## Teacher Answer Key

### Case Study

1. The AI was not retrieving real clinical trial data from a medical database. It was generating text that matched the statistical pattern of what a detailed medical research answer looks like, including realistic-sounding medication names and trial descriptions, even though those specific trials did not exist. The model has no internal fact-checking mechanism to know when its confident-sounding output is fabricated.

2. For high-stakes topics like medicine, law, or finances, users should treat every AI-generated fact as unverified and check all specific claims against authoritative sources such as medical journals, official government health sites, or licensed professionals. AI can be a useful starting point for understanding general concepts, but precise details, especially drug names, legal precedents, or clinical results, must always be independently verified before acting on them.

### Multiple Choice

1. D
2. B
3. C

### Matching

1-B, 2-D, 3-E, 4-A, 5-C

### Fill-in-the-Blank

1) obscure, 2) court case, 3) verify

### Short Answer (Sample Responses)

Q1: LLMs are trained to generate text that matches statistical patterns from training data, and they are rewarded during training for producing confident, detailed, helpful-sounding responses. The model has no internal mechanism to flag when it lacks reliable information, so it generates the most statistically probable response regardless of whether that response is actually accurate.

Q2: Obscure topics with little training data, very recent events past the knowledge cutoff, and questions requiring precise details like exact quotes or specific citations are most likely to cause hallucinations. In these cases, the model has weak statistical grounding and must rely on plausibility patterns rather than actual learned information, dramatically increasing the chance of generating fabricated details.

Q3: Answers will vary. A strong answer might describe a patient using AI for medication dosage information, a student submitting AI-generated legal or scientific citations in a research paper, or an employee using AI to interpret a contract. In each case, the safeguard is verifying the AI's output against authoritative human-produced sources before acting on it.