

glaucoma	134	0	0	0	1	0	0	0
jaundice	0	98	0	0	1	0	1	0
cyanosis	0	1	112	0	0	1	1	0
psoriasis	1	4	0	110	5	0	7	0
conjunctivitis	13	4	0	0	119	0	0	0
scoliosis	3	0	4	0	0	120	0	0
skin cancer	0	1	0	0	2	0	142	0
gingivitis	0	1	0	1	0	0	0	115
	glaucoma	jaundice	cyanosis	psoriasis	conjunctivitis	scoliosis	skin cancer	gingivitis

Robustness Test 1: gpt-3.5-turbo has generated the query dataset, text-embedding-ada-002 is the embedding model, and gpt-3.5-turbo has generated the ground truth dataset which is in the vector database

glaucoma	110	0	0	0	11	0	0	0
jaundice	5	128	14	3	3	0	2	2
cyanosis	5	11	107	0	1	0	0	0
psoriasis	1	15	0	107	6	0	0	0
conjunctivitis	1	4	0	2	128	0	0	0
scoliosis	6	3	4	2	0	101	0	0
skin cancer	0	0	0	1	0	0	99	0
gingivitis	0	3	0	0	8	0	0	119
	glaucoma	jaundice	cyanosis	psoriasis	conjunctivitis	scoliosis	skin cancer	gingivitis

Robustness Test 2: LLaMA 2 70b-chat has generated the query dataset, text-embedding-ada-002 is the embedding model, and gpt-3.5-turbo has generated the ground truth dataset which is in the vector database

glaucoma	188	0	0	0	12	0	0	0
jaundice	0	200	0	0	0	0	0	0
cyanosis	0	0	200	0	0	0	0	0
psoriasis	0	0	0	200	0	0	0	0
conjunctivitis	0	0	0	0	200	0	0	0
scoliosis	0	0	0	0	0	200	0	0
skin cancer	0	0	0	6	28	0	166	0
gingivitis	0	0	0	0	12	0	0	188
	glaucoma	jaundice	cyanosis	psoriasis	conjunctivitis	scoliosis	skin cancer	gingivitis

Robustness Test 3: google-flan-t5-xl has generated the query dataset, text-embedding-ada-002 is the embedding model, and gpt-3.5-turbo has generated the ground truth dataset which is in the vector database

glaucoma	132	0	0	0	3	0	0	0
jaundice	0	96	0	3	0	0	1	0
cyanosis	0	0	111	1	0	2	1	0
psoriasis	0	1	0	116	4	0	2	4
conjunctivitis	17	0	0	1	113	0	2	3
scoliosis	2	0	7	0	0	118	0	0
skin cancer	0	0	0	4	0	0	136	5
gingivitis	0	0	0	2	0	0	0	115
	glaucoma	jaundice	cyanosis	psoriasis	conjunctivitis	scoliosis	skin cancer	gingivitis

Robustness Test 4: gpt-3.5-turbo has generated the query dataset, text-embedding-ada-002 is the embedding model, and LLaMA 2 70b-chat has generated the ground truth dataset which is in the vector database

glaucoma	110	0	0	0	11	0	0	0
jaundice	5	119	12	1	7	0	7	6
cyanosis	11	4	106	2	0	1	0	0
psoriasis	0	1	0	111	10	2	0	5
conjunctivitis	3	0	0	0	132	0	0	0
scoliosis	1	0	3	1	0	109	1	1
skin cancer	0	0	0	0	0	0	100	0
gingivitis	0	1	1	1	0	0	0	127
	glaucoma	jaundice	cyanosis	psoriasis	conjunctivitis	scoliosis	skin cancer	gingivitis

Robustness Test 5: LLaMA 2 70b-chat has generated the query dataset, text-embedding-ada-002 is the embedding model, and LLaMA 2 70b-chat has generated the ground truth dataset which is in the vector database

glaucoma	188	0	0	0	12	0	0	0
jaundice	0	200	0	0	0	0	0	0
cyanosis	0	0	200	0	0	0	0	0
psoriasis	0	0	0	200	0	0	0	0
conjunctivitis	0	0	0	0	200	0	0	0
scoliosis	0	0	0	0	0	200	0	0
skin cancer	0	0	0	25	21	0	153	1
gingivitis	0	0	0	0	0	0	0	200
	glaucoma	jaundice	cyanosis	psoriasis	conjunctivitis	scoliosis	skin cancer	gingivitis

