

## Appendix 1

### Examples of paraphrased question stems and the answer provided by ChatGPT

Although the question stems were paraphrased in this supplement to preserve the intellectual property of the Australian Medical Council, the questions were provided to ChatGPT unedited.

### Answer 1

#### Question

The question describes the clinical presentation of a urinary tract infection. It then asks the most appropriate initial therapy with the following options:

- A. Trimethoprim
- B. Norfloxacin
- C. Amoxicillin
- D. Vancomycin
- E. Gentamicin

#### ChatGPT responds

*A. Trimethoprim is the most appropriate initial therapy for this patient, as it is effective against urinary tract infections caused by Escherichia coli, the most common cause of bacterial urinary tract infections in children. B. Norfloxacin is not typically used to treat urinary tract infections in children due to concerns about safety and efficacy. C. Amoxicillin is not typically effective against the types of bacteria that cause urinary tract infections in children. D. Vancomycin is not typically used to treat urinary tract infections in children, as it is reserved for severe or life-threatening infections caused by resistant bacteria. E. Gentamicin is not typically used to treat urinary tract infections in children, as it is reserved for severe or life-threatening infections caused by resistant bacteria.*

This response was marked 'correct'.

### Answer 2

#### Question

The question describes a teenage girl's increasing pain and tenderness overlying her second metatarsal area. She plays sports and is afebrile. The question provides an image of a nuclear medicine scan with increased tracer

uptake in the second metatarsal. As an image-dependent question, this question was provided to ChatGPT without the scan. The most likely diagnosis out of the following options was requested:

- A. Osteomyelitis
- B. Plantar fasciitis
- C. Stress fracture
- D. Ewing bone tumour
- E. Plantar digital neuroma (Morton metatarsalgia)

#### ChatGPT responds

*Based on the information provided, the most likely diagnosis is stress fracture of the right forefoot. The nuclear bone scan shows increased uptake in the region of the second metatarsal, which is consistent with a stress fracture. Stress fractures are common in athletes and result from repetitive stress on the bone. Treatment typically involves rest and physical therapy to promote healing.*

This response was marked 'correct'.

### Answer 3

#### Question

The question gives a history of a middle-aged woman presenting with multiple painful subcutaneous lumps. It describes the lumps in moderate detail. The question mentions her mother having similar breast lumps. It then asks for the most likely diagnosis out of the following options:

- A. Neurofibromatosis type 1 (Von Recklinghausen disease of nerve)
- B. Adiposis dolorosa (Dercum disease)
- C. Multiple symmetrical subcutaneous lipomas
- D. Multiple desmoid tumours
- E. Multiple epidermoid cysts

#### ChatGPT responds

*The most likely diagnosis for this patient is neurofibromatosis type 1 (Von Recklinghausen disease of nerve). Neurofibromatosis type 1 is a genetic disorder that is characterized by the development of multiple subcutaneous lumps called neurofibromas. These lumps can range in size from small to large and can be located on the trunk, limbs, and other areas of the body. They are often soft and lobulated, and can be mildly*

*tender. Neurofibromas are caused by the growth of nerve tissue, and can sometimes be associated with other abnormalities, such as changes in skin color and bone deformities. The patient's history of multiple lumps that are similar to those her mother had removed, as well as the characteristics of the lumps described on examination, are consistent with a diagnosis of neurofibromatosis type 1. Other potential diagnoses, such as adiposis dolorosa, multiple symmetrical subcutaneous lipomas, multiple desmoid tumors, or multiple epidermoid cysts, are less likely based on the information provided.*

This response was marked 'incorrect'.

On the repeat trial, ChatGPT was asked the same question again and responded as follows:

*The most likely diagnosis in this case is multiple symmetrical subcutaneous lipomas. Lipomas are benign (non-cancerous) tumors that develop from fat cells, and they are often found under the skin. They are typically soft and lobulated, as described in the patient's symptoms, and they can range in size from a few millimeters to several centimeters in diameter.*

*Multiple symmetrical subcutaneous lipomas are characterized by the presence of multiple lipomas on both sides of the body, typically in the upper and lower limbs and on the trunk. This pattern of distribution is consistent with the patient's symptoms. In addition, the patient's history of her mother having similar lumps removed is consistent with a genetic predisposition to developing lipomas.*

*Other potential diagnoses, such as neurofibromatosis type 1, adiposis dolorosa, multiple desmoid tumors, and multiple epidermoid cysts, are less likely given the specific symptoms and characteristics of the lumps.*

This response was marked 'correct'.