed tomography (CT) allows for better examination of intra- and extraluminal structures than barium X-rays but also involves ionising radiation exposure. Small intestinal magnetic resonance imaging (MRI) is an alternative to CT and, because it does not involve radiation exposure, has an increasingly important role to play in patients who require repeated imaging such as those with small intestinal Crohn’s disease.

Breath tests

Breath tests are used to identify incomplete absorption of sugars such as lactose and fructose by the small intestine. If incompletely absorbed, the sugars will be fermented by colonic microbiota, resulting in the production of hydrogen which can be detected in exhaled breath. The test results may aid dietary advice on carbohydrate restriction.

Bacterial overgrowth in the proximal small intestine can be detected by using lactulose or glucose as the test sugar. The concept of distal small intestinal bacterial overgrowth is somewhat more controversial and is probably overdiagnosed by lactulose breath tests, the majority of positive results probably being explained by rapid transit of lactulose through the small intestine to the caecum [9].

Tests for malabsorption

Several other tests are sometimes used for the investigation of malabsorption, including xylose absorption (through detection of urinary xylose concentrations 5 h after oral ingestion), measurement of faecal alpha-1 antitrypsin (as a marker of protein-losing enteropathy) and measurement of faecal fat over 3 days following an orally administered fat load of 70 g to detect fat malabsorption. The latter is sometimes combined with measurement of faecal elastase as a marker of exocrine pancreatic function, with low enzyme concentrations denoting pancreatic insufficiency.

Bile malabsorption study

Normally bile acids are produced in the liver, stored in the gallbladder and released into the duodenum
1.4 Physiology and function of the small intestine

in response to a meal; 90% of bile acids are reabsorbed in the terminal ileum and circulated via the portal vein back to the liver. The presence of excess bile acids in the colon can result in diarrhoea. Malabsorption of bile salts may occur in patients with terminal ileal disease or following its resorption. In such cases, diarrhoea may respond to treatment with bile salt sequestrants such as cholestyramine which chelates bile salts/acids \cite{10,11}. In patients with short ileal resections (<1 m), the liver is able to synthesize sufficient replacement bile acids to maintain normal fat absorption. However, in long ileal resections (>1 m), it is unable to do so, resulting in fat malabsorption and, consequently, steatorrhoea \cite{12}. Such patients will usually respond to dietary fat restriction. To maintain energy requirements, dietary fats may need to be substituted with medium-chain triglyceride (MCT) oil, since MCTs are absorbed directly into the portal vein without the need for bile acids/salts. Bile acid malabsorption can be detected using a radiolabelled synthetic bile acid scan (SeHCAT).

1.4.4 Pathology

**Villous atrophy**

Malabsorption occurs when there is a failure to absorb nutrients from the GI tract. It is either generalised or specific to a particular molecule (e.g. lactose), which may have associated clinical consequences. For example, fat malabsorption leads not only to steatorrhoea but also to malabsorption of the fat-soluble vitamins A, D, E and K. The causes of malabsorption are myriad; the major ones are listed in Table 1.4.2. The symptoms of malabsorption are variable but diarrhoea, undernutrition and fatigue are common (Table 1.4.3).

In Western countries, the most common cause of villous atrophy is coeliac disease, in which small intestinal inflammation occurs in response to the ingestion of gluten. The disease usually affects the duodenum and jejunum and is characterised by loss of the normal finger-like villi (villous atrophy), which decreases the surface area available for absorption. Coeliac serology and duodenal biopsy are diagnostic and the majority of patients respond to a gluten-free diet. Infections such as *Tropheryma whippelii* (Whipple’s disease), tropical sprue and giardia may also lead to villous atrophy.

**Inflammation**

Inflammation affecting the small intestine may be either acute or chronic. Acute inflammation is often related to infection or medications (for example, non-steroidal inflammatory drugs).

Small intestinal Crohn’s is a cause of chronic inflammation. Crohn’s disease most commonly affects the terminal ileum and right colon whereas isolated duodenal or jejunal disease is rare \cite{13}. Crohn’s disease affecting the proximal small intestine can present with weight loss, iron deficiency anaemia and protein-losing enteropathy, whereas distal small intestinal inflammation presents with the more typical symptoms associated with Crohn’s disease, including abdominal pain and diarrhoea \cite{14}. Over time, complications such as strictures, penetration of inflammation through the intestinal wall or perforation may occur and 50–60% of patients will require surgery for Crohn’s disease within 5 years of diagnosis \cite{13,15}.

**Infection**

There are several different mechanisms by which bacterial and viral infections may interfere with the normal absorptive and secretory functions of the small intestine to cause diarrhoea and malabsorption. For example, the enterotoxin of *Vibrio cholerae* stimulates chloride secretion, leading to secretory diarrhoea and copious fluid losses \cite{16}. Enteropathogenic *Escherichia coli* (EPEC), which is a common cause of traveller’s diarrhoea, impairs intestinal permeability \cite{17} while rotavirus, the major cause of infantile gastroenteritis, limits fluid absorption \cite{18}.

Hypertonic oral rehydration solution (ORS) is able to reduce fluid losses from the small intestine by stimulating Na⁺ and glucose transport, which concomitantly facilitates the absorption of water. This simple intervention is responsible for saving many millions of lives, particularly in developing countries.
Tumours of the small intestine

Benign tumours of the small intestine are rare and are usually derived from either the smooth muscle layers (leiomyoma), fat within the submucosa (lipoma) or from the enteric nervous system (neuroma). Occasionally large lesions may present with small intestinal obstruction or occult gastrointestinal bleeding but most are asymptomatic. Primary malignant lesions of the small intestine are very rare. In up to 1% of the population, tumours derived from enterochromaffin cells (carcinoid tumours) may be found.

### Table 1.4.2 Causes of malabsorption

<table>
<thead>
<tr>
<th>Changes in small intestinal contents</th>
<th>Changes in the small intestinal mucosa</th>
<th>Changes outside the small intestine</th>
<th>Inadequate small intestinal length</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bacterial overgrowth</td>
<td>Inflammation of small intestinal mucosa (Crohn’s disease, diffuse small intestinal vasculitis, NSAID-induced enteropathy, infection)</td>
<td></td>
<td>Malabsorption following gastric bypass surgery</td>
</tr>
<tr>
<td>Lack of bile salts (obstruction to the flow of bile, disruption of the enterohepatic circulation)</td>
<td></td>
<td></td>
<td>Enterocolic and enteroenteric fistulation</td>
</tr>
<tr>
<td>Exocrine pancreatic dysfunction</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

NSAID, non-steroidal anti-inflammatory drug.

### Table 1.4.3 Clinical consequences of specific micronutrient deficiencies

<table>
<thead>
<tr>
<th>Deficient micronutrient</th>
<th>Clinical consequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vitamin A</td>
<td>Night blindness</td>
</tr>
<tr>
<td>Vitamin B1</td>
<td>Wernicke’s encephalopathy</td>
</tr>
<tr>
<td>Vitamin B6</td>
<td>Dermatitis, peripheral neuropathy, angular cheilitis</td>
</tr>
<tr>
<td>Vitamin B12</td>
<td>Anaemia, altered mood, subacute combined degeneration of the spinal cord</td>
</tr>
<tr>
<td>Vitamin C</td>
<td>Scurvy</td>
</tr>
<tr>
<td>Vitamin D</td>
<td>Osteomalacia</td>
</tr>
<tr>
<td>Vitamin E</td>
<td>Neuropathy, myopathy, immunosuppression</td>
</tr>
<tr>
<td>Vitamin K</td>
<td>Impaired clotting (extrinsic clotting pathway – factors 2, 7, 9, 10)</td>
</tr>
<tr>
<td>Iron</td>
<td>Anaemia, glossitis, angular stomatitis</td>
</tr>
<tr>
<td>Calcium</td>
<td>Osteopenia/osteoporosis, muscle spasm, cardiac arrhythmias</td>
</tr>
<tr>
<td>Copper</td>
<td>Myelopathy, peripheral neuropathy, optic neuropathy</td>
</tr>
<tr>
<td>Magnesium</td>
<td>End-organ resistance to parathyroid hormone leading to hypocalcaemia.</td>
</tr>
<tr>
<td>Selenium</td>
<td>Cardiac arrhythmias, myopathy, fatigue</td>
</tr>
<tr>
<td>Zinc</td>
<td>Acrodermatitis enteropathica, reduced fertility</td>
</tr>
</tbody>
</table>
In general, these are small, benign lesions found in the appendix but are more likely to become malignant if located in the small intestine. Such tumours may cause symptoms such as flushing, diarrhoea, sweating and shortness of breath (carcinoid syndrome).