Robustness Test 6: google-flan-t5-xl has generated the query dataset, text-embedding-ada-002 is the embedding model, and LLaMA 2 70b-chat has generated the ground truth dataset which is in the vector database

glaucoma	134	0	0	0	1	0	0	0
jaundice	20	70	0	9	0	0	1	0
cyanosis	6	5	65	21	2	0	16	0
psoriasis	1	2	0	123	0	0	1	0
conjunctivitis	77	0	0	8	51	0	0	0
scoliosis	20	4	0	1	0	98	4	0
skin cancer	1	7	0	102	0	0	35	0
gingivitis	2	0	0	4	0	0	0	111
	glaucoma	jaundice	cyanosis	psoriasis	conjunctivitis	scoliosis	skin cancer	gingivitis

Robustness Test 7: gpt-3.5-turbo has generated the query dataset, text-embedding-ada-002 is the embedding model, and google-flan-t5-xl has generated the ground truth dataset which is in the vector database

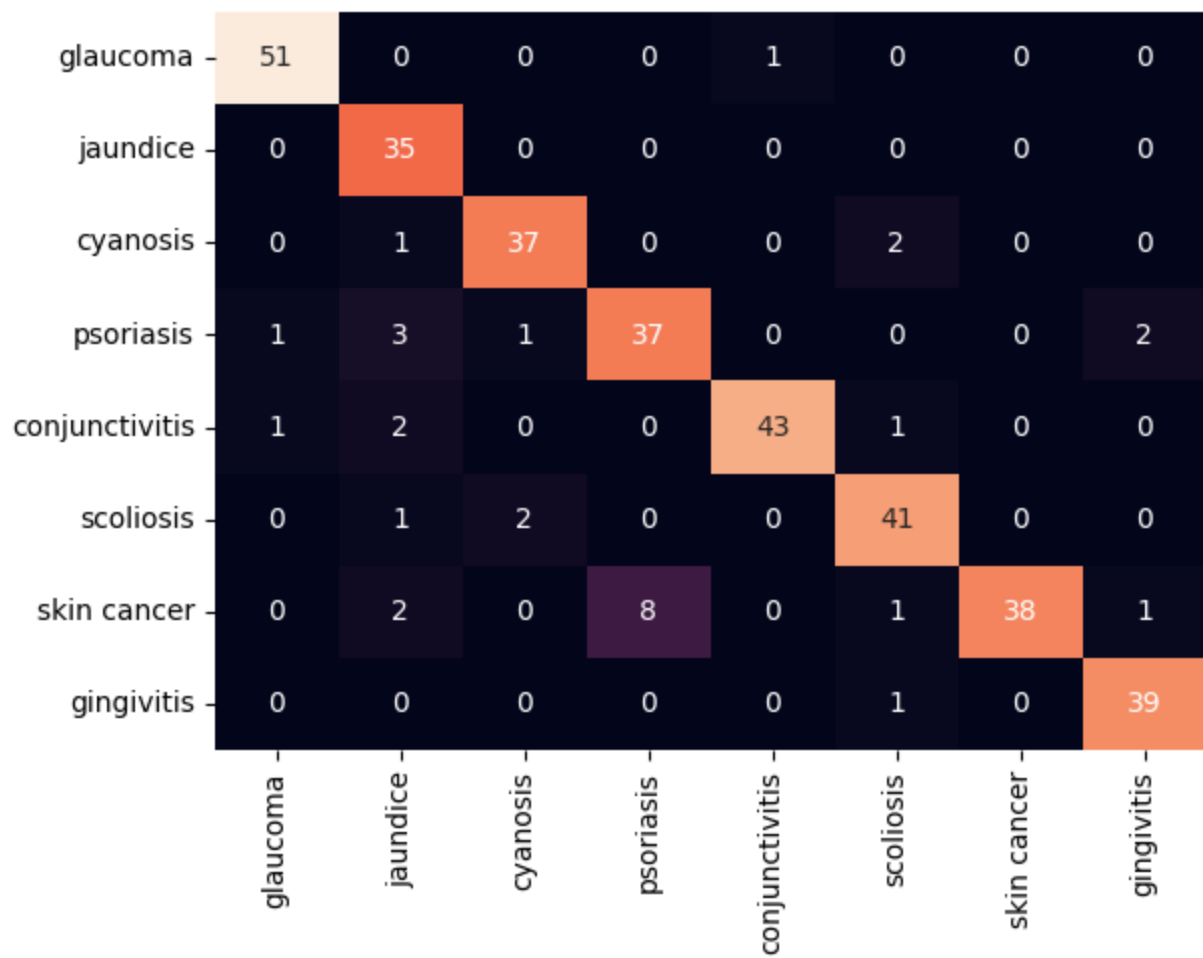
glaucoma	120	0	0	0	1	0	0	0
jaundice	53	70	0	11	2	0	21	0
cyanosis	20	0	35	10	1	0	58	0
psoriasis	11	0	0	116	0	0	2	0
conjunctivitis	63	0	0	1	71	0	0	0
scoliosis	28	0	0	0	0	70	18	0
skin cancer	0	0	0	7	0	0	93	0
gingivitis	7	0	0	9	1	0	5	108
	glaucoma	jaundice	cyanosis	psoriasis	conjunctivitis	scoliosis	skin cancer	gingivitis

