Table S5: Raw Score and Rating of each LLM-Chatbot, Attained for Each Question Concerning Other Ocular Symptoms, Related to Figures 2-3 and Tables 2-4

Domain	Question No.	ChatGPT-3.5			ChatGPT-4.0			Google Bard		
		Score Sum (Acc) Max: 9	Majority Consensus (Acc)	Average Score (Comp) Max: 5	Score Sum (Acc) Max: 9	Majority Consensus (Acc)	Average Score (Comp) Max: 5	Score Sum (Acc) Max: 9	Majority Consensus (Acc)	Average Score (Comp) Max: 5
Red Eye	1	6	Poor*	-	8	Good	5.00	8	Good	5.00
	2	5	Borderline	-	9	Good	4.00	8	Good	4.50
Painful Eye	1	8	Good	5.00	8	Good	5.00	8	Good	4.50
	2	7	Borderline	-	8	Good	5.00	6	Poor*	-
Pruritus	1	9	Good	4.33	8	Good	5.00	9	Good	4.67
	2	9	Good	4.67	9	Good	4.67	9	Good	4.67
Pressure Sensation	1	8	Good	4.50	8	Good	5.00	5	Borderline	-
	2	9	Good	4.33	8	Good	4.50	6	Poor*	-
Epiphora	1	8	Good	5.00	8	Good	5.00	9	Good	4.67
	2	8	Good	5.00	9	Good	4.67	8	Good	5.00

Ptosis	1	6	Borderline	-	9	Good	4.67	5	Borderline	-
	2	6	Borderline	-	8	Good	5.00	5	Borderline	-
Blepharospasm	1	9	Good	4.33	8	Good	4.50	7	Borderline	-
	2	9	Good	4.67	9	Good	5.00	8	Good	4.50
Foreign Body Sensation	1	8	Good	5.00	9	Good	4.67	7	Borderline	-
	2	8	Good	4.00	9	Good	4.67	8	Good	4.50
Trauma	1	8	Good	4.50	8	Good	4.00	5	Borderline	-
Asthenopia	2	7	Borderline	-	8	Good	4.50	9	Good	4.67
	1	8	Good	5.00	9	Good	4.33	9	Good	4.67

Where consensus on final accuracy rating was not reached (i.e., each grader provided a different rating), the lowest score ('poor'*) was assigned.