**Short bowel syndrome**

Short bowel syndrome describes problems arising from a reduced small intestinal length, usually following surgical resection. If less than 1 m of small intestine remains, enteral nutrition alone may be inadequate. Patients with short bowel syndrome may present with dehydration, malabsorption, weight loss and micronutrient deficiencies. Occasionally, parenteral nutrition is required to supplement feeding but oral feeding should still be encouraged where possible to prevent GI atrophy.

**1.4.5 Conclusion**

The small intestine is a highly adapted organ with the capacity to absorb water and nutrients far in excess of the body’s requirements. Whilst large portions of the small intestine can be resected without deleterious effects, changes in absorption and secretion homeostasis, for example in gastroenteritis, can cause copious diarrhoea, malabsorption and undernutrition. There are myriad diagnostic tests for the detection of small intestinal pathology and their application should be directed by accurate clinical history taking and examination. Furthermore, disease of the small intestine should be considered in any patient presenting with systemic symptoms occurring as a result of micronutrient deficiency.

**References**

Chapter 1.5

Physiology and function of the colon

Emma V. Carrington and S. Mark Scott
Queen Mary University of London, London, UK

The colon is the principal organ of the distal gastrointestinal tract. It plays a vital role in fluid and electrolyte homeostasis, digestion of food, absorption of nutrients, propulsion of intestinal contents and ultimately expulsion of waste products. Disorders of colonic function such as irritable bowel syndrome, inflammatory bowel disease, chronic constipation and diarrhoea are highly prevalent and cause significant morbidity with a negative impact on quality of life and consequently high socioeconomic costs. A keen understanding of colonic function is therefore required for the successful management of gastrointestinal disease.

1.5.1 Anatomy

Embryology

The primitive intestine begins to form in the third week of gestation. It arises secondary to ventral folding of the embryonic yolk sac and results in a tubular structure, lined with endoderm (ultimately forming the colonic mucosa) and covered with mesoderm (from which arises the surrounding muscle and serosa) [1]. This subsequently develops into foregut, midgut and hindgut regions, an understanding of which allows an appreciation of each section’s resultant blood supply, lymphatic drainage and neuronal innervation. The colon is derived from the midgut and hindgut regions (the midgut spanning from the second part of the duodenum to the middle third of the transverse colon and the hindgut extending from the middle third of the transverse colon to the rectum) [2].

Structure

The colon begins at the caecum and terminates with the rectum. It comprises six sections: caecum, ascending colon, transverse colon, descending colon, sigmoid colon and rectum. The junction of the ascending and transverse colon is commonly referred to as the hepatic flexure and the junction of the transverse and descending colon as the splenic flexure.

The caecum lies in the right iliac fossa, is completely covered by peritoneum and is therefore intraperitoneal. It is approximately 6 cm in length, without mesentery and is relatively mobile. Longitudinal muscle bands called teniae coli (which continue throughout the colon) converge at the base of the appendix, a vestigial organ that originates from its posterior surface.

The ascending colon is a continuation of the caecum. This extends upwards along the lateral side of the abdominal wall towards the right upper quadrant of the abdomen. It is approximately 15 cm in length and is covered on its anterior and lateral surfaces by peritoneum (therefore considered retroperitoneal). Once it has reached the inferior surface of the right lobe of the liver, it turns to form the transverse colon, which passes in front of the second part of the duodenum and the head of the pancreas. Following a further turn beneath the spleen, it
carries on to become the descending colon, which lies retroperitoneally on the left lateral side of the abdominal wall and is approximately 30 cm long. As the colon continues into the pelvis, it becomes known as the sigmoid colon, which finally terminates as the rectum [3].

Aside from its location at the periphery of the abdominal cavity, the colon may be characterised by the presence of taeniae coli and appendices epiploicae (small fatty tags attached to the serosa surface). It is thrown into concertina-like saccular folds referred to as haustra, which are thought to be important for mixing of intestinal contents.

**Vascular supply**

The vascular supply of the colon is determined by its embryological origin. Structures derived from the midgut receive arterial supply from branches of the superior mesenteric artery and are drained by tributaries of the superior mesenteric vein and thence into the portal system. Distal to the middle third of the transverse colon (hindgut in origin), tissue receives arterial supply derived from the inferior mesenteric artery. Similarly, venous drainage is via tributaries of the inferior mesenteric vein (which also subsequently drains into the portal system).

**Neuronal innervation**

Colonic innervation is derived from four sources: the enteric, extrinsic afferent, sympathetic and parasympathetic nervous systems.

The enteric nervous system is composed of a number of nerve plexi within the GI wall and is principally responsible for regulation of colonic motility. The two major plexi are the myenteric plexus and the submucosal plexus. The interstitial cells of Cajal provide the functional link between the nerve processes of the plexi and the muscle cells.

The extrinsic afferent nerves provide sensory innervation of the colon and rectum. The proximal colon receives this supply from the vagus nerve and the distal colon and rectum receive this supply from S1 and S2. It is thought that this innervation is primarily responsible for the conscious perception of rectal filling as well as the initiation of propulsion required for defaecation [4].

Sympathetic and parasympathetic supplies also act to modulate sensory and motor activity. Midgut structures derive this innervation from the superior mesenteric plexus and hindgut structures from the pelvic splanchnic nerves via the inferior mesenteric plexus. Generally speaking, parasympathetic activity exhibits an excitatory effect on colonic function, increasing colonic motility and secretory activity. By contrast, sympathetic activity inhibits colonic tone and motility [5].

**1.5.2 Function**

The colon has evolved to perform four major functions:

1. propulsion of colonic contents towards the rectum and anus for eventual expulsion
2. absorption of water and electrolytes from intraluminal contents
3. absorption of short-chain fatty acids produced by resident microbiota
4. defaecation.

**Propulsion of intestinal contents**

The term ‘colonic motility’ is used to describe the mixing and propulsive movements of the colon that allow for digestion, absorption and transit of intraluminal contents.

The mechanisms responsible for absorption in the colon are slow and the colonic microbiota are facilitated by the speed and orientation of mixing movements. Distal propulsion of contents is therefore gradual to allow for mixing and uniform contact with the colonic mucosa. Contents take roughly 12–30 h to traverse the length of the colon, compared to 2–4 h in the small intestine (which is four or five times greater in length).

Colonic motility patterns are complex. Co-ordinated activity between the terminal ileum, caecum and proximal colon is required to deliver chyme from the terminal ileum to the colon. Contents become increasingly solid as water is absorbed and they are transported aborally toward the rectum for eventual evacuation.

Two forms of colonic contractile activity have been described: propagating pressure sequences
(PSs) (sometimes referred to as high- and low-amplitude propagating pressure sequences) and segmental contractions. Antegrade movement of colonic contents is generally as a result of proximally originating PSs. Frequency significantly increases after waking and/or meal ingestion and may be of high (with a >100 mmHg rise in colonic pressure over a significant length) or lower amplitude (2–5 mmHg increase in pressure). Both high- and low-amplitude PSs are equally likely to produce colonic movement [6].

Localised mixing of colonic contents is achieved through segmental contractions, accounting for the majority of colonic activity. These can be considered as more limited areas of activity facilitating contact with the colonic mucosa principally for absorption of water and other contents.

Disorders of colonic motility can result in impaired stool propulsion and abdominal pain, distension, constipation and diarrhoea.

**Water and electrolyte homeostasis**

Absorption of water and electrolytes is one of the principal functions of the colon. Roughly 1500 mL of effluent reaches the ileocaecal valve each day and the healthy colon will generally resorb 90% of the fluid from this, resulting in the formation of 200 g of solid stool.

Water absorption is intimately associated with sodium reabsorption and occurs primarily in the ascending and transverse colon [5]. Intraluminal Na⁺ passively diffuses into colonocytes via apical channels in response to a negative electrochemical gradient (maintained through the presence of electrogenic Na⁺/K⁺ pumps present on their basolateral membrane). Absorption of water then follows the resultant osmotic gradient via a paracellular pathway and is controlled by both aldosterone and antidiuretic hormone (ADH). Aldosterone acts to increase K⁺ secretion/Na⁺ conservation and ADH increases apical membrane water permeability.

Chloride is also actively absorbed from the colonic lumen. It is transported through the apical membrane via Cl⁻/HCO₃⁻ channels. Secretion of HCO₃⁻ acts to neutralise acidic compounds produced through bacterial fermentation.

**Digestion and absorption**

Although the majority of digestion and absorption occurs in the stomach and small intestine, the colon also plays a role in nutrient salvage through the process of fermentation [5]. The colon contains approximately $10^{11}$–$10^{12}$ bacteria per gram of contents [7] and these bacteria have the ability to break down carbohydrates and proteins into short-chain fatty acids (SCFAs). If required, this can supply up to 15% of an individual’s total energy requirements [8].