Robustness Test 8: LLaMA 2 70b-chat has generated the query dataset, text-embedding-ada-002 is the embedding model, and google-flan-t5-xl has generated the ground truth dataset which is in the vector database



glaucoma	177	3	2	0	18	0	0	0
jaundice	0	200	0	0	0	0	0	0
cyanosis	0	0	200	0	0	0	0	0
psoriasis	0	0	0	200	0	0	0	0
conjunctivitis	0	0	0	0	200	0	0	0
scoliosis	0	0	0	0	0	200	0	0
skin cancer	0	0	0	67	16	0	117	0
gingivitis	0	0	0	15	3	0	0	182
	glaucoma	jaundice	cyanosis	psoriasis	conjunctivitis	scoliosis	skin cancer	gingivitis

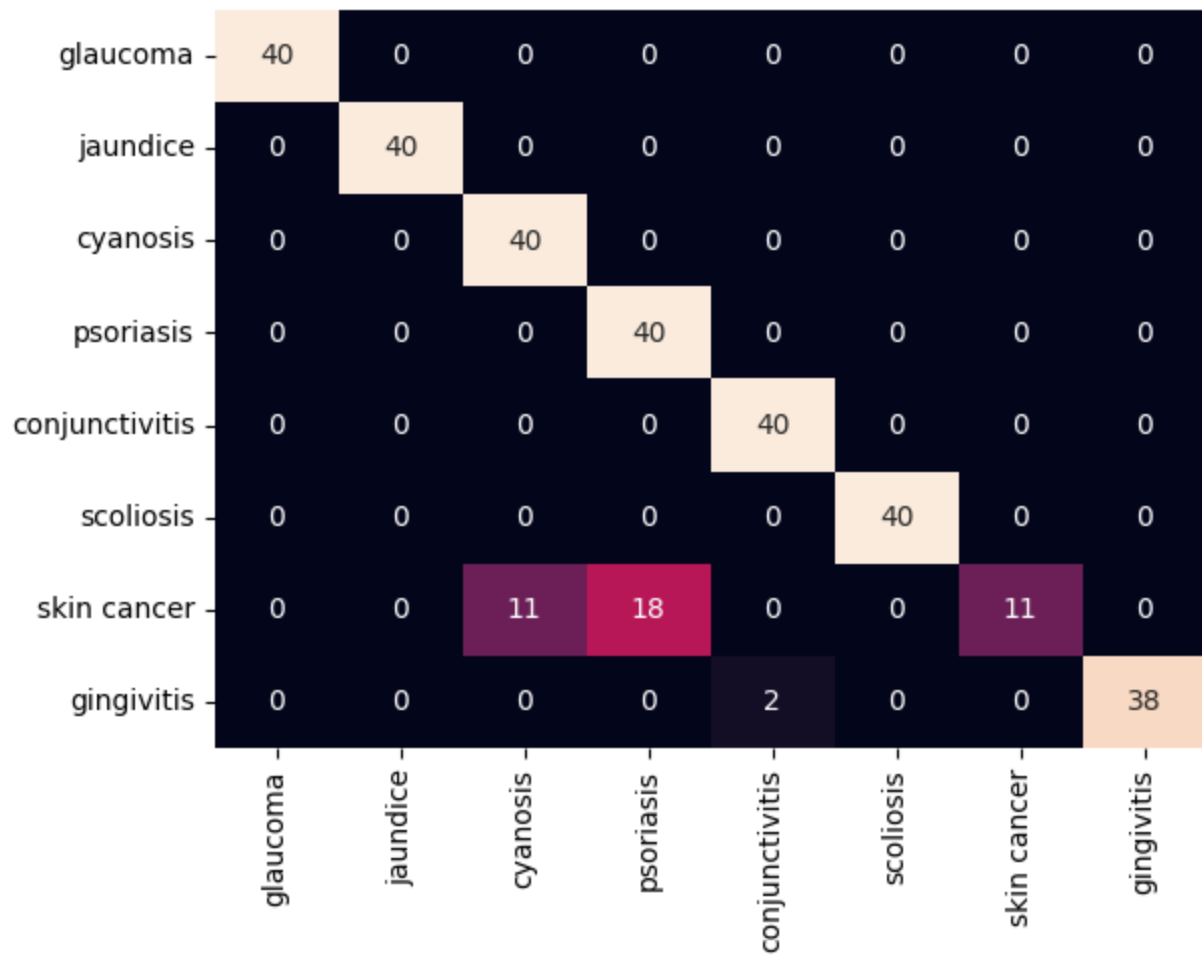
Robustness Test 9: google-flan-t5-xl has generated the query dataset, text-embedding-ada-002 is the embedding model, and google-flan-t5-xl has generated the ground truth dataset which is in the vector database



