

Supplement Table 3

All scenario	DeepSeek (1)	ChatGPT (2)	Deep Research (3)	p-value	post hoc test
Total					
Hallucination Score (Mean ± SD)	4.57 (0.81)	1.96 (1.37)	1.33 (1.38)	<0.0001	1>2>3
Hallucination Score (Median [Q1, Q3])	5.00 (4.00 ,5.00)	2.00 (1.00 ,3.00)	1.00 (0.00 ,2.00)	<0.0001	
Hallucination (n (%))	67(95.71)	19(27.14)	11(15.71)	<0.0001	
Preeclampsia					
Hallucination Score (Mean ± SD)	4.70 (0.95)	2.60 (0.52)	1.70 (1.95)	<0.0001	1>2, 1>3
Hallucination Score (Median [Q1, Q3])	5.00 (5.00 ,5.00)	3.00 (2.00 ,3.00)	1.00 (0.00 ,3.00)	0.0007	
Hallucination (n (%))	9(90.00)	6(60.00)	3(30.00)	0.0291	
Cancer drug resistance					
Hallucination Score (Mean ± SD)	4.90 (0.32)	1.00 (1.41)	0.60 (0.84)	<0.0001	1>2, 1>3
Hallucination Score (Median [Q1, Q3])	5.00 (5.00 ,5.00)	0.50 (0.00 ,1.00)	0.00 (0.00 ,1.00)	<0.0001	
Hallucination (n (%))	10(100.00)	2(20.00)	0(0.00)	<0.0001	
Ferroptosis in neurodegenerative disease					
Hallucination Score (Mean ± SD)	4.50 (1.27)	1.70 (0.95)	1.40 (0.97)	<0.0001	1>2, 1>3
Hallucination Score (Median [Q1, Q3])	5.00 (5.00 ,5.00)	1.50 (1.00 ,2.00)	1.50 (1.00 ,2.00)	0.0005	
Hallucination (n (%))	9(90.00)	1(10.00)	1(10.00)	0.0001	
Cuproptosis					
Hallucination Score (Mean ± SD)	5.00 (0.00)	2.50 (1.58)	1.60 (1.07)	<0.0001	1>2, 1>3
Hallucination Score (Median [Q1, Q3])	5.00 (5.00 ,5.00)	2.00 (2.00 ,3.00)	1.50 (1.00 ,2.00)	<0.0001	
Hallucination (n (%))	10(100.00)	4(40.00)	1(10.00)	0.0002	
Tumor in obesity					
Hallucination Score (Mean ± SD)	4.90 (0.32)	1.70 (1.57)	1.30 (1.06)	<0.0001	1>2, 1>3
Hallucination Score (Median [Q1, Q3])	5.00 (5.00 ,5.00)	2.00 (0.00 ,2.00)	1.00 (1.00 ,2.00)	<0.0001	
Hallucination (n (%))	10(100.00)	2(20.00)	2(20.00)	<0.0001	
Lung Cancer treatment					
Hallucination Score (Mean ± SD)	3.80 (0.92)	2.00 (0.47)	2.00 (1.89)	0.0035	1>2, 1>3
Hallucination Score (Median [Q1, Q3])	4.00 (3.00 ,4.00)	2.00 (2.00 ,2.00)	2.00 (0.00 ,3.00)	0.0041	
Hallucination (n (%))	9(90.00)	1(10.00)	3(30.00)	0.0013	
Breast cancer treatment					
Hallucination Score (Mean ± SD)	4.20 (0.42)	2.20 (1.99)	0.70 (1.25)	<0.0001	1>2, 1>3
Hallucination Score (Median [Q1, Q3])	4.00 (4.00 ,4.00)	1.00 (1.00 ,5.00)	0.00 (0.00 ,1.00)	0.0007	
Hallucination (n (%))	10(100.00)	3(30.00)	1(10.00)	<0.0001	