The production of SCFAs depends on a number of factors including the constituents of luminal contents, gut transit time and microbial variety. As colonic microbiota will preferentially ferment carbohydrates over proteins, saccharolytic fermentation is predominant in the proximal ascending colon whereas proteolysis is more common in the distal colon.

The three end-products of fermentation are acetate, butyrate and propionate. Although butyrate only accounts for approximately 20% of total SCFA production, it is of particular importance as it is the primary energy source for the colonic mucosa and plays a major role in cellular differentiation and proliferation [9]. Additionally, there is a degree of evidence suggesting that butyrate has anti-inflammatory and anticarcinogenic properties.

Absorption of SCFAs is both passive and active in nature and results in sodium absorption and bicarbonate excretion. The colon is particularly proficient at SCFA absorption (as only 10% of those produced are excreted) [9].

Recognition of the benefits of saccharolytic fermentation has led to the development of prebiotics, probiotics and synbiotics.

In addition to the production of SCFAs, colonic microbiota also play a role in the production of vitamin synthesis, notably vitamin K, biotin (vitamin B7) and niacin (vitamin B3).

**Defaecation**

Effective defaecation is a result of the successful co-ordinated function of the colon, rectum and anus and is under central, spinal and enteric neural control [10]. A number of factors influence defaecatory frequency including diet, intraluminal contents,
colonic transit, behaviour and posture. There is variability in defaecatory frequency between individuals but studies suggest that 99% of the healthy population open their bowels between three times per day and three times per week [11].

Consumption of food produces a near immediate increase in colonic motor activity. Food content affects the degree of colonic response as it has been shown that a fat-rich meal induces colonic motor activity to a greater extent than a protein- or carbohydrate-rich meal [12].

The process of defaecation constitutes four distinct phases: the basal phase (characterised by a change in colonic motor activity, usually precipitated by waking or meal ingestion), the predefaecatory phase (during which gradual rectal distension produces an awareness of rectal filling), expulsion (following a conscious desire to evacuate) and termination (characterised by contraction of the external anal sphincter and closure of the anal canal) [10].

1.5.3 Measurement and assessment of function

Assessment of colonic motility and transit is usually indicated in patients with symptoms of infrequent evacuation. At the present time, two radiological methods for the assessment of transit are routinely employed to look indirectly at colonic motor function: radiopaque marker studies and colonic scintigraphy. Colonic manometry can be employed to look directly at colonic contractile activity.

Radiopaque marker studies

This technique is the simplest method for study of colonic transit times (either total or segmental) [13]. This study is often used for the evaluation of patients with symptoms of persistent constipation. Radiopaque markers contained in a degradable capsule are ingested and plain abdominal X-rays are subsequently taken to determine marker distribution. Segmental colonic transit times can be estimated by administration of different-shaped markers on consecutive days.

Conventionally, this study involves the ingestion of 24 markers on day 1 and the performance of a single plain radiograph on day 6, but in reality there are significant differences in practice between centres, making interinstitution comparisons difficult. The upper limit of normal colonic transit time is around 72 h and delayed transit may be secondary to a primary colonic dysmotility or disorders of evacuation. Further testing with either colonic scintigraphy and/or evacuation proctography is often required to establish a definitive diagnosis.

Colonic scintigraphy

Colonic scintigraphy is a radioscintigraphic method for studying colonic motility. A radioisotope, usually $^{111}$In (indium), is bound to a non-absorbable substance, e.g. diethylenetriamine penta-acetic acid (DTPA), and is either ingested orally with water, or delivered direct to the colonic lumen via an ingestible enteric-coated capsule (i.e. that degrades in the caecum). For data analysis, the colon is generally divided into regions of interest. Time-distribution analysis enables information about activity in a given region of interest to be determined at any one time or for the overall study [14,15].

Colonic manometry

Direct evaluation of changes in intracolonic pressure is termed ‘colonic manometry’. This technique utilises an intraluminal device with the ability to detect pressure changes that occur as a result of phasic contractions of circular colonic muscle. As opposed to radiopaque marker studies and colonic scintigraphy, which provide an indirect assessment of intraluminal movement, colonic manometry is able to characterise specific patterns and phases of colonic motor activity, giving the clinician a greater appreciation of differences in regional function. Regrettably, colonic manometry is an invasive procedure (requiring colonoscopy for placement of the catheter) and for this reason a paucity of data are available in both health and disease, limiting its use in clinical practice [16].

1.5.4 Conclusion

The colon is a complex and responsive structure with function dependent not only on external events and influence of other organs, but also the maintenance of a stable microbiota. The impact of diet should not be underestimated. Diseases characterised by disorders of colonic function are common; for
example, chronic constipation affects 3% of individuals and diverticular disease may affect up to 60% of the population over 60 years of age [17,18]. Maintenance of colonic health is therefore fundamental to ensure a satisfactory quality of life.

References

Chapter 1.6

Physiology and function of the pancreas

Yiannis N. Kallis\(^1\) and David Westaby\(^2\)

\(^1\)Barts Health NHS Trust, London, UK
\(^2\)Imperial College Healthcare NHS Trust, London, UK

1.6.1 Anatomy, physiology and function

The pancreas is a retroperitoneal organ located in the upper abdomen. It extends transversely between the concavity of the duodenum and the spleen, and lies posteroinferior to the stomach. The medial aspect of the pancreas receives a blood supply from branches of the gastroduodenal and superior mesenteric arteries, whilst branches of the splenic artery supply the bulk of the pancreatic body and tail. Blood drains into the portal venous system via the superior mesenteric and splenic veins. Lymph drainage of the pancreas is via splenic, coeliac and superior mesenteric lymph nodes, which are common sites of metastatic cancer spread.

The pancreas performs both exocrine and endocrine functions and plays a central role in digestion and glucose metabolism. Pancreatic exocrine secretions drain into the medial aspect of the second part of the duodenum via tributaries that form the main pancreatic duct. The duct enters the duodenum at the ampulla of Vater, into which the common bile duct also drains. Outflow is controlled by a smooth muscle sphincter termed the sphincter of Oddi. Endocrine cells of the pancreas release hormones directly into the bloodstream.

The exocrine pancreas consists of units called acini, which are arranged into lobules and drain into the main pancreatic duct. They make up 98% of the pancreatic mass and are responsible for the production of digestive enzymes and pancreatic fluid. Enzymes, such as trypsin, chymotrypsin, lipase, phospholipase A2 and amylase, are stored as inactive precursors within secretory granules and are released under neurohormonal control. These proenzymes are activated by intestinal enteropeptidases only upon reaching the duodenum, thereby preventing pancreatic autodigestion due to premature activation. Dysregulation of these mechanisms is thought to underpin the pathogenesis of pancreatitis. Enzymes are released in large quantities for the early digestion of proteins, fats and carbohydrates. Pancreatic ductal cells secrete approximately 2 L of bicarbonate-rich fluid daily to neutralise duodenal chyme and optimise conditions for digestion.

Exocrine function is governed by multiple neurohormonal pathways triggered by the process of eating. The autonomic nervous system directly induces pancreatic enzyme release via vagal parasympathetic efferents in response to cephalic stimuli (e.g. the sight and smell of food), and also after gastric distension. Duodenal exposure to food and acidity induces the release of gut hormones from specialised intestinal enteroendocrine cells. Cholecystokinin governs acinar cell degranulation and secretin is primarily responsible for alkaline pancreatic secretion.

Endocrine cells are distributed throughout the pancreas in spherical clusters called islets of Langerhans, which are criss-crossed by a dense network of capillaries. Beta-cells, the predominant cell type, are responsible for insulin production and alpha-cells synthesise glucagon, both key hormones
in glucose homeostasis. Other cell types secrete inhibitors of pancreatic exocrine secretion, such as somatostatin, which also have inhibitory effects on islet cell function.

1.6.2 Measurement and assessment of function

Structural assessment

Cross-sectional imaging with arterial phase contrast-enhanced computed tomography (CT) scanning is the gold standard for detection of pancreatic lesions and for assessment of complications of acute and chronic pancreatitis [1]. Magnetic resonance imaging (MRI) is a suitable alternative and can provide more detailed information about pancreatic ductal anatomy, via magnetic resonance cholangiopancreatography (MRCP). Ultrasound scanning is of limited utility because overlying intestinal gas frequently obscures views.

The role of endoscopic ultrasound in the investigation of pancreatic disease is rapidly expanding. It is particularly helpful in the staging of pancreatic cancer and the diagnosis of pancreatic cystic lesions [2]. Fine needle aspiration facilitates histological diagnosis and large pancreatic cysts can be managed by transgastric stent insertion.

Assessment of function

Acute pancreatic inflammation is typified by release of enzymes into the bloodstream, and simple assays are widely available to quantify serum amylase and lipase in the diagnosis of acute pancreatitis.