Robustness Test 10: gpt-3.5-turbo has generated the query dataset, textembedding-gecko@001 is the embedding model, and gpt-3.5-turbo has generated the ground truth dataset which is in the vector database

glaucoma	39	0	0	0	2	0	0	0
jaundice	0	36	5	3	9	0	0	0
cyanosis	0	1	41	0	0	0	0	0
psoriasis	0	7	0	32	3	1	0	0
conjunctivitis	3	1	0	1	40	0	0	0
scoliosis	1	0	12	0	0	26	0	0
skin cancer	0	1	0	3	0	0	30	0
gingivitis	0	3	2	0	3	0	0	36
	glaucoma	jaundice	cyanosis	psoriasis	conjunctivitis	scoliosis	skin cancer	gingivitis

Robustness Test 11: LLaMA 2 70b-chat has generated the query dataset, textembedding-gecko@001 is the embedding model, and gpt-3.5-turbo has generated the ground truth dataset which is in the vector database



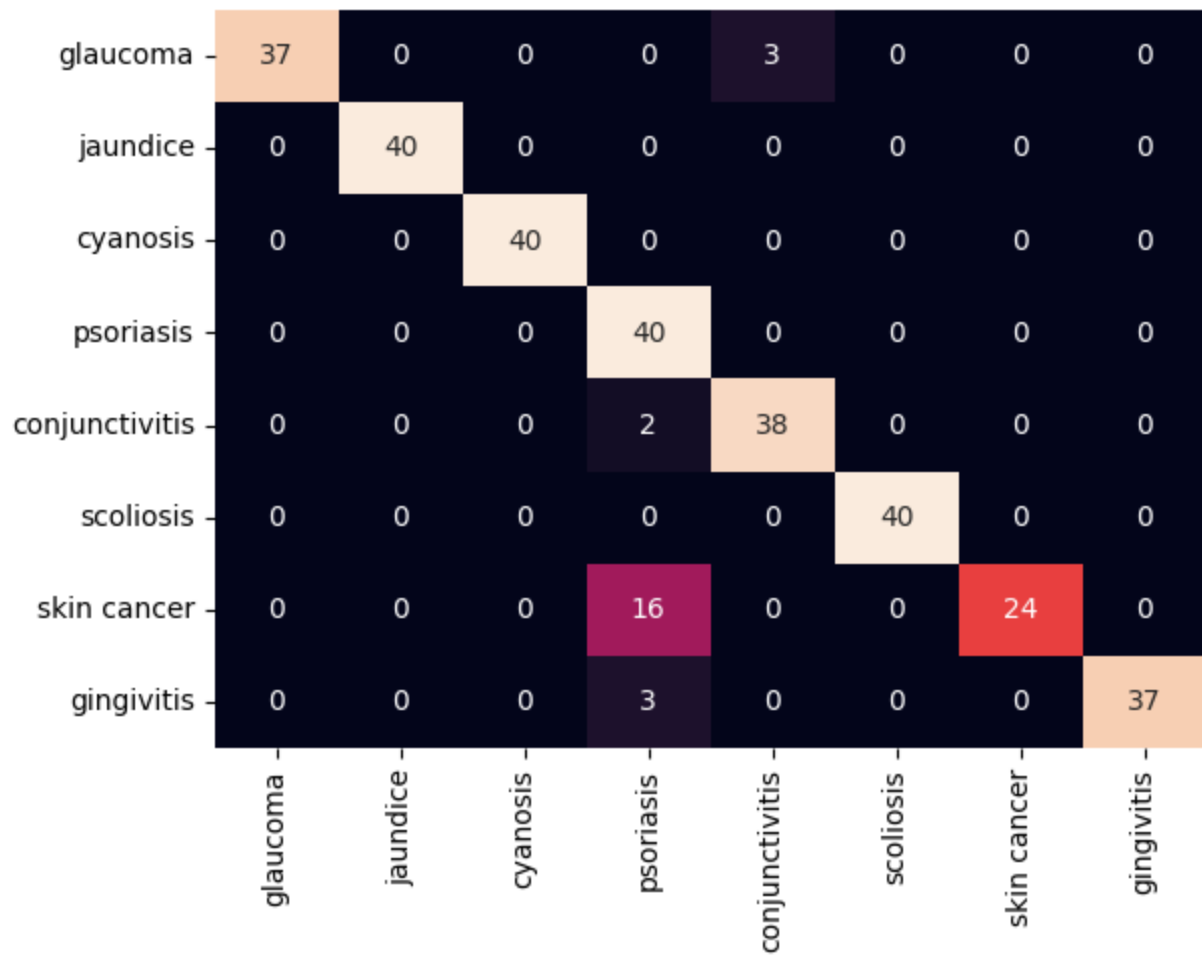
Robustness Test 12: google-flan-t5-xl has generated the query dataset, textembedding-gecko@001 is the embedding model, and gpt-3.5-turbo has generated the ground truth dataset which is in the vector database