Pancreatic exocrine insufficiency is common in chronic pancreatic disease, though clinically evident malabsorption only usually occurs after 85–90% reduction in enzyme production. Endocrine insufficiency is typically a late feature and is manifest by the development of diabetes.

A variety of assays of exocrine function are available though many are expensive and time consuming, and most are rarely performed. Measurement of faecal elastase is a useful screening tool for moderate or severe pancreatic insufficiency and is commercially available [3]. Elastase is exclusively produced by the pancreas, is not enterally absorbed, and has largely replaced quantification of faecal fat or chymotrypsin as the assay of choice. Faecal fat excretion of more than 7 g per day (100 g daily fat intake) is considered indicative of fat malabsorption, though may represent intestinal disorders as well as pancreatic disease. Direct analysis of endoscopically obtained duodenal aspirates following pancreatic stimulation is laborious and costly and largely obsolete.

Quantification of exocrine function can also be performed by indirect assessment of enzymatic activity. The pancreolauryl and PABA (N-benzoyl-L-tyrosyl-p-aminobenzoic acid) tests are two such assays, whereby orally administered, labelled compounds are digested by luminal pancreatic enzymes, releasing substrates quantifiable in urine. They are only reliable for detecting severe insufficiency and are not widely available.

1.6.3 Pathology

Acute pancreatitis is a condition of sudden onset usually precipitated by acute pancreatic injury. Common causes include gallstones and alcohol, though drugs, trauma and other rare triggers are also recognised. Its severity can range from mild to life threatening, and a number of prognostic scoring systems have been developed [4]. Its pathophysiology involves the premature activation of pancreatic enzymes leading to tissue autodigestion and necrosis, which can also trigger a systemic inflammatory response syndrome. Early treatment is largely supportive, and later management is often focused on local complications such as fluid collections or abscesses [5]. The role and timing of enteral feeding during an acute episode are under ongoing review [6]. Identification and removal of precipitating factors are important to prevent future episodes. Occasionally patients may experience recurrent discrete episodes classified as acute relapsing pancreatitis.

Chronic pancreatitis is defined by abdominal pain, which is often intractable, and involves irreversible fibrosis, atrophy and calcification of the gland. It may involve features of exocrine or endocrine insufficiency. The most common cause is
1.6 Physiology and function of the pancreas

Excess alcohol consumption, though many cases are classified as idiopathic. Other aetiologies include inherited genetic abnormalities, autoimmune disorders and conditions associated with impaired pancreatic drainage (e.g. pancreas divisum). The pathogenesis of chronic pancreatitis is increasingly recognised as multifactorial [7]. Treatment incorporates avoidance of precipitating factors and adequate analgesia, though supplementation of exocrine/endocrine function and maintenance of nutrition are also important aspects. Endoscopic or surgical procedures to improve pancreatic duct drainage may be required.

Ductal adenocarcinoma is the most common pancreatic malignancy and is often locally advanced or metastatic upon presentation. Surgical resection offers the only prospect of a cure, but a majority of cancers are inoperable at the time of diagnosis [8]. Palliative chemotherapy and relief of biliary or gastric outlet obstruction with endoscopic stenting or surgery are the mainstays of treatment.

Neuroendocrine tumours are commonly located in the pancreas. They are a heterogenous group of tumours of variable metastatic potential, with a more favourable prognosis [9]. They may cause symptoms by dysregulated release of hormones such as insulin, glucagon or gut hormones (e.g. gastrin). Surgical resection and somatostatin analogues are mainstays of treatment.

Cystic lesions of the pancreas are often found incidentally on abdominal imaging. They can be divided into simple cysts, pseudocysts or true cystic neoplasms. Management of these lesions is focused on defining their malignant potential [10]. Size and characterisation of the cyst are key, and endoscopic ultrasound has advanced this field significantly. Surgical resection is considered where lesions are symptomatic or have high malignant potential.

References

1.7.1 Anatomy, physiology and function

The liver is situated in the right hypochondrium and is split into right and left lobes. It can be further divided into eight functional segments according to vascular supply and biliary drainage [1]. Inflow of blood is via a dual supply, with approximately 25% derived from the hepatic artery and 75% from the portal vein, which drains the gastrointestinal tract, pancreas and spleen. Both vessels enter the liver at the hilum and subdivide into smaller branches, running in structures called portal tracts, also composed of bile ducts and lymphatics. Blood perfuses the liver within sinusoids before draining via tributaries to form the hepatic vein, which enters the inferior vena cava just beneath the diaphragm.

Histologically, liver parenchyma is arranged into units called lobules, defined by a central hepatic venule and multiple peripheral portal tracts [2]. Hepatocytes are organised into three-dimensional plates separated by sinusoids, which are lined by fenestrated endothelium through which blood can readily permeate. Kupffer cells (liver macrophages) and hepatic stellate cells (fibroblast-like collagen-producing cells) lie in close approximation [3]. Spatially, hepatocytes are classified into three zones, with those around portal tracts in zone 1 receiving the most oxygenated blood and those in zone 3 around hepatic venules the least [4].

Bile canaliculi form a dense meshwork around hepatocytes and fuse to form bile ducts within portal tracts. The right and left hepatic ducts join at the hilum to form the common hepatic duct. The cystic duct drains the gallbladder and fuses with the common hepatic duct to form the common bile duct. The latter enters the duodenum at the ampulla of Vater (see Chapter 1.5).

The gallbladder is located in the abdomen under the right lobe of the liver. It is a pear-shaped sac, with average volume of 50 mL and length of 9 cm in healthy adults (range 4–14 cm), and is connected to the gastrointestinal tract via the cystic duct and common bile duct [5]. Its function is to concentrate and store bile produced in the liver and to release this into the duodenum when required for digestion. Release of bile is achieved by muscular contraction of the gallbladder in response primarily, but not exclusively, to gut hormones, especially cholecystokinin (CCK), that are stimulated by the products of digestion in the gastrointestinal tract. In the small intestine, bile is essential for the emulsification of dietary fat.

Carbohydrate metabolism

The liver plays a major role in glucose homeostasis and is the main store of glycogen. It assimilates excess glucose into glycogen within hepatocytes under insulin control. Conversely, hepatic glycogenolysis makes
glucose readily available to other organs as the first response to starvation. Thereafter, the liver can also release glucose via gluconeogenesis from non-carbohydrate substrates such as lactate, amino acids (alanine and glutamine) and glycerol. During prolonged starvation, hepatic fatty acids are metabolised by beta-oxidation to ketone bodies, an important energy source for organs such as the brain.

**Protein metabolism**

The liver regulates plasma amino acid levels by controlling amino acid transamination and gluconeogenesis. All circulating plasma proteins except gamma-globulins are synthesised by the liver. Approximately 10–12 g of albumin is produced daily, helping to maintain oncotic pressure and transport water-insoluble compounds. Coagulation cascade factors, including fibrinogen, are produced in the liver, as are acute phase and complement proteins. The liver is the primary site of nitrogen excretion via amino acid transamination and oxidative deamination, leading to the formation of ammonia, which is subsequently converted to urea and excreted renally.

**Lipid metabolism**

Following dietary fat absorption, the liver synthesises triglycerides from free fatty acids for redistribution around the body within very low-density lipoproteins (VLDL). Likewise, hepatic fatty acid oxidation can be utilised for energy release. The liver controls production of circulating lipoproteins which transport insoluble fats through the bloodstream. Cholesterol formation, excretion and redistribution are also under hepatic regulation. Circulating cholesterol is taken up via hepatic low-density lipoprotein (LDL) receptors, and cholesterol can be formed de novo from hepatic acetyl-CoA. Cholesterol esterification to fatty acids also takes place in the liver.

The liver is a store for several vitamins and minerals, including vitamins A, D, B12, iron and copper, and is the site of 25-hydroxylation of cholecalciferol.

**Bile acid and bilirubin metabolism**

Bile is composed of water, electrolytes, bile acids, bilirubin, phospholipids, cholesterol and conjugated waste products, with approximately 600 mL produced each day [6]. In fasted states, approximately half is syphoned off to the gallbladder where it is concentrated. It is actively secreted at the hepatocyte canicular membrane via bile transporter proteins, with biliary ductular epithelium also contributing. It facilitates the emulsification and digestion of fats and provides an alkaline pH for optimal pancreatic enzyme function. It is also the main vehicle for the elimination of hydrophobic waste products such as bilirubin. Bile formation is stimulated by secretin and inhibited by somatostatin. After ingestion of a meal, cholecystokinin stimulates gallbladder contraction and sphincter of Oddi relaxation.

Primary bile acids are synthesised from cholesterol in the liver and these are converted into secondary bile acids by GI microbiota. Ninety-five per cent of these are reabsorbed in the terminal ileum, returned to the liver in the portal venous circulation and resecreted into bile, termed the enterohepatic circulation [7]. Bilirubin is formed from erythrocyte breakdown, conjugated with glucuronic acid in the liver to render it water soluble and excreted within bile. Some is reabsorbed via the enterohepatic circulation after bacterial hydrolysis to urobilinogen.

**Drug and hormone metabolism**

Most xenobiotics, including alcohol, are inactivated in the liver by cytochrome P450-mediated processes such as methylation or hydroxylation, and excreted in bile or urine after conjugation by hepatic transferases [8]. The liver is a key site of the catabolism of hormones such as oestrogens, insulin, growth hormone, glucocorticoids and parathyroid hormone. Angiotensinogen is produced in the liver and helps regulate blood pressure.

**Immunological function**

As part of the reticuloendothelial system, the liver is a crucial site of gut-derived antigen presentation to Kupffer cells, NK cells and sinusoidal endothelium transported in the portal venous circulation [9]. It is also involved in adaptive immunity via lymphocyte trafficking from the GI tract. Impaired immunity and recurrent sepsis are common consequences of hepatic dysfunction.
1.7.2 Measurement and assessment of function

Blood tests
Liver function tests (LFTs) can be subdivided into true tests of liver synthetic and excretory function, and markers of liver injury measured by cellular enzyme release. Alanine aminotransferase (ALT) and aspartate aminotransferase (AST) are enzymes predominantly located in hepatocytes and their release into serum reflects hepatocellular damage. Alkaline phosphatase (ALP) and gamma-glutamyl transpeptidase (gamma-GT) are chiefly found in biliary canaliculae and serum rises indicate intrahepatic cholestasis as well as extrahepatic bile duct damage or obstruction. ALP is also located in other tissues such as bone and placenta, and gamma-GT is an enzyme inducible by certain drugs and alcohol. Commonly measured indices of liver function are bilirubin, albumin and prothrombin time. Advanced liver impairment can be associated with renal failure, hence serum sodium and creatinine concentrations are included in scoring systems of liver dysfunction such as MELD and UKELD [10,11].

Blood tests can help determine the cause of liver injury, and include viral hepatitis serology, immunoglobulins, autoantibody profiles, serum iron and copper indices and a metabolic screen. Alpha-fetoprotein (AFP) and Ca19.9 are tumour markers associated with hepatocellular carcinoma (HCC) and cholangiocarcinoma, respectively, though their clinical utility is limited.

Imaging
Ultrasound scanning is usually the first modality employed to define anatomy in hepatobiliary investigation. The presence of liver lesions, gallbladder stones, biliary obstruction, portal hypertension or ascites can be readily detectable. Fatty liver infiltration gives an echo ‘bright’ signal. Doppler is used to interrogate blood flow and detect portal or hepatic venous occlusion.

Cross-sectional imaging with contrast-enhanced triple phase computed tomography (CT) or magnetic resonance imaging (MRI) can better define hepatic or biliary mass lesions and gives a clearer indication of liver size and architecture. HCC can be diagnosed on radiological characteristics alone, obviating the need for biopsy. CT or MRI scanning is also invaluable in the management of cholangiocarcinoma to define anatomical relationships to blood vessels and bile ducts and thus determine surgical resectability. MR cholangiopancreatography is the ‘gold standard’ modality to delineate the biliary tree and can detect calculi, strictures or diffuse cholangiopathy.

Hepatobiliary iminodiacetic acid (HIDA) scanning is a dynamic radionucleotide test of bile flow and is still sometimes used in the investigation of the jaundiced patient to determine if cholestasis is of hepatic or biliary origin.

Endoscopy
Upper gastrointestinal (GI) endoscopy is used for the diagnosis and treatment of complications of portal hypertension such as oesophageal or gastric variceal bleeding. Endoscopic retrograde cholangiopancreatography (ERCP) is the therapeutic modality of choice in the management of bile duct stones or strictures.

Interventional radiology
Percutaneous transhepatic cholangiography (PTC) is sometimes performed when ERCP is unsuccessful. The biliary tree is accessed under fluoroscopic guidance and obstruction can be relieved by internal or external stenting. It is particularly useful in the context of complex hilar strictures.

Fluoroscopically guided cannulation of the hepatic vein via a transjugular approach can allow the measurement of hepatic and portal venous pressure in the management of portal hypertension, and can offer an alternative approach to obtaining a liver biopsy. The transjugular placement of an intrahepatic portosystemic shunt (TIPSS) can control variceal bleeding when endoscopy has failed [12]. Hepatic arterial angiography is employed to embolise liver tumours such as HCC, which have a dense hepatic arterial supply.
Liver biopsy

Histological examination of liver tissue is frequently invaluable in the diagnosis of parenchymal liver disease and in the characterisation of liver or biliary mass lesions [13]. Estimation of liver fibrosis permits the staging of chronic liver disease and establishes a diagnosis of cirrhosis. Most liver biopsies are performed percutaneously under ultrasound guidance, though transjugular and laparoscopic approaches are alternatives.

1.7.3 Pathology

Liver disease is often asymptomatic and can be present for many years before diagnosis. Symptoms occur late and are characterised by jaundice or features of hepatic decompensation such as ascites, peripheral oedema, encephalopathy or portal hypertensive GI bleeding.

Severe acute liver injury is rare though may rapidly lead to liver failure. Paracetamol overdose is the most common cause, though other aetiologies such as idiosyncratic drug reactions, viral hepatitis and autoimmune liver disease are recognised.

Chronic liver disease is characterised by a balance between progressive fibrosis and attempts at liver regeneration. Cirrhosis is the hallmark of advanced liver disease and can herald the development of liver dysfunction and HCC. Undernutrition and loss of muscle mass are common features. Excess alcohol consumption, non-alcoholic fatty liver disease (NAFLD) and chronic hepatitis C are the most common causes in the developed world. Other important aetiologies include hepatitis B, haemochromatosis and immune disorders such as primary biliary cirrhosis (PBC), primary sclerosing cholangitis (PSC) and autoimmune hepatitis. Management includes the withdrawal or treatment of the causative agent where possible, and the treatment of complications of liver decompensation. Liver transplantation can offer good medium-term survival in end-stage disease [14].

Biliary disorders are characterised by upper abdominal pain, jaundice, fever or weight loss. Gallstone disease can present with biliary colic, obstructive jaundice or cholangitis. Isolated strictures of the biliary tree may be benign or malignant and usually present with LFT derangement and jaundice. Diffuse stricturing of the biliary tree is distinctive of a cholangiopathy, e.g. PSC or autoimmune cholangiopathy. Biliary obstruction can also occur from extrinsic compression, such as from intrahepatic or lymph node metastases.

Worldwide, common disorders affecting the gallbladder are the formation of stones, functional dyskinesia, cancer of the gallbladder and steatocholecystitis [15–18].

Hepatocellular carcinoma typically arises in the context of cirrhosis, though it can occur in the precirrhotic liver in conditions such as hepatitis B and NAFLD. Benign liver lesions such as haemangiomata, adenoma and focal nodular hyperplasia (FNH) are also common. Primary biliary tract cancer (gallbladder carcinoma and cholangiocarcinoma) is a relatively common malignancy and carries a poor overall prognosis [19].

References


Chapter 1.8

Gastrointestinal microbiota

Katherine Stephens, Gemma E. Walton and Glenn R. Gibson

University of Reading, Reading, UK

Bacteria are associated with all areas of the human body from the skin to the genitourinary, respiratory and gastrointestinal (GI) tracts [1]. The GI tract is the most heavily populated, with the majority of the total bacterial population of humans residing therein. A highly diverse ecosystem exists, with the collective bacterial species within the human GI tract totalling in the thousands [2,3]. The results of the MetaHIT Consortium (Metagenomics of the Human Intestinal Tract, www.metahit.eu/) indicate that any one of 1000–1150 different species could populate the human GI tract, with at least 160 species residing in an individual [4]. Given these large numbers, although there is great potential for diversity in the GI microbiota between different humans, there is considerable stability in some species, with a core of 18 species being found in all those in the MetaHIT Consortium, and a core of 57 species found in 90% of subjects [4].

1.8.1 Composition

The GI tract has evolved to become a functional organ comprising anatomically distinct areas. The digestive process starts in the oral cavity, then moves through the stomach, small and large intestine and finally the rectum. This passage allows the presence of several microbial niches due to different environmental conditions, such as acidity in the stomach, varying retention times and different nutrient availabilities (Table 1.8.1). Physicochemical variables are contributing factors to the diverse community of micro-organisms residing in the GI tract (see Table 1.8.1). Within the intestinal tract, genomic analysis has shown the number of micro-organisms to be approximately $10^{13}$ to $10^{14}$ in total [5], with the overall microbiome (the combined genome of all the micro-organisms) approximately 100 times greater than the human genome [4]. Within the large intestine, there is also variation in diversity of species within specific compartments, such as the mucosa, lumen and epithelium [6]. The small intestinal sites, duodenum, jejunum and ileum, also comprise differing numbers and species.