glaucoma	44	0	0	0	1	0	0	0
jaundice	0	34	0	0	0	0	0	0
cyanosis	0	2	37	0	0	0	0	0
psoriasis	0	3	1	34	3	0	1	1
conjunctivitis	10	2	2	0	32	0	0	0
scoliosis	0	0	12	0	0	31	0	0
skin cancer	0	0	1	0	0	0	48	0
gingivitis	0	0	1	0	0	0	0	38
	glaucoma	jaundice	cyanosis	psoriasis	conjunctivitis	scoliosis	skin cancer	gingivitis

Robustness Test 13: gpt-3.5-turbo has generated the query dataset, textembedding-gecko@001 is the embedding model, and LLaMA 2 70b-chat has generated the ground truth dataset which is in the vector database

glaucoma	39	0	0	0	2	0	0	0
jaundice	4	36	2	1	7	0	0	3
cyanosis	0	3	39	0	0	0	0	0
psoriasis	0	7	0	33	0	0	0	3
conjunctivitis	3	1	0	0	40	1	0	0
scoliosis	0	1	9	0	0	29	0	0
skin cancer	0	0	0	0	0	0	34	0
gingivitis	0	4	0	0	2	0	0	38
	glaucoma	jaundice	cyanosis	psoriasis	conjunctivitis	scoliosis	skin cancer	gingivitis

Robustness Test 14: LLaMA 2 70b-chat has generated the query dataset, textembedding-gecko@001 is the embedding model, and LLaMA 2 70b-chat has generated the ground truth dataset which is in the vector database



Robustness Test 15: google-flan-t5-xl has generated the query dataset, textembedding-gecko@001 is the embedding model, and LLaMA 2 70b-chat has generated the ground truth dataset which is in the vector database