Microorganisms residing within the GI tract carry out many necessary roles, for example in metabolism, immune defence and GI physiology [7]. Some are associated with health benefits whereas others are known to be potentially pathogenic. Lactobacilli and bifidobacteria are associated with many positive effects and have been used in various health food products as probiotics. A possible reason for this could be their ability to prevent commensal and potentially pathogenic microbial population levels from increasing through various inhibitory mechanisms [8–11]. Potential pathogens include Clostridium difficile, Escherichia coli and Helicobacter pylori which have been connected with antibiotic-associated diarrhoea, vomiting and stomach ulcers respectively [12,13].

Although each individual has a distinctive microbiome, the majority of key players remain the same but in varying quantities.
1.8.2 Functions of the human gastrointestinal tract

A main function of the GI microbiota is modulation of the immune system. Germ-free mice have been extensively used in studies investigating the involvement of the microbiota in immune response development [14]. The microbiota can form a protective barrier which decreases the chance of pathogen invasion by possibly occupying receptor sites in the GI tract [14]. The microorganisms compete by several different mechanisms, such as nutrient scavenging, receptor occupation and the production of antimicrobial substances, which can elicit a specific or non-specific effect such as the modulation of pH. Antimicrobial substances produced in the GI tract include acids, antimicrobial peptides (AMPs), defensins, cathelicidins and C type lectins, all of which are capable of targeting bacterial cell walls, thus controlling population levels of commensal organisms or aiding protection against pathogens [15,16].

Competition plays a vital role in immune defence, helping to prevent potential pathogen invasion.

Table 1.8.1 Summary of microbiota associated with the GI tract in humans

<table>
<thead>
<tr>
<th>Site</th>
<th>Approximate numbers per mL</th>
<th>Examples of microbial types</th>
<th>Environmental factors</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stomach</td>
<td>$10^{3}$</td>
<td><em>H. pylori</em> <em>Lactobacillus</em> spp. <em>Veillonella</em> spp. <em>Staphylococcus</em> spp. <em>Streptococcus</em> spp.</td>
<td>Microaerophilic Low pH due to gastric acidity from hydrochloric acid Presence of pepsin Rapid transit</td>
<td>14, 58, 59</td>
</tr>
<tr>
<td>Large intestine</td>
<td>$10^{12}$</td>
<td></td>
<td>Anaerobic</td>
<td>60</td>
</tr>
</tbody>
</table>
Specialised GI tract lymphoid tissues produce secretory immunoglobulin A (IgA) [17] which neutralises receptors on target bacteria, allowing some control over the GI microbiota [18]. Activation of IgA is due to localised GI dendritic cells, which sample the luminal micro organisms; therefore antibodies against GI microbiota have already been developed.

A number of features aid in the control of GI population levels, for example IgA and AMPs. Dendritic cells (DCs) are specialised white blood cells which act as antigen-presenting cells (APCs); they sample the intestinal lumen, and therefore GI microbiota, and are able to secrete antibodies to neutralise any potential growing threat [18]. Distinguishing between threats involves Toll-like receptors which are expressed on eukaryotic cells; these have a unique function of recognising conserved regions within bacterial membranes [19]. Due to this ability, signalling molecules such as cytokines can elicit an inflammatory response [20]. Antimicrobial peptides have the ability to work across the GI tract; they are localised towards the intestinal mucosa, preventing the expansion of microbes throughout the lumen and minimising contact with host GI tract epithelium [21]. Lactic acid bacteria produce lactate and acetate, which can be detrimental to other microbes, through their ability to disrupt bacterial outer membranes [22].

The GI tract must also be able to tolerate microbes and not always elicit an immune response. This can be achieved in three different ways: a physical barrier between host cells and bacterial cells, antigen modification on bacterial cells or modifying immune responsive cells in the GI tract [14]. DC’s are specialised in the GI tract to induce and stimulate T-cell differentiation into T-helper cells and T-regulatory cells, an alternative to cytotoxic T-cells which can damage the GI tract epithelial lining [23]. Another potential problem is lipopolysaccharide (LPS) on the gram-negative bacterium’s outer membrane; host recognition of LPS can lead to septic shock or low-grade chronic inflammation [24]. To overcome this, LPS toxicity can be reduced by phosphorylation [25]. In mice, it has been shown that GI epithelial cells inherit a tolerance to LPS endotoxin [26].

Bacterial metabolism is a key part of the microbiota. They are able to breakdown non-digestible food products into short-chain fatty acids (SCFA). Such substrates include non-starch polysaccharides (NSP), starch, oligosaccharides, proteins and amino acids [27]. These organic acids can be used for growth and energy, not only for themselves but as a secondary source for the host [28]. Acetate, propionate and butyrate are the main SCFAs produced and have various impacts on human metabolism and the immune system [28,29]. Butyrate is involved in cytokine development as an essential signalling molecule and provides structural aid in the intestinal epithelium; it also stimulates apoptosis and therefore is an important growth regulator for colonocytes [30,31]. Acetate can aid intestinal inflammation during an immune response, allowing for more immune cells to translocate to the infected site via G-protein-coupled receptors. Acetate is also metabolised in muscle and other systemic tissues [32]. Propionate has been shown to lower cholesterol concentration [33]. SCFAs also have abilities in AMP generation, aiding in immune system defence [5].

Studies have shown that microbial GI composition plays a role in human brain development and behaviour, with germ-free mice displaying higher anxiety issues and less motor control than conventionally raised animals [34]. *Bifidobacterium infantis* has been shown to regulate the metabolism of tryptophan, an amino acid involved in the production of serotonin showing a potential link between GI micro-organisms and neurotransmitter concentrations [35]. As such, the GI microbiota may have an additional impact on host psychology. The microbiota have also been shown to interfere with the hypothalamic-pituitary-adrenal axis—interactions between the hypothalamus, pituitary and adrenal glands [36]. The GI microbiota have been associated with the control of different signalling molecules such as neurotransmitters. These connections suggest that the GI microbiota have an impact on host response to stress as well as mood/psychological disorders [37,38].

1.8.3 Factors influencing composition of the microbiota

The establishment of the native microbiota can be observed from birth and continues to develop throughout life. During pregnancy, the infant’s
intestine tract is thought to be devoid of microorganisms. The delivery method of the infant can result in distinctive colonisation patterns. Natural birth delivery, where the infant passes through the birthing canal, results in the infant ingesting the mother’s commensal vaginal and faecal microbiota [39]. A caesarean birth results in the first colonisers being those from the hospital environment; species such as Staphylococcus epidermis and other Staphylococcus spp and Propionibacterium have been noted in caesarean births [40,41]. Facultative anaerobes are the first GI tract colonisers due to the infant GI tract having positive oxidation/reduction potential [8]. Examples include E. coli, Enterococcus faecium, Enterococcus faecalis, Pseudomonas spp, Aeromonas spp, Klebsiella spp and Enterobacter spp [36,38]. Oxygen is rapidly utilised by initial invaders, thus creating anaerobic conditions which allow the colonisation of strict anaerobes [37]. Examples of such strict anaerobic bacterial genera include Bifidobacterium spp, Bacteroides spp and Clostridium spp [42].

From birth, diet will also affect the initial colonisers, particularly in the case of breast as opposed to formula feeding, with the former having a preponderance of bifidobacteria. Much research has shown that the initial microbiota composition can have an impact on subsequent colonisation which may later influence the health of the individual [14]. For example, early colonisers of lactobacilli have been associated with a lower number of allergies [43]. After 3 years of age, post breastfeeding and weaning, the GI tract starts to stabilise and over time a more established microbiota is developed [44]. In general, breastfed infants have reduced risks of infections and more chronic issues in later life [39].

In the elderly population the microbiota is more changeable [45]. Composition varies and the diversity of micro-organisms has been observed to decrease [46]. Factors which can affect this altered organisation over time include loss of appetite and therefore less nutrient availability, decrease in saliva secretion, decrease in vitamin synthesis, tooth decay, potential mutations in cancer suppression genes, immunological changes, decreases in nutrient absorption and intestinal transit time and sensitivity [47,48]. Lactobacillus spp and Bifidobacterium spp have been observed to be lower in elderly volunteers whereas Bacteroides, enterococci, enterobacteria and Clostridia levels were fairly similar or even higher than in younger adults [49].

Diet can affect bacterial diversity in the GI tract. Non-digestible nutrients will become available to the microbiota; certain species may thrive depending on substrate availability and type [42]. The energy can be harboured for their own metabolic processes or can be available to the host. Diet cannot always provide the vital nutrients the body needs to function, and in this context, the microbiota is important for the synthesis of certain vitamins [28].

Microbial infection can occur at any stage of life. The usual treatment consists of a recommended antibiotic to which the proposed bacterial infection shows sensitivity. However, antibiotic use may also have an impact on the normal indigenous microbiota which can lead to complex issues such as diarrhoea or pseudomembranous colitis [50]. Commensal micro-organisms such as lactobacilli and bifidobacteria may decrease [20]. As a result, opportunistic micro-organisms such as C. difficile and yeasts such as Candida albicans may be better able to multiply due to less competition; these then may have the ability to cause further illnesses, such as antibiotic-associated diarrhoea (AAD) [13,20]. Other issues include vitamin deficiencies, as members of the microbiota contribute to the vitamin requirements of the host [51]. The severity of deficiencies is dictated by a series of factors including dosage and duration of antibiotic treatment, range of potential microbial targets, route of transmission of the treatment, pharmacokinetics of the drug and how easily it can be metabolised [52]. Clindamycin is an antibiotic commonly used for individuals suffering with a health issue caused by an anaerobe, its wide range of targets making it a useful drug on its own and in combination with others. Although this drug is effective, it has a negative impact on commensal GI micro-organisms, allowing for an increase in C. difficile and therefore the risk of colitis, diarrhoea and bloating [51,53].

A growing predicament with the use of antibiotics is increasing bacterial resistance – commensal to commensal or commensal to pathogen [52]. With growing resistance to antibiotic treatments, it is more likely that micro organisms can transfer resistant genes
1.8 Gastrointestinal microbiota

to one another via horizontal transfer, e.g. bacterial conjugation. An example of this was the transfer of beta-lactamase on a plasmid from a resistant *E. coli* strain to an initially susceptible strain in a child taking amoxicillin[54]. It is now thought that over a short antibiotic treatment period, resistant strains can remain for several years which may lead to less successful treatment and higher costs due to failure to eradicate pathogenic infection and the evolution of superbugs[55].

Dysbiosis of the GI microbiota can occur, influencing microbiota composition and leading to potential health problems (Box 1.8.1). A common cause of dysbiosis is inflammation, which is associated with many GI-related diseases[56]. Current research has shown a connection between the microbiota and health issues such as obesity, diabetes, cancer, autism, allergies, inflammatory bowel disease and irritable bowel syndrome. There is much debate about the exact species involved in these disorders, with some studies contradicting one another. However, what is clear is that the microbiota can markedly affect host health.

The microbiota are a crucial component of the human body, required not only in manufacturing necessities such as vitamins and SCFAs but also for the digestive process to occur optimally. Their role in immune defence is of great magnitude and the human body would be at much higher risk of infection without the protection of the commensal microbiota. However, there are also negative effects, which can be mostly controlled through a healthy lifestyle. Our knowledge of the relationship between the host and their microbiota is developing further, with new studies providing more insight into the complex network of our GI system.

**References**


**Box 1.8.1 Examples of factors that may affect the composition of the GI microbiota, adapted from Fooks and Gibson [61]**

- Other microbiota
- Type of feeding
- Amount, chemical composition and availability of growth substrates
- Availability of colonisation sites
- Immunological interactions
- Individual fermentation
- Strategies by the bacteria
- Intestinal transit time
- Gut pH
- Redox potential
- Availability of inorganic electron acceptors
- Production of bacterial metabolites
- Presence of antimicrobial compounds
- Xenobiotic compounds
- Age of the host
- Peristalsis
- Host genetics
- Physical activity levels
- Antibiotics
- Disease state
- Stress


Despite fluctuations in food intake and physical activity, healthy adults maintain a relatively constant weight over decades. However, as stated by the laws of thermodynamics, if less energy is expended than consumed then the excess energy will be stored. It is calculated that an average North American man will increase his weight by 9.1 kg between 25 and 35 years of age, as a consequence of a mere 0.3% imbalance between energy consumed and energy expended over this period [1]. It is therefore of no surprise that, as a result of readily available high-energy food and our sedentary lifestyles, obesity has become a growing global epidemic. The converse is true in disorders that culminate in reduced energy intake such as anorexia nervosa. An understanding of the mechanisms that control body weight, by co-ordinating food intake and energy expenditure, is key in unravelling the pathogenesis of disordered energy homeostasis in gastrointestinal disease.

1.9.1 Role of the gut neuroendocrine system in appetite regulation

Several neural, hormonal and psychological factors control the complex process known as appetite. The hypothalamus and brainstem receive these peripheral neural and hormonal signals and co-ordinate a response in order to achieve energy homeostasis. Two discrete populations of neurones present in the arcuate nucleus (ARC) of the hypothalamus with opposing effects on food intake are crucial in this process: medially located orexigenic neurones (i.e. those stimulating appetite) express neuropeptide Y (NPY) and agouti-related peptide (AgRP), and anorexigenic neurones (i.e. those inhibiting appetite) in the lateral ARC express pro-opiomelanocortin (POMC) and cocaine and amphetamine-regulated transcript (CART) [2]. Additionally, ARC neurones also project onto the paraventricular nucleus (PVN) of the hypothalamus where important efferent pathways regulating energy expenditure arise.

The important inputs to this intricate neural network are twofold. Firstly, the short-term signals that govern meal ingestion are primarily regulated by the ‘gut–brain axis’ and secondly, information regarding long-term energy stores is signalled via leptin, an adipose-derived hormone [3]. This ‘gut–brain axis’ exists to contribute to the short-term feelings of satiety and hunger, by transmitting information from the gastrointestinal tract to the hypothalamus and brainstem, via gut hormones and the vagus nerve (see Figure 1.9.1). The majority of these gut hormones are anorexigenic and include peptide tyrosine-tyrosine (PYY), pancreatic polypeptide (PP), glucagon-like peptide-1 (GLP-1), oxyntomodulin (OXM) and cholecystokinin (CCK). The only truly orexigenic hormone to be discovered thus far is ghrelin. These hormones act in concert as meal initiators and terminators [4].

These endogenous gut hormones act on the central nervous system either via the circulation
through areas deficient in the blood–brain barrier, such as the median eminence of the hypothalamus and the area postrema [5] or via receptors of vagal afferents, together with stretch receptors and nutrient chemoreceptors [6]. These signals converge in the nucleus of the tractus solitarius (NTS) of the brainstem and are integrated with information from higher brain centres relaying reward drive and mood to regulate appetite and control energy expenditure. In this chapter, a review of the key gut hormones implicated in appetite regulation is undertaken.

### 1.9.2 Peptide tyrosine-tyrosine

Peptide tyrosine-tyrosine is a member of the PP-fold family of peptides, which also includes the anorexigenic PP and the orexigenic neurotransmitter NPY, all sharing a common tertiary structure. PYY exists endogenously in two forms: peptide tyrosine-tyrosine$_{1-36}$ and PYY$_{3-36}$ [7]. Enzymatic cleavage of secreted PYY$_{1-36}$ at the amino terminal by the enzyme dipeptidyl peptidase IV (DPP-IV) gives rise to PYY$_{3-36}$, the predominant form of circulating PYY.

Peptide tyrosine-tyrosine is released postprandially by the L-cells of the distal gut in proportion to the energy content consumed. Plasma PYY$_{3-36}$ concentrations rise within 15 min of food ingestion, well before nutrients reach the colon or rectum where it is released, implicating a neural or hormonal mechanism for its release [9]. PYY$_{3-36}$ concentrations peak 1–2 h postprandially and remain elevated for up to 6 h [10]. Protein-rich meals cause the greatest increase in PYY concentrations compared to other macronutrients [11]. The effects of PYY are thought to be mediated centrally via the G-protein-coupled Y receptors, in particular the Y2 receptor [12], which is densely expressed in the ARC. Peptide tyrosine-tyrosine may also exert its actions via the vagus nerve which also expresses the Y2 receptor [13]. Interestingly, obese individuals have lower fasting PYY$_{3-36}$ concentrations than...
their lean counterparts, suggesting that obesity is a ‘PYY-deficient’ state [10].

Peptide tyrosine-tyrosine changes occur in several gastrointestinal disorders. PYY concentration in tissue extracts from the ileum and colon of patients with Crohn’s colitis and ulcerative colitis has been found to be lower than in controls [14]. Furthermore, rectal and fasting plasma PYY concentrations have been reported to be reduced in patients with ulcerative colitis. This decrease in PYY could contribute to the development of the symptoms seen in inflammatory bowel disease; for example, diarrhoea may be brought about by hindering the ileal brake and loss of inhibition of intestinal transit. Conversely, total PYY concentrations have recently been found to be increased, both pre- and postprandially, in patients with Crohn’s disease affecting the small intestine. It appears that enhanced enteroendocrine cell responses may play a role in feeding disturbance and weight loss [15]. In patients with coeliac disease, basal and postprandial plasma concentrations of PYY are elevated. These elevated concentrations of PYY normalise within 8 months on a gluten-free diet [16]. PYY concentrations in cerebrospinal fluid (CSF) have been found to be elevated in normal-weight bulimic patients abstinent for a month from pathological eating behaviour. However, CSF PYY concentrations were not affected in women who had recovered from anorexia [17]. Plasma concentrations of PYY have not yet been investigated in this group of patients.