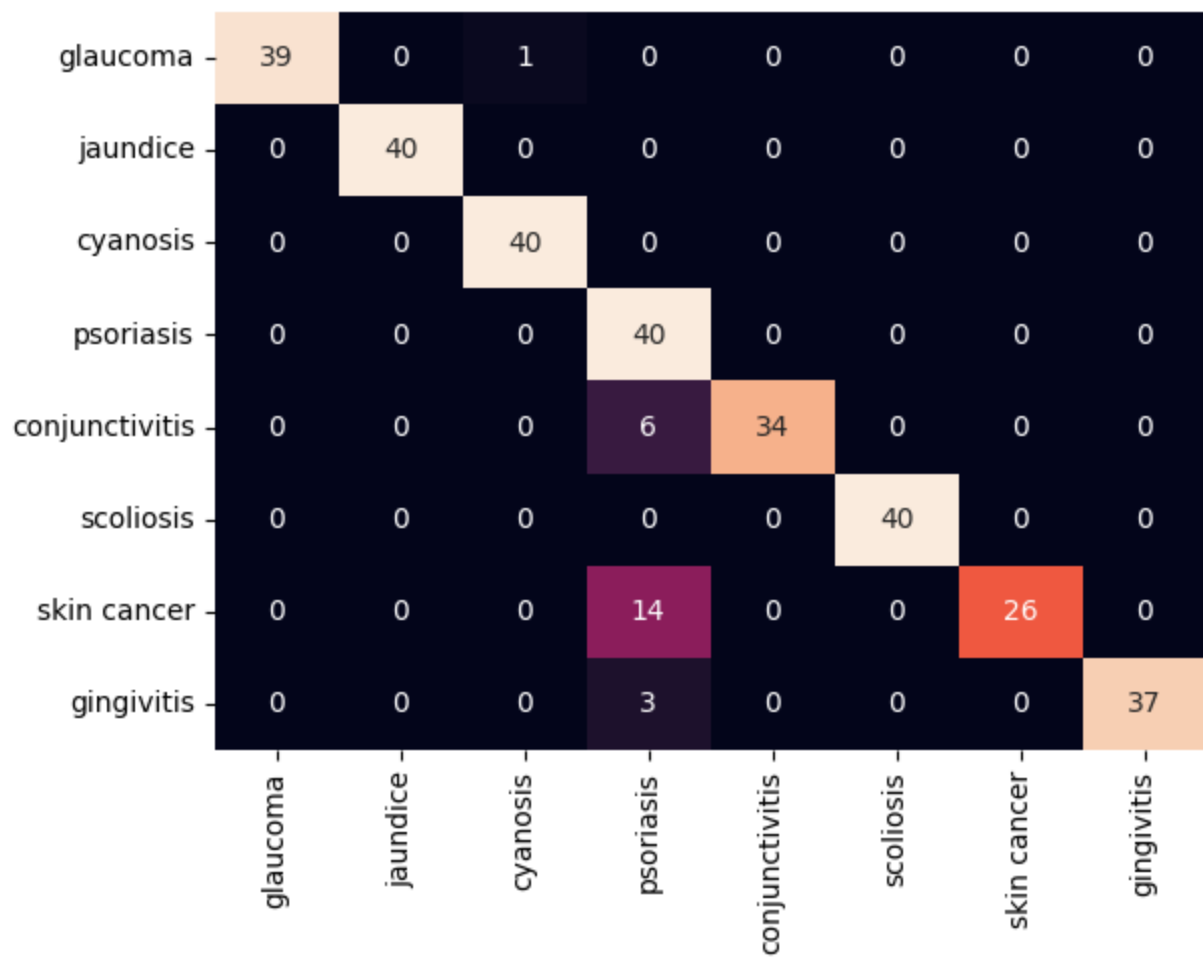
glaucoma	45	0	0	0	0	0	0	0
jaundice	4	28	1	1	0	0	0	0
cyanosis	0	2	32	5	0	0	0	0
psoriasis	1	1	2	38	0	0	1	0
conjunctivitis	33	1	2	0	10	0	0	0
scoliosis	3	0	6	5	0	29	0	0
skin cancer	1	0	3	32	0	0	13	0
gingivitis	0	0	1	0	0	0	0	38
	glaucoma	jaundice	cyanosis	psoriasis	conjunctivitis	scoliosis	skin cancer	gingivitis

Robustness Test 16: gpt-3.5-turbo has generated the query dataset, textembedding-gecko@001 is the embedding model, and google-flan-t5-xl has generated the ground truth dataset which is in the vector database



glaucoma	40	0	0	0	1	0	0	0
jaundice	6	28	0	17	2	0	0	0
cyanosis	1	5	15	21	0	0	0	0
psoriasis	1	0	0	42	0	0	0	0
conjunctivitis	29	0	0	2	13	0	0	1
scoliosis	3	0	6	5	0	25	0	0
skin cancer	0	0	0	14	0	0	20	0
gingivitis	0	0	0	7	1	0	0	36
	glaucoma	jaundice	cyanosis	psoriasis	conjunctivitis	scoliosis	skin cancer	gingivitis

Robustness Test 17: LLaMA 2 70b-chat has generated the query dataset, textembedding-gecko@001 is the embedding model, and google-flan-t5-xl has generated the ground truth dataset which is in the vector database



Robustness Test 18: google-flan-t5-xl has generated the query dataset, textembedding-gecko@001 is the embedding model, and google-flan-t5-xl has generated the ground truth dataset which is in the vector database