1.9.3 Pancreatic polypeptide

Pancreatic polypeptide is an amidated 36-amino acid peptide, structurally similar to PYY. It is released postprandially under vagal control from pancreatic islet PP cells, in proportion to caloric intake. A biphasic response is observed postprandially and concentrations remain elevated for up to 6 h after a meal [18]. Pancreatic polypeptide is thought to mediate its effects via the Y4 receptor, which is highly expressed in the brainstem and ARC [19].

Fasting plasma concentrations of PP are lower in obese individuals, as is the second phase of release after food consumption [20]. Children with Prader–Willi syndrome, a condition characterised by hyperphagia and obesity, have reduced fasting and postprandial concentrations of PP [21]. In contrast, increased concentrations of PP are observed in patients with anorexia nervosa [22]. Pancreatic polypeptide release is reduced in patients with slow transit constipation, but increased in those with functional diarrhoea [23]. Levels are unchanged in patients with coeliac disease [24] but patients with ulcerative colitis have significantly elevated fasting PP concentrations [25].

1.9.4 Glucagon-like peptide 1

Glucagon-like peptide 1 is a product of proglucagon, secreted by enteroendocrine L-cells in the intestine. After a meal, it is released in response to direct L-cell stimulation by nutrients within the GI lumen, and indirectly via neuronal pathways within the enteric nervous system. Glucagon-like peptide 1 is processed intracellularly to generate the 7–37 and 7–36 amide peptides. Glucagon-like peptide 1 is inactivated by dipeptidyl peptidase-4 (DPP-4) which processes GLP-17-37 and GLP-17-36 amide to GLP-19-36 amide.

Glucagon-like peptide 1 binds the G-protein-coupled GLP-1 receptor expressed by pancreatic islet cells as well as brain, heart and lung tissue [26]. The actions of GLP-1 are best characterised in the beta cell where GLP-1 exerts an incretin effect, the stimulation of glucose-dependent insulin release from pancreatic beta cells [27]. Therefore, until recently, the focus on GLP-1 has been largely as an antidiabetic agent and two long-acting analogues of GLP-1, exenatide and liraglutide, are licensed for the treatment of diabetes.

In contrast to insulin, however, GLP-1 causes a decrease in body weight [28]. Acute intravenous injection of GLP-1 reduces appetite and energy intake [29]. This effect has been observed in lean, obese and diabetic volunteers. The actions of GLP-1 on appetite are likely to be related to a direct effect on the central nervous system (CNS) via activation of POMC-expressing neurones in the ARC, as well as through delayed gastric emptying [30].

In gastrointestinal disease, circulating concentrations of GLP-1 can be affected. In both ulcerative colitis and Crohn’s disease, postprandial GLP-1 responses are augmented [31]. In coeliac disease,
concentrations are unaffected [32]. Research into GLP-1 physiology in these inflammatory gastrointestinal conditions is limited and thus far, the emphasis of investigation has been in the field of diabetes and obesity. It is well established that obese subjects have diminished postprandial GLP-1 responses compared to lean controls, but improved secretion is observed after weight loss [33]. The mechanism by which obesity affects GLP-1 secretion is not known but may be related to the insulin resistance that accompanies weight gain, and which also impairs GLP-1 release. The prospect of using long-acting analogues of GLP-1 as a treatment for obesity is eagerly anticipated.

1.9.5 Oxyntomodulin

Also formed by the cleavage of proglucagon, OXM is secreted by the L-cells postprandially, in proportion to energy intake [34]. In addition to reducing appetite, OXM delays gastric emptying and reduces gastric acid secretion [35]. Furthermore, chronic administration of OXM causes rats to lose more weight than pair-fed controls, suggesting that its weight loss effect may be mediated by an increase in energy expenditure [36].

Although OXM is a dual agonist at both the glucagon and GLP-1 receptors, the anorectic effects of OXM are abolished in GLP-1 receptor knockout mice, suggesting that its action on appetite control is primarily via the GLP-1 receptor [37]. OXM concentrations in gastrointestinal disease have not been well studied.

1.9.6 Cholecystokinin

Cholecystokinin is released postprandially from the I-cells of small intestine, and also co-localises with PYY in L-cells. After a meal, CCK is secreted in response to saturated fat, long-chain fatty acids, amino acids and small peptides that would normally result from protein digestion [38]. After lipid ingestion, CCK is released and binds CCK1 receptors, thereby stimulating release of PYY and inhibition of ghrelin [39]. Both of these hormones act to further inhibit food intake. Postprandial secretion of CCK also stimulates pancreatic enzyme secretion and gallbladder contraction, leading to release of bile salts into the duodenum, promoting protein and fat digestion [38]. In addition, CCK is implicated in other gastrointestinal functions, including gastric emptying [40].

Cholecystokinin appears to have a role in the pathogenesis of gallstone disease. Recent studies have shown that postprandial concentrations of CCK are higher in patients with reduced gallbladder contractility, predisposing them to gallstone disease, compared to their healthy counterparts. Interestingly, CCK1 receptor expression in the gallbladders of these patients was lower [41]. This finding is supported by the phenotype of CCK1 receptor knockout mice, in which the prevalence of gallstones is markedly increased [42].

In patients with Crohn’s disease, postprandial CCK concentrations were markedly increased compared to a control group, and patients with ulcerative colitis and diverticulitis. This excessive postprandial release of CCK may be responsible for the delayed gastric emptying observed in these patients [31]. Similarly in Giardia enteritis, elevated postprandial CCK concentrations were correlated with anorectic symptoms upon feeding and treatment led to normalisation of CCK concentrations and symptoms [43].

1.9.7 Ghrelin

Ghrelin is a 28-amino acid peptide that is acylated by the enzyme ghrelin O-acyltransferase (GOAT) and secreted from the gastric fundus. It binds to the growth hormone secretagogue receptor type 1 (GHSR) and stimulates release of growth hormone from the pituitary gland [44]. It is the only true orexigenic gut hormone to have been discovered thus far.

Plasma ghrelin concentrations are highest postprandially, both when meals are provided at scheduled times and when individuals are allowed to eat at will [45]. Fasting ghrelin concentrations are low in obese subjects and chronically high in patients with weight loss due to anorexia nervosa or dietary restriction [46]. In lean people, concentrations decrease after a test meal but do not change in
patients with obesity [47]. There is recent evidence that diet-induced obesity may blunt the orexigenic effects of ghrelin. High-fat feeding renders NPY/AgRP neurones relatively ghrelin resistant [48], and diets high in fat directly inhibit the hyperphagic effect of ghrelin [49]. Furthermore, ghrelin interacts with neurones in the ventral tegmental area of the brain and may provide a link between the GI tract and central control of stress-induced eating of ‘comfort foods’ [50].

Ghrelin seems to be affected in several gastrointestinal diseases. Serum ghrelin concentrations are significantly higher in patients with active ulcerative colitis and Crohn’s disease compared to those in remission or controls. Levels were positively correlated with erythrocyte sedimentation rate and C-reactive protein and negatively correlated with nutritional status parameters [51]. In children with coeliac disease, serum ghrelin was higher than those of controls and negatively correlated with Body Mass Index (BMI). Ghrelin concentrations decreased after 6 months of gluten-free diet compared with the concentrations detected on admission [52]. Helicobacter pylori infection is associated with reduced circulating ghrelin concentrations independent of sex and BMI. Ghrelin concentrations increased, however, 12 weeks after successful H. pylori eradication [53]. In patients with gastric cancer, ghrelin in the gastric mucosa is affected. Gastrectomy decreased the plasma concentration, regardless of the extent of gastric resection [54]. Ghrelin serum concentrations were significantly lower in colon cancer patients compared with controls [55].

1.9.8 Conclusion

In conclusion, the GI tract is now recognised as an endocrine organ that secretes a variety of hormones. These gut hormones play a crucial role in the control of appetite and hence energy homeostasis. Changes in gut hormones concentrations have been identified in several gastrointestinal disorders but the molecular mechanisms by which individual diseases are affected still need to be resolved. Furthermore, the clinical implications of these changes are yet to be elucidated.


30. Willms B, Werner J, Holst JJ, Orskov C, Creutzfeldt W, Nauck MA. Gastric emptying glucose responses, and insulin secretion...


